JNEST EISSN: 2961-8738 ● PISSN: 2961-8916 ● DOI: 10.56741/jnest.v5i03.2259 250
Article Received 21 April 2026, Revised 19 August 2026

Explainable Machine Learning Framework for Predicting and Optimizing Fly Ash Slag Concrete Mixture Composition Based on SHAP Analysis and Multi-Algorithm Benchmarking

Ichwan Hadi Saputra*, Eko Wahyu Abryandoko, Widyo Nugroho, Moh. Nurudduja, Nayla Farikha Zahra

***Abstract*—Optimizing fly ash slag concrete mixtures is a complex material-engineering problem because mixture variables interact non-linearly, making conventional single variable approaches inadequate. This study develops an interpretable prediction-and-optimization framework comprising three contributions: (1) benchmarking five machine-learning algorithms Support Vector Regression (SVR), Artificial Neural Network (ANN), Random Forest, XGBoost, and LightGBM using hyperparameters optimized through Bayesian optimization (Optuna/TPE); (2) applying SHAP analysis to identify the contribution of each mixture variable; and (3) using a Genetic Algorithm (GA) to optimize concrete mixtures for three strength classes based on the best performing surrogate model. Starting from the 1,030-sample UCI concrete strength benchmark dataset, records were screened according to three Indonesian standards (SNI 2847:2019, SNI 2461:2011, and SNI 8297:2016), resulting in 542 samples that satisfied the specified technical limits. Data augmentation was evaluated using six noise levels, with 2% identified as optimal, producing a final dataset of 1,074 samples. Five-fold cross-validation demonstrated that SVR achieved the highest predictive performance (C=193.39, gamma=0.658, epsilon=0.146), with R^2=0.939±0.020 and RMSE=3.30±0.45 MPa, followed by XGBoost (R^2=0.934), LightGBM (R^2=0.931), ANN (R^2=0.925), and Random Forest (R^2=0.920). SHAP analysis identified cement, age, and superplasticizer as the most influential variables, consistent with cement hydration mechanisms and Abrams' law. GA optimization produced three mix designs corresponding to target strengths of 15.00, 27.50, and 54.99 MPa for Classes I-III, respectively. The proposed framework provides an accurate, interpretable, and standards-compliant approach for preliminary sustainable concrete mixture design across different structural applications.**

***Index Terms*—Concrete Compressive Strength, Fly Ash, Genetic Algorithm, SHAP, Support Vector Regression.**

I. INTRODUCTION

Sustainable concrete development requires a balance between environmental considerations and the mechanical performance demanded by construction applications. One approach that has received considerable attention is the partial replacement of Portland cement with supplementary materials derived from industrial by-products, including fly ash and blast furnace slag. The use of these materials can reduce the amount of Portland cement required in a mixture and consequently contribute to lowering the environmental burden associated with concrete production [1]–[2]. However, determining suitable proportions remains challenging because concrete strength is governed by the combined effects of binder composition, water content, aggregate proportions, admixtures, and curing conditions. Conventional trial-and-error mixture design can therefore become time-consuming, material-intensive, and difficult to reproduce consistently when several variables interact simultaneously [3].

Machine learning has become an increasingly useful tool for estimating concrete compressive strength because it can learn nonlinear relationships between mixture characteristics and measured mechanical performance. Previous studies have applied approaches such as Artificial Neural Networks (ANN), Support Vector Machines or Support Vector Regression (SVM/SVR), Random Forest, and gradient-boosting algorithms to concrete-strength prediction [4]–[9]. Ensemble-based models have demonstrated strong predictive capability, while ANN-based methods have also been reported to capture complicated interactions among mixture constituents effectively. Similar applications to geopolymer and waste-derived concrete further demonstrate that machine-learning models can represent relationships between constituent materials and mechanical properties that are difficult to formulate using conventional empirical equations [10]–[11].

Despite the progress achieved in machine-learning-based concrete prediction, several methodological gaps remain. A considerable number of previous investigations evaluate only a small selection of learning algorithms, limiting the extent to which their predictive performance can be compared with alternative ensemble techniques such as XGBoost and LightGBM [12]–[13]. In addition, hyperparameter selection is frequently conducted through manual experimentation, which may not efficiently explore the available parameter space and may produce results that are difficult to reproduce. Another important limitation concerns model interpretability. A highly accurate prediction does not necessarily explain the contribution of individual mixture variables or the reason for a particular prediction. This issue becomes particularly relevant in engineering applications, where model transparency is important for examining whether computational results remain compatible with established material behavior [14]–[15].

SHAP (SHapley Additive exPlanations) provides a framework for examining the contribution of individual

Ichwan Hadi Saputra (Corresponding Author) and **Widyo Nugroho** are with the Department of Civil Engineering; **Eko Wahyu Abryandoko** and **Moh. Nurudduja** and **Nayla Farikha Zahra** are with the Department of Industrial Engineering, Universitas Bojonegoro, Bojonegoro, Indonesia. Corresponding Author: *ichwanhs@gmail.com

© 2026 IISTR

input variables to machine-learning predictions. The method is derived from the Shapley-value concept in cooperative game theory and assigns a contribution value to each feature according to its effect on the model output. In material-property prediction, SHAP can therefore be used not only to establish a global ranking of influential variables but also to investigate the direction and magnitude of their contributions for individual observations. Previous applications to modified concrete systems indicate that this type of analysis can expose response patterns that can subsequently be interpreted in relation to known engineering mechanisms [14]–[15].

Concrete mixture formulation involves several competing considerations, including the required strength level, material proportions, durability, constructability, and environmental impact. Because the feasible design space may contain nonlinear relationships and interacting constraints, evolutionary search methods provide an alternative to conventional sequential trial-and-error approaches. Genetic Algorithms (GA) are particularly suitable for this type of problem because candidate solutions can be evolved through population-based selection, crossover, and mutation operations [6], [16]–[17]. In the present study, the optimization search is restricted by the technical requirements adopted from SNI 2847:2019, SNI 2461:2011, and SNI 8297:2016, allowing the generated candidate mixtures to remain within the composition limits defined for the analysis.

Building on the ANN- and GA-based prediction-optimization framework developed in previous research, this study aims to extend and strengthen that framework through three methodological contributions: (1) comprehensive benchmarking of five prediction algorithms using systematic and reproducible Bayesian-optimized hyperparameters, addressing the need for cross-algorithm validation that has rarely been performed thoroughly in similar studies; (2) SHAP analysis to reveal the relative contribution of composition variables to compressive strength prediction, providing interpretability insights that structural engineers can directly use to understand and verify the rationality of model decisions; and (3) determination of the synthetic data augmentation level based on quantitative sensitivity analysis, rather than assumption-based determination. Together, these three contributions produce a more accurate, transparent, and scientifically accountable prediction-optimization framework to support the design of fly ash slag concrete mixtures that comply with national standards.

II. Material and Method

A. Data Collection and Preprocessing

The experimental database used in this research is Yeh's Concrete Compressive Strength dataset obtained from the UCI Machine Learning Repository [18]. It consists of 1,030 laboratory observations, with each observation containing eight explanatory variables and one response variable. The predictors comprise cement, slag, fly ash, water, superplasticizer, coarse aggregate, fine aggregate, and curing age, whereas compressive strength in MPa is used as the target variable. The combination of mixture composition and curing-age information provides the basis for modeling the relationship between concrete formulation and its measured compressive performance.

Before the prediction models were trained, the original observations underwent a standards-based screening process. The filtering criteria were established with reference to SNI 2847:2019 for structural concrete, SNI 2461:2011 for aggregate requirements, and SNI 8297:2016 for geopolymer concrete. The retained mixtures were required to have a water-to-binder ratio between 0.3 and 0.8, a fly-ash fraction of no more than 30% of the total binder mass, and a slag fraction not exceeding 70% of the binder mass. Application of these criteria reduced the original database to 542 valid observations, which were subsequently used as the basis for data augmentation and model development. Descriptive statistics for the eight input variables and the target variable in the filtered dataset are presented in Table 1, which provides an overview of the technical value ranges and data variability underlying the subsequent modeling stage.

TABLE I. DESCRIPTIVE STATISTICS OF THE DATASET AFTER SNI FILTERING

Variable	Unit	Min	Max	Mean	Std
Cement	kg/m³	140.00	500.00	290.03	88.68
Slag	kg/m³	0.00	342.10	83.82	90.34
Fly Ash	kg/m³	0.00	148.10	31.79	49.05
Water	kg/m³	145.00	247.00	190.39	18.76
Superplasticizer	kg/m³	0.00	20.00	4.37	4.97
Coarse Aggregate	kg/m³	801.00	1145.00	963.80	74.77
Fine Aggregate	kg/m³	594.00	945.00	759.56	78.94
Age	days	7.00	365.00	56.25	75.93
Compressive Strength	MPa	8.00	81.75	36.81	13.87

Table 1 shows a wide range of values for cement (140 500 kg/m³) and slag (0 342.1 kg/m³), indicating representative composition variation across different concrete quality classes. The high standard deviation of age (75.93 days) reflects a broad distribution of testing ages, ranging from 7 to 365 days, so that models trained on this dataset are likely to capture cement hydration dynamics across a wide curing time range a data characteristic relevant to the non-linear pattern of the age variable revealed through the SHAP analysis in the Results and Discussion section.

B. Concrete Quality Classification

For subsequent optimization and engineering interpretation, the measured compressive-strength values were divided into three predefined quality categories. Class I represents strengths below 20 MPa and is associated with applications having relatively low structural demand. Class II covers the interval from 20 to 35 MPa and represents commonly encountered residential-scale structural applications. Class III includes strengths greater than 35 MPa and corresponds to applications requiring a higher level of load-carrying capacity. The three categories provide target ranges that can be used to evaluate the candidate mixtures generated by the optimization procedure, as summarized in Table 2.

TABLE II. CONCRETE QUALITY CATEGORIES BASED ON COMPRESSIVE STRENGTH

Class	Main Usage Description	Concrete Strength (fc')	Category
I	Applications with minimal structural demand walkways, unpaved-road overlays, and other lightly loaded elements not intended to carry significant load	< 20 MPa	Low
II	Typical residential-scale construction, including slab-on-grade floors, modest beam sections, and columns in low-rise housing	20 35 MPa	Medium
III	Structural elements requiring higher load capacity tall building frames, bridge components, and construction in earthquake-prone regions	> 35 MPa	High

C. Sensitivity Analysis of the Data Augmentation Level

To enrich the variation of training data without distorting the physical validity of the mixture composition, synthetic data augmentation was applied in the form of proportional Gaussian noise added to the original numerical values. To determine the optimal noise level in a data-driven manner, a sensitivity analysis was conducted across six noise levels (0%, 2%, 5%, 10%, 15%, 20%) using an SVR probe model with a consistent 5-fold cross-validation evaluation scheme. Each augmented sample was re-validated against the SNI technical limits so that the resulting synthetic data remained within a physically realistic composition range and did not violate concrete engineering technical constraints. The results of the sensitivity analysis are presented in Table 3.

TABLE III. SENSITIVITY ANALYSIS RESULTS FOR THE NOISE AUGMENTATION LEVEL

Noise Level	Number of Samples	R^2 (mean±std)	RMSE (MPa)
0%	542	0.825 ± 0.053	5.64 ± 0.83
2%	1,074	0.881 ± 0.030	4.65 ± 0.47
5%	1,062	0.840 ± 0.049	5.38 ± 0.77
10%	1,034	0.750 ± 0.034	6.79 ± 0.52
15%	1,002	0.664 ± 0.045	7.92 ± 0.49
20%	969	0.617 ± 0.030	8.54 ± 0.44

The results in Table 3 show that the best performance is achieved at a 2% noise level (R^2=0.881±0.030), higher than both no augmentation and higher noise levels. The consistent performance decline at noise levels above 5% indicates that excessive augmentation distorts the physical relationships among composition variables and introduces prediction bias, so the 2% level was chosen as the optimal point between adding data variation and preserving the validity of the concrete mixture composition. The final dataset used in the modeling stage consists of 1,074 samples, comprising 542 original samples and 532 augmented samples that have been re-validated against the SNI limits.

D. Development and Tuning of Prediction Models

Five different learning approaches were considered in the comparative modeling stage: Support Vector Regression (SVR) with an RBF kernel, a multilayer-perceptron Artificial Neural Network (ANN), Random Forest [19], XGBoost [20] and LightGBM [21]. The selected methods represent distinct learning paradigms. SVR provides a margin-based regression approach, ANN represents nonlinear neural-network learning, and Random Forest, XGBoost, and LightGBM represent tree-based ensemble strategies. Evaluating these different model families under the same dataset and validation procedure provides a more consistent basis for determining their relative predictive capability.

Hyperparameter configuration was performed using Bayesian optimization through the Tree-structured Parzen Estimator (TPE) sampler available in Optuna [22]–[23]. The optimization process uses information obtained from previous trials to identify promising regions of the parameter space, rather than evaluating all combinations through an exhaustive grid. The average RMSE obtained from the internal validation process was used as the optimization criterion. The resulting parameter configurations for each prediction algorithm are provided in Table 4.

TABLE IV. BEST HYPERPARAMETERS FROM BAYESIAN TUNING (OPTUNA/TPE) FOR THE FIVE MODELS

Model	Best Hyperparameters
SVR	Kernel=RBF; C=193.39; gamma=0.658; epsilon=0.146
Random Forest	n_estimators=420; max_depth=17; min_samples_split=2; min_samples_leaf=1
XGBoost	n_estimators=416; max_depth=6; learning_rate=0.059; subsample=0.588; colsample_bytree=0.975
LightGBM	n_estimators=636; num_leaves=63; max_depth=8; learning_rate=0.189; subsample=0.945
ANN	64-128-32 neuron architecture (ReLU); dropout=0.084; learning_rate=0.00122

All models were evaluated using an identical 5-fold cross-validation scheme: in each iteration, four data subsets (80%) were used for training and one subset (20%) for testing, so that each sample was used exactly once as test data across all iterations. This uniform evaluation scheme ensures a fair performance comparison among models, with evaluation metrics including the coefficient of determination (R^2), root mean squared error (RMSE), mean absolute error (MAE), and mean absolute percentage error (MAPE) The use of multiple complementary error and fitness indicators provides a more comprehensive assessment of predictive performance than relying on a single metric [24].

E. Explainability Analysis (SHAP)

Model interpretation was conducted using SHAP (SHapley Additive exPlanations), which provides feature-level attribution based on the Shapley-value framework [25]. For each prediction, SHAP decomposes the model output into contributions associated with the individual input variables relative to a baseline prediction. This makes it possible to distinguish variables that increase the predicted compressive strength from those that contribute negatively for a particular observation.

SHAP analysis was applied to the model with the best cross-validation performance from the benchmarking stage. Three forms of analysis were produced: (1) a global feature ranking based on the mean absolute SHAP value, ranking the eight composition variables by the average magnitude of their influence on predictions across all samples; (2) a summary plot visualizing the distribution of

SHAP values for each variable along with the direction of influence (positive or negative) based on the variable's value level; and (3) dependence plots for key variables revealing non-linear relationship patterns and potential interactions among composition variables. These three forms of analysis are presented and discussed in detail in the Results and Discussion section, with interpretations directly linked to civil engineering principles concerning cement hydration processes and the pozzolanic reaction mechanisms of supplementary materials.

F. Genetic Algorithm Optimization for Mixture Design

The final stage of the framework is mixture composition optimization using a Genetic Algorithm (GA), aimed at producing mixture designs representative of the three concrete quality classes (Class I, II, and III) as classified in Table 2, rather than merely maximizing compressive strength in absolute terms. An initial population of 30 individuals was randomly generated within the limits of SNI 2847:2019, SNI 2461:2011, and SNI 8297:2016, then evolved over 50 generations using blend crossover (BLX-α) and Gaussian mutation operations. Each individual in the population represents one candidate mixture composition (cement, slag, fly ash, water, superplasticizer, coarse aggregate, fine aggregate, and age content), whose fitness is evaluated using the best prediction model from the benchmarking stage as a surrogate function unlike conventional approaches that fix the surrogate model a priori, the model selection in this study was based on empirical cross-validation performance evidence, so the GA objective function represents the most accurate prediction model available.

The fitness function was designed based on per-class targets: for each class, the GA is directed to produce a composition with a predicted compressive strength as close as possible to the midpoint of that class's range (Class I: target 15 MPa; Class II: target 27.5 MPa; Class III: target 55 MPa), while remaining within the targeted class range. Solutions falling outside the class range are penalized proportionally to their distance from that range, so the GA remains driven toward the correct range even before reaching an optimal solution. The three optimization processes were run independently, producing three mixture designs, each representing a different structural application scenario non-structural, general residential, and multi-story/high-seismic-zone structures.

These same code-derived bounds a 0.3-0.8 water-to-binder window, a 30%-of-binder ceiling on fly ash, and a 70%-of-binder ceiling on slag were enforced as hard constraints throughout every generation of the evolutionary search, so no candidate mixture violating the SNI limits could survive selection. Applying these constraints during both the crossover and mutation stages ensures that all candidate solutions explored by the GA not only the final recommendation remain within realistic composition limits and comply with concrete structural engineering practices, so the optimization results can be directly considered as a starting point for trial mix formulation in the field.

III. Results and Discussion

A. Prediction Model Benchmarking Results

The five models were evaluated using an identical 5-fold cross-validation scheme as described in the Methods section. A summary of each model's performance is presented in Table 5 to facilitate comparison among models.

TABLE V. BENCHMARKING RESULTS OF THE FIVE MODELS USING AN IDENTICAL 5-FOLD CROSS-VALIDATION SCHEME

Model	R^2 (mean±std)	RMSE (MPa)	MAE (MPa)	MAPE (%)
SVR	0.939 ± 0.020	3.30 ± 0.45	2.07 ± 0.19	6.34 ± 0.63
XGBoost	0.934 ± 0.027	3.41 ± 0.55	2.34 ± 0.28	7.20 ± 0.58
LightGBM	0.931 ± 0.026	3.52 ± 0.53	2.38 ± 0.28	7.40 ± 0.77
ANN	0.925 ± 0.026	3.66 ± 0.52	2.72 ± 0.36	8.25 ± 1.01
Random Forest	0.920 ± 0.033	3.78 ± 0.60	2.62 ± 0.32	8.14 ± 0.55

Among the five evaluated algorithms, SVR obtained the highest mean R^2 of 0.939, with XGBoost and LightGBM following at 0.934 and 0.931, respectively. ANN achieved an R^2 of 0.925, while Random Forest recorded 0.920. SVR also produced the lowest RMSE, MAE, and MAPE values reported in Table 5. Although the differences between the models are relatively small, the combination of these indicators identifies SVR as the model with the strongest overall validation performance for the present dataset. Although the performance gap across all models is relatively small (R^2 range of 0.920 0.939), the consistent advantage of SVR indicates that for a medium-sized dataset (1,074 samples) with eight continuous numerical variables that are physically correlated with one another for example through the water-cement ratio and binder proportion the RBF kernel in SVR is able to capture the non-linearity of the composition-strength relationship well, without the higher overfitting risk associated with higher-capacity models such as tree-based ensembles.

From a materials engineering perspective, this finding is consistent with the characteristics of the concrete compressive strength dataset, which essentially follows a relatively smooth physical relationship among the water-cement ratio, binder content, and curing age a data structure that naturally aligns with the maximum-margin assumption of SVR. Per-fold result analysis further reveals that Random Forest, XGBoost, LightGBM, and ANN experienced a consistent performance decline at the 4th fold (R^2 dropping to a range of 0.86 0.89), while SVR remained relatively stable on the same fold (R^2=0.924), indicating SVR's superior robustness to data subset variability likely containing extreme concrete-age combinations or less common mixture proportions.

G. Explainability Analysis of the Best Model

Because SVR produced the strongest cross-validation performance, this model was selected for the subsequent explainability analysis. SHAP was then used to quantify the contribution of the eight input variables to the predicted compressive strength. The global feature ranking is

presented in Fig. 1, while the distribution of individual SHAP contributions is illustrated in Fig. 2.

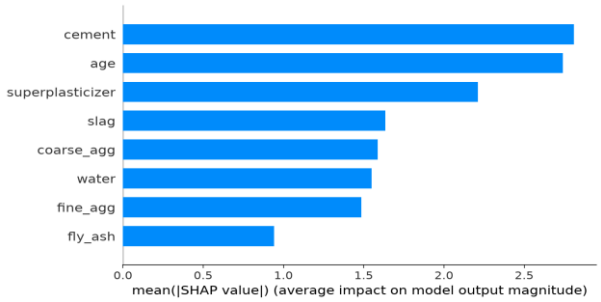


Fig. 1. Global ranking of variable influence on compressive strength prediction (mean |SHAP value|)

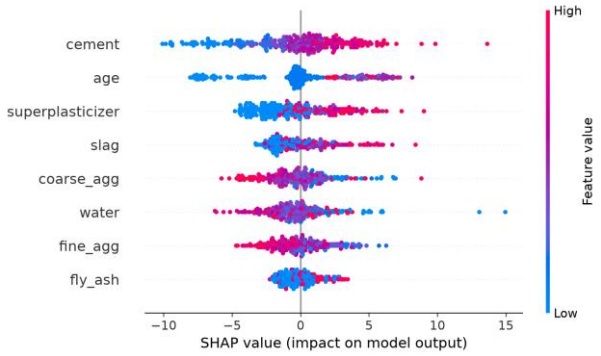
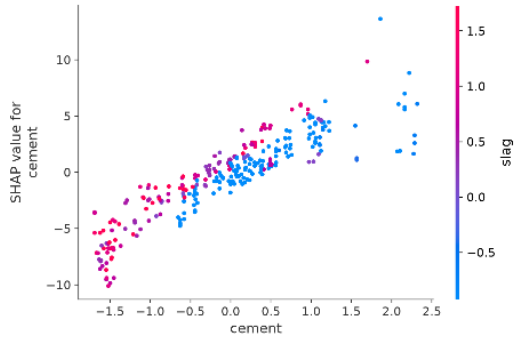
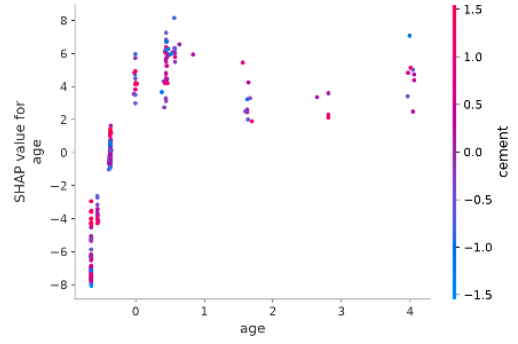


Fig. 2. SHAP summary plot: distribution and direction of each variable's influence across all samples

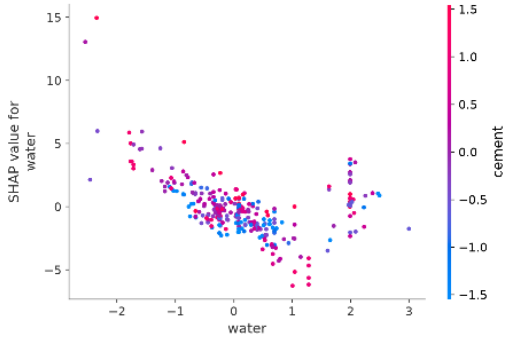
Cement and age were identified as the two most influential variables in compressive strength prediction, consistent with the fundamental civil engineering principle regarding the role of cement hydration as the primary determinant of concrete strength development. To reveal more detailed relationship patterns, Fig. 5 presents SHAP dependence plots for three key variables side by side: (A) cement, (B) age, and (C) water.



(A)



(B)



(C)

Fig. 3. SHAP dependence plots for (A) cement, (B) age, and (C) water

Fig. 3 (A) shows a strong, nearly linear, positive monotonic relationship between cement and the SHAP value: the higher the cement content, the greater its positive contribution to compressive strength prediction. The coloring pattern in this panel also indicates a cement-slag interaction, in which, at high cement levels, samples with low slag content tend to have a larger SHAP contribution than samples with high slag content at an equivalent cement level indicating a partial cement-substitution effect by slag that proportionally reduces cement's marginal contribution to compressive strength, consistent with the secondary hydration mechanism of slag, which proceeds more slowly than Portland cement hydration.

Fig. 3 (B) reveals a non-linear pattern with a saturating effect on the age variable: the SHAP contribution rises sharply from very young concrete ages to intermediate ages, then levels off at later ages. This pattern aligns with the classic cement hydration curve, which approaches an asymptote after a certain curing period (generally 28 90 days), and confirms that the SVR model successfully captures the physical dynamics of the cement hydration process without needing to be explicitly stated in the form of reaction kinetics equations in the input data.

Fig. 3 (C) Water exhibits a predominantly negative association with the SHAP contribution to predicted compressive strength. Higher water content is generally accompanied by a lower contribution to the model output. This tendency is compatible with the established relationship between water-to-cement ratio and concrete strength, in which excessive water can increase the capillary pore structure of hardened cement paste and consequently reduce compressive performance. The agreement between the learned response and this established engineering relationship provides supporting evidence that the model has captured physically meaningful patterns in the dataset.

Fly ash showed the lowest SHAP influence among the eight variables analyzed (Fig. 1), indicating that its contribution to compressive strength within the testing age range in this dataset is relatively smaller than that of slag. This finding is consistent with the pozzolanic reaction of fly ash, which proceeds more slowly than the cementitious properties of slag, so its contribution to early-age strength tends to be lower although in the long term, fly ash still plays an important role in durability and reducing concrete permeability, aspects that fall outside the scope of the short-term compressive strength target variable in the

dataset used.

H. Genetic Algorithm Optimization Results

GA optimization was run independently for the three concrete quality classes, with SVR as the surrogate model (selected based on its best cross-validation performance in the benchmarking stage). The optimal composition results for each class are presented in Table 6.

TABLE VI. OPTIMAL MIXTURE COMPOSITIONS FROM GENETIC ALGORITHM OPTIMIZATION FOR THE THREE CONCRETE QUALITY CLASSES (SVR AS SURROGATE MODEL)

Variable	Class I (target 15 MPa)	Class II (target 27.5 MPa)	Class III (target 55 MPa)
Cement (kg/m³)	268.29	141.03	495.69
Slag (kg/m³)	350.00	204.96	12.53
Fly Ash (kg/m³)	65.44	48.60	101.13
Water (kg/m³)	249.99	236.67	232.65
Superplasticizer (kg/m³)	1.73	8.77	13.97
Coarse Aggregate (kg/m³)	1191.42	1117.84	782.23
Fine Aggregate (kg/m³)	736.49	500.66	650.63
Age (days)	20.53	206.54	74.97
Predicted Compressive Strength (MPa)	15.00	27.50	54.99
Water-Binder Ratio (w/b)	0.366	0.600	0.382
Slag Substitution (% binder)	51.2%	51.9%	2.1%
Fly Ash Substitution (% binder)	9.6%	12.3%	16.6%

The results in Table 6 show that the GA successfully achieved predicted compressive strengths very close to the target of each class Class I reached 15.00 MPa (target 15 MPa), Class II reached 27.50 MPa (target 27.5 MPa, an exact match), and Class III reached 54.99 MPa (target 55 MPa) demonstrating excellent GA convergence toward the designed target-based fitness function. All three compositions were also validated as meeting all SNI constraints, including the water-binder ratio within the 0.3 0.8 range and the fly ash and slag substitution limits relative to total binder.

The binder substitution pattern in Table 6 reveals an interesting and engineering-relevant trend: for Class I and II (low-medium strength), the GA tends to utilize a high proportion of slag substitution (51 52% of binder) to achieve the strength target with lower cement content, consistent with the principle of cost efficiency and industrial waste utilization in non-structural and light structural applications that do not demand high early strength. Conversely, for Class III (high strength), the GA drastically reduces the slag proportion (2.1%) and increases cement content close to the SNI upper limit (495.69 kg/m³ out of a limit of 500 kg/m³), consistent with the SHAP analysis findings in Table 6 that cement is the most influential variable for high compressive strength, while slag which reacts through slower secondary hydration is less effective for pursuing high-strength targets within a reasonable curing age range. This pattern shows that the GA framework not only produces mathematically optimal figures but also captures meaningful engineering trade-offs between the use of supplementary materials and achieving structural strength targets.

From a construction sustainability perspective, the Class III results, which show a low slag substitution (2.1%), indicate that for very high strength targets, the contribution of supplementary materials to reducing Portland cement use is relatively limited compared to Class I and II. If a specific sustainability goal demands a certain minimum substitution proportion even for high-strength classes, the GA objective function would need to be reformulated into a multi-objective optimization that explicitly considers the binder substitution proportion alongside achieving the strength target a development direction discussed further in the future research recommendations section.

IV. CONCLUSION

An integrated machine-learning framework was developed to predict and optimize fly ash–slag concrete mixtures while incorporating the composition restrictions adopted from the relevant SNI standards. The sensitivity analysis of synthetic-data augmentation identified 2% proportional noise as the most suitable setting, producing an SVR probe-model R^2 of 0.881. Comparative evaluation of five prediction algorithms after Bayesian hyperparameter tuning subsequently showed that SVR achieved the highest predictive performance, with an R^2 of 0.939, followed by XGBoost, LightGBM, ANN, and Random Forest. These findings demonstrate the usefulness of systematic hyperparameter optimization and uniform cross-validation when evaluating alternative learning methods for concrete-strength prediction.

The SHAP results indicate that cement content and curing age have major contributions to the predicted compressive strength, while superplasticizer is also identified as an influential variable in the overall model interpretation. The dependence patterns observed for cement, age, and water exhibit trends that are compatible with established concepts of cement hydration and the relationship between water-to-cement ratio and strength. These results provide an interpretable perspective on the behavior learned by the SVR model. In the optimization stage, the Genetic Algorithm generated candidate mixtures with predicted strengths of 15.00 MPa for Class I, 27.50 MPa for Class II, and 54.99 MPa for Class III. The candidate solutions satisfied the specified SNI-based composition constraints.

The combination of model benchmarking, augmentation-level sensitivity analysis, SHAP interpretation, and target-based Genetic Algorithm optimization provides a unified computational workflow for preliminary fly ash–slag concrete mixture formulation. The framework can assist in narrowing the mixture-design search space while also providing information about the variables associated with the model predictions. Nevertheless, the optimized compositions represent computationally generated candidates rather than experimentally confirmed concrete formulations. Laboratory preparation and compressive-strength testing, or validation using independent experimental datasets outside the UCI database, are therefore required before the proposed mixtures can be considered for practical implementation.

Further development of the framework should consider multiple design objectives simultaneously. Future optimization could combine compressive strength with fly

ash and slag utilization, mixture cost, and other indicators of industrial-waste incorporation. Algorithms such as Non-dominated Sorting Genetic Algorithm II (NSGA-II) could be investigated to generate a set of Pareto-optimal mixtures representing different compromises between mechanical and sustainability objectives. In addition, expanding the database with fly ash and slag obtained from different regions and material sources in Indonesia would provide broader variability in material characteristics and could improve the applicability of the model to local concrete-material conditions.

ACKNOWLEDGMENT

The authors would like to express their sincere gratitude to Universitas Bojonegoro for providing academic support and facilities during the completion of this research. The authors also appreciate the contributions and support of all individuals who assisted in the preparation, analysis, and completion of this study.

REFERENCES

[1] X.-Y. Wang, "Impact of climate change on the optimization of mixture design of low-CO2 concrete containing fly ash and slag," *Sustainability*, vol. 11, no. 12, e3394, Jun. 2019, doi: 10.3390/su11123394.

[2] S. Barbhuiya, B. B. Das, D. Adak, K. Kapoor, and M. Tabish, "Low carbon concrete: Advancements, challenges and future directions in sustainable construction," *Discover Concrete and Cement*, vol. 1, no. 1, e3, Mar. 2025, doi: 10.1007/s44416-025-00002-y.

[3] Q. Chen *et al.*, "Compressive strength prediction of high-performance concrete: Integrating multi-ingredient influences and mix proportion insights," *Constr. Build. Mater.*, vol. 451, e138791, Nov. 2024, doi: 10.1016/j.conbuildmat.2024.138791.

[4] M. R. Raju *et al.*, "Machine learning for predicting strength properties of waste iron slag concrete," *Heliyon*, vol. 11, no. 3, e42133, Feb. 2025, doi: 10.1016/j.heliyon.2025.e42133.

[5] A. A. B. Loureiro and R. Stefani, "Comparing the performance of machine learning models for predicting the compressive strength of concrete," *Discover Civil Eng.*, vol. 1, no. 1, e19, Jun. 2024, doi: 10.1007/s44290-024-00022-w.

[6] S. I. Hassan *et al.*, "Systematic literature review on the application of machine learning for the prediction of properties of different types of concrete," *PeerJ Comput. Sci.*, vol. 10, e1853, May 2024, doi: 10.7717/peerj-cs.1853.

[7] Z. Yuan, W. Zheng, and H. Qiao, "Machine learning based optimization for mix design of manufactured sand concrete," *Constr. Build. Mater.*, vol. 467, e140256, Mar. 2025, doi: 10.1016/j.conbuildmat.2025.140256.

[8] M. Kioumarsi, H. Dabiri, A. Kandiri, and V. Farhangi, "Compressive strength of concrete containing furnace blast slag; optimized machine learning-based models," *Clean. Eng. Technol.*, vol. 13, e100604, Apr. 2023, doi: 10.1016/j.clet.2023.100604.

[9] H. X. Mohammed, A. S. Mohammed, and A. M. T. Hassan, "Soft computing models to evaluate the effect of fly ash and ground granulated blast furnace slag (GGBS) on the compressive strength of concrete in normal and high strength ranges," *Structures*, vol. 58, e105459, Dec. 2023, doi: 10.1016/j.istruc.2023.105459.

[10] S. K. Parhi and S. K. Patro, "Prediction of compressive strength of geopolymer concrete using a hybrid ensemble of grey wolf optimized machine learning estimators," *J. Building Eng.*, vol. 71, e106521, Jul. 2023, doi: 10.1016/j.jobe.2023.106521.

[11] X. Luo, Y. Li, Q. Wang, J. Mu, and Y. Liu, "Machine learning based modeling for predicting the compressive strength of solid waste material-incorporated magnesium phosphate cement," *J. Clean. Prod.*, vol. 442, e141172, Feb. 2024, doi: 10.1016/j.jclepro.2024.141172.

[12] S. Elhishi, A. M. Elashry, and S. El-Metwally, "Unboxing machine learning models for concrete strength prediction using XAI," *Sci. Rep.*, vol. 13, no. 1, e19892, Nov. 2023, doi: 10.1038/s41598-023-47169-7.

[13] M. S. N. Tak, Y. Feng, and M. Mahgoub, "Advanced machine learning techniques for predicting concrete compressive strength," *Infrastructures (Basel).*, vol. 10, no. 2, e26, Jan. 2025, doi: 10.3390/infrastructures10020026.

[14] S. Wang, P. Xia, F. Gong, Q. Zeng, K. Chen, and Y. Zhao, "Multi objective optimization of recycled aggregate concrete based on explainable machine learning," *J. Clean. Prod.*, vol. 445, e141045, Mar. 2024, doi: 10.1016/j.jclepro.2024.141045.

[15] Y. Gu, R. Fan, Y. Li, J. Zhao, Z. Song, and H. Chu, "Multi-objective optimization for nano-silica-modified concrete based on explainable machine learning," *Nanomaterials*, vol. 15, no. 18, e1423, Sep. 2025, doi: 10.3390/nano15181423.

[16] R. K. Tipu, V. R. Panchal, and K. S. Pandya, "Multi-objective optimization of the concrete mixture blended with mineral admixture using machine learning and NSGA-II algorithms," in *Lecture Notes in Mechanical Eng.*, 2023, pp. 165–174. doi: 10.1007/978-981-19-9285-8_16.

[17] W. J. Park, T. Noguchi, and H. S. Lee, "Genetic algorithm in mix proportion design of recycled aggregate concrete," *Computers & concrete*, vol. 11, no. 3, pp. 183–199, Mar. 2013, doi: 10.12989/cac.2013.11.3.183.

[18] I. Yeh, "Concrete Compressive Strength," UCI Machine Learning Repository, 1998, doi: 10.24432/C5PK67.

[19] L. Breiman, "Random Forests," *Mach. Learn.*, vol. 45, no. 1, pp. 5–32, Oct. 2001, doi: 10.1023/A:1010933404324.

[20] T. Chen and C. Guestrin, "XGBoost," in *Proc. 22nd ACM SIGKDD Int. Conf. Knowl. Discov. Data Mining*, New York, Aug. 2016, pp. 785–794. doi: 10.1145/2939672.2939785.

[21] G. Ke *et al.*, "LightGBM: A highly efficient gradient boosting decision tree," in *Adv. Neural Inf. Processing Systems*, Curran Associates, Inc., 2017, pp. 3146–3154. [Online]. Available: https://proceedings.neurips.cc/paper_files/paper/2017/file/6449f44a102fde848669bdd9eb6b76fa-Paper.pdf

[22] T. Akiba, S. Sano, T. Yanase, T. Ohta, and M. Koyama, "Optuna," in *Proc. 25th ACM SIGKDD Int. Conf. Knowl. Discov. Data Mining*, New York, Jul. 2019, pp. 2623–2631. doi: 10.1145/3292500.3330701.

[23] J. Bergstra, R. Bardenet, Y. Bengio, and Y. Kegl, "Algorithms for hyper-parameter optimization," in *Adv. Neural Inf. Processing Systems*, J. Shawe-Taylor, R. Zemel, P. Bartlett, F. Pereira, and K. Weinberger, Eds., Curran Associates, Inc., 2011, pp. 2546–2554. [Online]. Available: https://proceedings.neurips.cc/paper_files/paper/2011/file/86e8f7ab32cfd12577bc2619bc635690-Paper.pdf

[24] M. Z. Naser and A. H. Alavi, "Error metrics and performance fitness indicators for artificial intelligence and machine learning in engineering and sciences," *Architecture, Structures and Construction*, vol. 3, no. 4, pp. 499–517, Dec. 2023, doi: 10.1007/s44150-021-00015-8.

[25] S. M. Lundberg and S.-I. Lee, "A unified approach to interpreting model predictions," in *Adv. Neural Inf. Processing Systems*, 2017, pp. 4765–4774. [Online]. Available: https://proceedings.neurips.cc/paper_files/paper/2017/hash/8a20a8621978632d76c43dfd28b67767-Abstract.html



Ichwan Hadi Saputra is a Lecturer in the Department of Civil Engineering at the Faculty of Science and Engineering, Universitas Bojonegoro (UNIGORO), Bojonegoro, Indonesia. His current research focuses on sustainable construction materials, machine learning applications in civil engineering, and construction project management. (email: ichwanhs@gmail.com).



Eko Wahyu Abryandoko is a Lecturer in the Department of Industrial Engineering at the Faculty of Science and Engineering, Universitas Bojonegoro (UNIGORO), Bojonegoro, Indonesia, where he currently serves as the Head of the Department of Industrial Engineering. (email: abryandoko@gmail.com).



Widyo Nugroho is a Lecturer in the Department of Civil Engineering at the Faculty of Science and Engineering, Universitas Bojonegoro (UNIGORO), Bojonegoro, Indonesia. (email: widyo.nugroho@gmail.com).



Moh. Nurudduja is a graduate of the Industrial Engineering Study Program at the Faculty of Science and Engineering, Universitas Bojonegoro (UNIGORO), Bojonegoro, Indonesia. (email: m.nurudduja@gmail.com).



Nayla Farikha Zahra is an undergraduate student in the Department of Industrial Engineering at the Faculty of Science and Engineering, Universitas Bojonegoro (UNIGORO), Bojonegoro, Indonesia. (email: nayfarik937@gmail.com).