



International Journal of Artificial Intelligence and Machine Learning

Publisher's Home Page: <https://www.svedbergopen.com/>



Open Access

Machine Learning Based Prediction of Flexural Strength of Reinforced Concrete Using Concrete Mix and Material Properties

Amrendra Kumar¹, Ajay Kumar^{2*}, Mani Bhushan³, Aviral Kumar⁴, Md Irfan Ansari⁵, Abhishek Kumar⁶

¹Assistant Professor, Department of Civil Engineering, Sandip University, Madhubani, Bihar India-847235. Email: amrendraroy2k8@gmail.com, ORCID: 0000-0001-6092-8271

^{2*}Assistant Professor, Department of Civil Engineering, Sandip University, Madhubani, Bihar India-847235. Email: ajaykumar8mail@gmail.com, ORCID: 0009-0002-8418-0135

³Department of Civil Engineering, Government Engineering College Khagaria, Bihar, India (Affiliated to Bihar Engineering University, Patna, Bihar, India; under the Department of Science, Technology and Technical Education, Government of Bihar). Email: mani@geckhagaria.org.in, ORCID: 0009-0007-0182-1045

⁴Department of Civil Engineering, Government Engineering College Khagaria, Bihar, India (Affiliated to Bihar Engineering University, Patna, Bihar, India; under the Department of Science, Technology and Technical Education, Government of Bihar). Email: shahiaviral1@gmail.com, ORCID: 0009-0003-5452-2939

⁵Department of Civil Engineering, Government Engineering College Khagaria, Bihar, India (Affiliated to Bihar Engineering University, Patna, Bihar, India; under the Department of Science, Technology and Technical Education, Government of Bihar). Email: er.mdirfan2k9@gmail.com, ORCID: 0000-0002-3592-9093

⁶Department of Civil Engineering, Government Engineering College Khagaria, Bihar, India (Affiliated to Bihar Engineering University, Patna, Bihar, India; under the Department of Science, Technology and Technical Education, Government of Bihar). Email: abhi.geckh@gmail.com, ORCID: 0009-0006-6079-0675

ABSTRACT

Accurate estimation of flexural strength is important for the assessment and proportioning of reinforced concrete, particularly when conventional concrete mixtures are modified through the incorporation of recycled aggregates and supplementary cementitious materials. The interaction among aggregate replacement, binder composition, water-cement ratio, and reinforcement characteristics makes direct prediction of flexural performance difficult using simple empirical relationships. Machine learning (ML) provides an alternative data-driven approach for capturing such interactions and identifying useful relationships between concrete composition, reinforcement parameters, and mechanical response. Previous research has demonstrated the potential of ML for predicting concrete strength and has highlighted the importance of appropriate feature selection, model validation, and interpretation when dealing with complex cementitious systems.

In this study, an ML-based framework is developed to predict the 28-day flexural strength of reinforced concrete using concrete mix and material properties. A structured dataset comprising 36 mix designs and four replicate observations for each mix, resulting in 144 observations, was developed for the analysis. The predictor variables include recycled coarse aggregate replacement, supplementary cementitious material replacement, silica fume content, water-cement ratio, steel grade, reinforcing-bar diameter, reinforcement ratio, and effective depth. The dataset was subjected to descriptive statistical analysis, Pearson correlation analysis, and variance inflation factor assessment to identify redundant predictors and control multicollinearity. Nine regression algorithms Linear Regression, Ridge, Lasso, Elastic Net, Support Vector Regression, K-Nearest Neighbors, Random Forest, Gradient Boosting, and Extra Trees were comparatively evaluated using grouped five-fold cross-validation, with observations from the same mix retained within the same validation group.

Among the evaluated models, Lasso regression produced the highest reported coefficient of determination, with an R^2 value of 0.984623, followed by Elastic Net ($R^2 = 0.952104$) and Linear Regression ($R^2 = 0.912514$). The corresponding Lasso model achieved a mean absolute error of 0.125826 MPa, root mean square error of 0.157385 MPa, and mean absolute percentage error of 2.124262%. The results indicate that a carefully selected combination of concrete mixture and reinforcement-related variables can provide a useful basis for data-driven estimation of flexural strength. The study also demonstrates the importance of feature screening before ML modelling, particularly when concrete mixture variables contain direct or indirect relationships. The proposed framework can support preliminary mixture assessment and reduce dependence on repeated trial-and-error evaluation, while laboratory testing remains essential for final structural and material qualification. ML-based prediction should therefore be considered a complementary engineering tool rather than a replacement for established experimental and design procedures.

Keywords: reinforced concrete; flexural strength; recycled coarse aggregate; supplementary cementitious materials; machine learning; Lasso regression; feature screening; grouped cross-validation; sustainable concrete

1. INTRODUCTION

Concrete remains one of the most widely used construction materials because of its versatility, availability, and ability to provide the compressive capacity required in a broad range of structural applications. At the same time, the conventional production of concrete relies heavily on natural aggregates and cementitious materials, creating considerable demand for virgin resources and contributing to the environmental burden associated with construction. The increasing generation of construction and demolition waste has consequently encouraged the development of recycled aggregate concrete (RAC), in which aggregates recovered from demolished concrete are incorporated into new concrete mixtures. The use of recycled aggregate can reduce dependence on natural aggregate resources and provide a productive pathway for diverting construction waste from disposal. However, the engineering performance of recycled aggregate concrete is influenced by the quality and proportion of recycled aggregate, the residual mortar attached to aggregate particles, water absorption, mixture proportions, and the characteristics of the parent concrete. These factors introduce greater variability than is generally encountered in conventional natural aggregate concrete (Ahmed et al., 2022; Nguyen et al., 2023; Zhao et al., 2024, McNeil & Kang, 2013).

The mechanical behaviour of concrete is not governed by a single mixture parameter. Changes in aggregate characteristics interact with cement content, supplementary cementitious materials (SCMs), water–cement ratio, admixture dosage, curing conditions, and the internal structure of the cementitious matrix. The incorporation of fly ash, ground granulated blast-furnace slag (GGBS), and silica fume can further modify hydration, pore structure, interfacial transition zones, and the resulting mechanical response. In recycled aggregate concrete, these interactions become particularly important because recycled aggregates can introduce additional adhered mortar, higher porosity, and different absorption characteristics. Consequently, conventional empirical relationships developed for relatively narrow material systems may not adequately represent the combined influence of mixture composition and material variability across a broader design space (Ahmed et al., 2022; Nguyen et al., 2023; Zhao et al., 2024, Momeni et al., 2022, Yuan et al., 2022).

Flexural strength is an important mechanical property because it characterizes the tensile response of concrete under bending and is relevant to the assessment of cracking behaviour, pavement applications, and structural material performance. Unlike compressive strength, flexural strength is particularly sensitive to the heterogeneous nature of concrete and to the interaction between the cementitious matrix and aggregate skeleton. When reinforcement-related parameters are also considered, the resulting response becomes dependent on additional variables such as reinforcement ratio, bar diameter, steel grade, and effective depth. Therefore, a predictive framework intended for reinforced concrete should not rely exclusively on conventional mixture proportions when the objective is to capture the combined effect of concrete composition and reinforcement characteristics.

Machine learning (ML) offers a useful alternative for representing these nonlinear and potentially interacting relationships. Rather than imposing a predetermined functional form, ML regression algorithms can learn relationships between multiple input variables and a target response from available observations. A growing body of research has demonstrated the application of artificial intelligence and ML techniques to concrete-property prediction, including compressive strength, tensile strength, durability indicators, and other performance characteristics. Studies on recycled aggregate concrete have particularly highlighted the potential of data-driven models to accommodate the influence of recycled aggregate content, water–cement ratio, cement content, and other mixture parameters (Ahmed et al., 2022; Nguyen et al., 2023; Zhao et al., 2024, Qureshi et al., 2020).

Despite this progress, several methodological issues remain important. First, many ML studies concentrate predominantly on compressive-strength prediction, while comparatively less attention is given to flexural-strength prediction using a combination of concrete mixture and reinforcement variables. Second, concrete datasets commonly contain strongly related variables. For example, replacement percentages may be mathematically related to the corresponding material masses, while water–cement ratio is related to water and binder contents. Including all such variables without screening can introduce multicollinearity and make model interpretation less reliable. Recent reviews of ML applications in recycled aggregate concrete have emphasized that model performance depends strongly on input-variable selection, dataset structure, algorithm choice, and hyperparameter settings (Nguyen et al., 2023; Zhao et al., 2024, Gao et al., 2020).

A further concern is validation methodology. When multiple observations originate from the same concrete mixture, randomly distributing individual observations between training and validation subsets can result in information leakage because closely related replicates may appear in both subsets. Such a procedure can produce overly optimistic estimates of predictive performance. Group-based validation provides a more conservative alternative by keeping observations belonging to the same mixture within the same fold. This consideration is particularly relevant to small or medium-sized engineering datasets, where the number of unique mix designs may be substantially lower than the total number of recorded observations. Recent reviews have identified dataset quality, validation strategy, and generalization to unseen material combinations as continuing challenges in data-driven concrete research (Zhao et al., 2024; Fan et al., 2025, Bui et al., 2017, Moein et al., 2023,).

Another important issue is interpretability. A high prediction score alone does not establish that an ML model

provides useful engineering knowledge. For concrete applications, researchers need to understand which mixture or reinforcement variables contribute most strongly to the predicted response and whether the resulting relationships are consistent with known material behaviour. Feature screening, correlation analysis, multicollinearity assessment, and model-based importance analysis can therefore complement conventional performance metrics such as the coefficient of determination (R^2), mean absolute error (MAE), root mean square error (RMSE), and mean absolute percentage error (MAPE). This combination provides a more transparent basis for evaluating data-driven prediction than reporting a single accuracy measure (Nguyen et al., 2023; Zhao et al., 2024, Kumar et al., 2024, Hassan et al., 2024).

Against this background, the present study develops a machine-learning framework for predicting the 28-day flexural strength of reinforced concrete using a combination of concrete mixture and reinforcement-related properties. The study considers recycled coarse aggregate replacement, SCM replacement, silica fume content, water–cement ratio, steel grade, reinforcing-bar diameter, reinforcement ratio, and effective depth as the principal predictive variables. Nine regression algorithms are comparatively examined: Linear Regression, Ridge, Lasso, Elastic Net, Support Vector Regression, K-Nearest Neighbors, Random Forest, Gradient Boosting, and Extra Trees. Before model development, the predictor space is examined using correlation analysis and variance inflation factor (VIF) assessment to reduce unnecessary redundancy. The models are then evaluated using grouped five-fold cross-validation based on the concrete mix identifier, thereby preventing replicate observations from the same mixture from being distributed across both training and validation folds.

The specific objectives of the study are therefore to: (i) develop a structured dataset for data-driven prediction of reinforced-concrete flexural strength; (ii) identify a compact and technically meaningful set of mixture and reinforcement predictors; (iii) investigate the influence of multicollinearity and feature relationships on model development; (iv) compare the predictive performance of nine widely used ML regression algorithms under grouped cross-validation; and (v) examine the relative importance of the selected predictors for engineering interpretation. The resulting framework is intended as a computational decision-support approach for preliminary assessment of flexural performance rather than as a replacement for standardized laboratory testing or structural design procedures.

2. LITERATURE REVIEW AND RESEARCH GAP

The application of machine learning in concrete engineering has expanded considerably over the last decade. Early studies largely focused on predicting compressive strength because of the availability of relatively large datasets and the importance of compressive strength in concrete characterization. More recently, researchers have extended data-driven methods to tensile strength, flexural strength, durability, workability, and other performance indicators. A critical review by Ahmad et al. (2020) showed that ML models can accommodate a larger number of mixture variables than conventional empirical formulations, although their reliability depends strongly on dataset quality, input selection, and model-development procedures. Similar observations have been reported in subsequent reviews, which identified artificial neural networks, support vector methods, ensemble learning, and regression-based approaches among the frequently used techniques for concrete-property prediction.

2.1 Machine Learning for Concrete Strength Prediction

Concrete strength is governed by complex interactions between cementitious materials, aggregate characteristics, water content, admixtures, curing conditions, and age. Such interactions make it difficult to represent all possible combinations through a single conventional empirical equation. ML algorithms provide a flexible alternative because they can approximate nonlinear relationships without requiring the researcher to specify the complete mathematical form of the relationship beforehand. Reviews of ML applications in concrete have consequently reported considerable growth in data-driven strength prediction and identified model selection, feature engineering, validation, and generalization as major methodological considerations (Ahmad et al., 2020; Nguyen et al., 2023; Zhao et al., 2024).

The choice of algorithm can substantially affect the resulting prediction accuracy. Linear and regularized regression methods remain useful when the relationship between predictors and response is approximately systematic, while kernel-based approaches such as Support Vector Regression can capture nonlinear patterns through suitable transformations. Tree-based ensemble methods, including Random Forest, Gradient Boosting, and Extra Trees, can represent interactions and nonlinear effects without requiring feature scaling. Studies comparing several algorithms on concrete datasets have therefore emphasized that there is no universally optimal ML algorithm; model performance depends on the characteristics, size, variability, and structure of the dataset being analysed (Ahmad et al., 2020; Nguyen et al., 2023; Zhao et al., 2024).

An important development has been the application of ML to sustainable concrete systems. The incorporation of recycled aggregates and supplementary cementitious materials introduces additional variability because recycled aggregates may contain adhered mortar, microcracks, and pores that differ substantially from natural aggregate surfaces. Zhao et al. (2024) emphasized that recycled aggregate concrete (RAC) properties are affected by factors such as recycled aggregate content, water–cement ratio, curing conditions, and aggregate quality. Nguyen et al. (2023) similarly identified water–cement ratio and cement-related

parameters as influential variables in AI-based prediction of RAC mechanical properties. These findings demonstrate why models developed for conventional concrete cannot necessarily be transferred directly to recycled concrete systems without considering their specific material characteristics.

2.2 Recycled Aggregate Concrete and Flexural Behaviour

The mechanical behaviour of RAC differs from that of conventional concrete because recycled aggregates commonly contain residual cement paste or mortar from the parent concrete. Their rougher surfaces, higher porosity, greater water absorption, and possible internal defects can modify the interfacial transition zone and consequently influence strength and deformation behaviour. These characteristics become particularly relevant to tensile and flexural performance, where cracking and the quality of the aggregate–paste interface have a pronounced influence on the measured response. Previous ML research on RAC has consequently incorporated variables such as recycled aggregate replacement, parent concrete strength, water–cement ratio, water absorption, and aggregate characteristics when predicting mechanical properties (Alam et al., 2022; Nguyen et al., 2023; Zhao et al., 2024).

Alam et al. (2022) developed ML models for both compressive and flexural strength of RAC and demonstrated that ensemble techniques could successfully capture relationships involving multiple material parameters. Their study is particularly relevant to the present work because it explicitly considered flexural strength rather than restricting the prediction problem to compressive strength. The reported results also demonstrated that variables associated with recycled aggregate characteristics and mixture composition can jointly influence RAC strength, supporting the need for multivariable predictive frameworks.

More recent work has moved closer to the specific problem addressed in the present study. A 2024 investigation combined experimental observations with ML models to predict flexural and tensile strengths of recycled concrete and used a dataset containing results from multiple studies. The authors identified effective water–cement ratio, recycled aggregate replacement ratio, aggregate-to-cement ratio, and recycled aggregate water absorption among the important variables affecting prediction. Their work demonstrates the growing recognition that flexural-strength prediction deserves separate treatment rather than simply being inferred from compressive strength models.

Another recent study examined the flexural behaviour of reinforced concrete beams containing recycled aggregates and CFRP reinforcement using ML-based prediction. This direction is important because it moves beyond material-only prediction toward structural behaviour and demonstrates that ML can also accommodate reinforcement-related parameters when sufficient structural information is available.

2.3 Role of Supplementary Cementitious Materials

SCMs provide an additional dimension to sustainable concrete design. Materials such as fly ash, GGBS, and silica fume can modify the binder system, hydration process, pore structure, and interfacial characteristics of concrete. Their effects are not necessarily independent of recycled aggregate content. For example, changes in binder composition can alter the quality of the cementitious matrix surrounding recycled aggregate particles, while the higher absorption and surface characteristics of recycled aggregates can influence the effective water availability of the mixture. Consequently, predicting mechanical properties of concrete containing both recycled aggregates and SCMs represents a multivariable problem rather than a simple single-factor relationship.

Recent reviews of RAC and sustainable concrete have highlighted the importance of simultaneously considering mixture composition and aggregate characteristics when developing predictive models. Nguyen et al. (2023) noted that the accuracy of AI-based RAC prediction depends not only on algorithm selection but also on the choice of input variables. Zhao et al. (2024) similarly emphasized that recycled aggregate ratio, water–cement ratio, curing conditions, and aggregate characteristics can interact to determine the resulting mechanical properties.

The growing literature on ML-based prediction of concrete properties also indicates a shift from simply maximizing prediction accuracy toward developing models that are more transparent and useful for engineering interpretation. A recent review of ML and deep-learning applications in concrete strength prediction identified sustainability, explainability, error analysis, and model generalization as increasingly important research themes. This shift is particularly relevant when models are intended to support engineering decisions rather than function solely as statistical predictors.

2.4 Feature Selection, Multicollinearity and Model Reliability

Feature selection is an important but sometimes overlooked component of concrete ML studies. Concrete mixture datasets often contain variables that convey overlapping information. For example, the replacement percentage of recycled aggregate can be directly related to the mass of recycled coarse aggregate, while SCM replacement percentage may be related to the individual quantities of fly ash, GGBS, or silica fume. Similarly, water–cement ratio is mathematically connected to water and cementitious-material quantities. Feeding all of these variables simultaneously into a regression model can introduce substantial multicollinearity.

Multicollinearity is particularly problematic for regression-based models because correlated predictors can make coefficient estimates unstable and complicate interpretation. It can also create apparently high predictive performance without necessarily improving the underlying generalizability of the model. For this reason, correlation analysis and variance inflation factor assessment provide useful preliminary screening

procedures before final model development. Reviews of ML-based RAC prediction have specifically identified input-variable selection as an important factor controlling model performance and interpretability (Nguyen et al., 2023; Zhao et al., 2024).

Regularized regression methods provide another approach to this issue. Ridge and Lasso regression introduce penalties that constrain model coefficients and can reduce sensitivity to correlated predictors. Lasso additionally provides a mechanism for reducing the contribution of less informative variables by shrinking some coefficients toward zero. These characteristics make regularized linear models particularly useful as benchmark models when the dataset contains several related engineering variables. Rather than assuming that a complex nonlinear algorithm must always provide superior performance, comparing regularized regression with tree-based and kernel-based models provides a more balanced assessment of the predictive structure within the dataset.

2.5 Validation Strategy in Small Engineering Datasets

Validation methodology is another critical consideration in ML-based concrete research. Many engineering datasets contain repeated measurements associated with the same mixture design. If replicate observations are randomly divided between training and testing subsets, the model may encounter highly similar observations during training and validation. Such leakage can produce an overly optimistic estimate of generalization performance.

The use of grouped cross-validation addresses this concern by assigning all observations associated with a particular mixture to the same fold. Consequently, the validation fold represents mixtures that were not directly used during model fitting. This provides a more meaningful assessment of whether the model can generalize beyond individual observations to other mix configurations. Recent work on RAC prediction has increasingly emphasized the importance of validation strategy and the distinction between row-level performance and generalization to genuinely unseen material groups. A 2026 physics-guided ML study, for example, reported substantially different behaviour when validation was performed at the study-group level rather than simply at the individual-record level.

This issue is directly relevant to the present dataset, which contains four observations for each of 36 mix designs. Accordingly, the present study uses the mix identifier as the grouping variable in five-fold cross-validation. This design prevents replicates belonging to the same mixture from being distributed across different folds and provides a more conservative evaluation of model generalization.

2.6 Research Gap

The existing literature establishes that ML can successfully model several properties of conventional and recycled concrete, but important gaps remain. First, a substantial proportion of previous work has concentrated on compressive strength, whereas flexural-strength prediction has received comparatively less attention. Second, studies dealing with RAC generally emphasize mixture and aggregate variables, while relatively fewer predictive frameworks combine concrete composition with explicit reinforcement characteristics. Third, several studies focus primarily on achieving high predictive accuracy without sufficiently addressing feature redundancy and multicollinearity. Fourth, the use of repeated observations from the same mixture creates a need for validation strategies that explicitly account for group structure.

The recent literature confirms that flexural-strength prediction is an active research direction. Alam et al. (2022) demonstrated the feasibility of ML for RAC flexural-strength prediction, while later work expanded the scope toward datasets containing results from multiple experimental sources and toward structural members incorporating recycled aggregates and reinforcement. Nevertheless, these studies differ substantially in their predictor variables, datasets, modelling strategies, and validation procedures. A consistent framework combining recycled aggregate content, SCM-related variables, water-cement ratio, and reinforcement parameters, while explicitly screening correlated predictors and evaluating models using grouped validation, remains useful.

The present study addresses this gap by developing a structured comparative ML framework for 28-day flexural-strength prediction using eight selected predictors: recycled coarse aggregate replacement, SCM replacement, silica fume content, water-cement ratio, steel grade, bar diameter, reinforcement ratio, and effective depth. The study compares nine regression algorithms under grouped five-fold cross-validation and incorporates correlation and VIF analysis before model evaluation. In addition, model-based feature importance is used to provide an engineering interpretation of the variables contributing to the prediction. The framework therefore combines prediction, feature screening, validation, and interpretation rather than treating model accuracy as the sole objective.

2.7 Positioning of the Present Study

The contribution of the present research is therefore methodological as well as predictive. Rather than proposing another single-algorithm concrete-strength model, the study establishes a comparative framework in which several fundamentally different regression approaches are evaluated using the same feature space and grouped validation strategy. The analysis also separates the influence of concrete mixture variables from reinforcement-related parameters and examines the effect of including compressive strength as an additional material-property predictor. This structure allows the resulting model performance to be interpreted in the context of the available input information rather than as an isolated accuracy value.

The literature indicates that ML has considerable potential for concrete-property prediction, but its engineering usefulness depends on disciplined dataset preparation, appropriate feature selection, transparent validation, and physically meaningful interpretation (Ahmad et al., 2020; Nguyen et al., 2023; Zhao et al., 2024). The present study builds on these observations by applying a structured and group-aware methodology specifically to flexural-strength prediction of reinforced concrete containing recycled and supplementary cementitious materials.

3. MATERIALS, DATASET DEVELOPMENT AND METHODOLOGY

3.1 Overall Research Framework

The methodology adopted in this study consists of five sequential stages: dataset development, exploratory statistical analysis, feature screening, machine-learning model development, and model evaluation and interpretation. The overall workflow was designed to examine whether concrete mixture and reinforcement-related properties can be used to predict the 28-day flexural strength of reinforced concrete through data-driven regression models.

The first stage involved developing a structured dataset containing concrete mixture characteristics, reinforcement parameters, and corresponding strength responses. The second stage involved descriptive statistical analysis and examination of the relationships among the available numerical variables. Because concrete mixture datasets frequently contain variables that convey overlapping information, feature screening was subsequently performed using correlation analysis and variance inflation factor (VIF) assessment. The screened variables were then used to develop and compare nine machine-learning regression models. Finally, grouped five-fold cross-validation was employed to evaluate model performance while ensuring that replicate observations belonging to the same mix design were not distributed across different validation folds.

This workflow follows the general principle that reliable machine-learning prediction requires not only an appropriate algorithm but also careful data preparation, feature selection, and validation (Kohavi, 1995; Pedregosa et al., 2011).

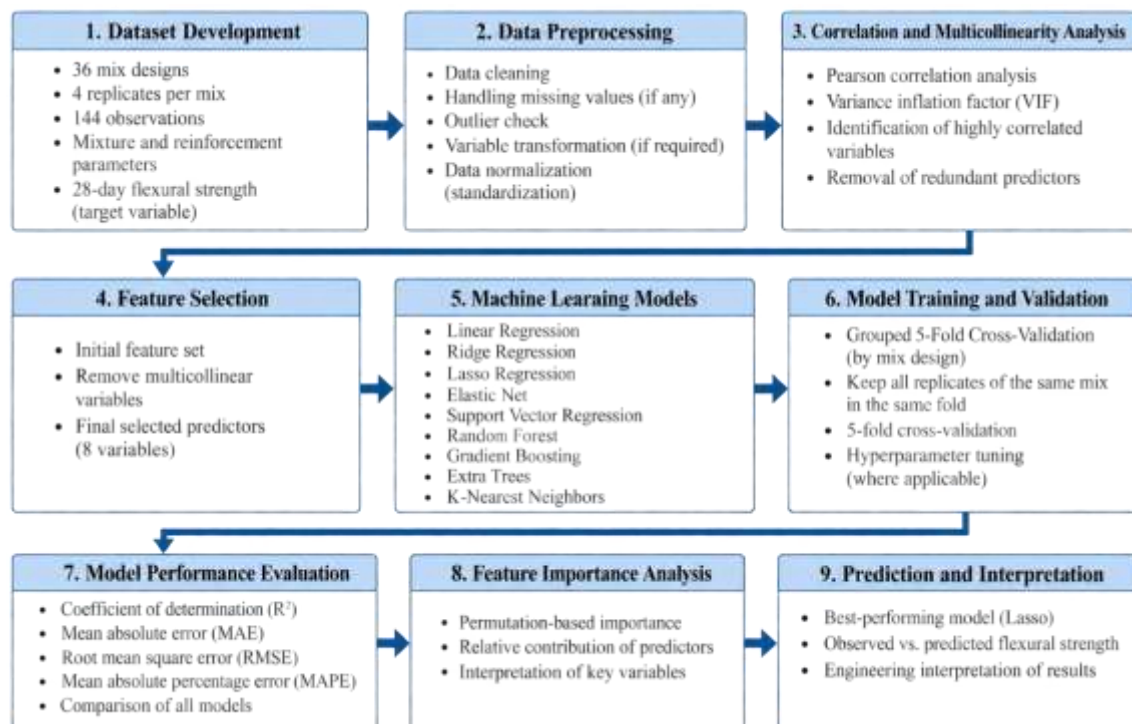


Figure 1. Overall research framework for machine-learning-based prediction of flexural strength

3.2 Dataset Development

A structured dataset comprising 36 unique concrete mix designs and four replicate observations for each mix was used, resulting in a total of 144 observations. The dataset was developed to represent variations in concrete mixture composition and reinforcement characteristics relevant to flexural-strength prediction. The mixture space included recycled coarse aggregate (RCA) replacement, supplementary cementitious material (SCM) replacement, and silica fume incorporation. Reinforcement-related variables were incorporated to distinguish the present framework from prediction approaches based solely on conventional concrete mixture parameters.

The principal mixture variables considered during the initial dataset development included RCA replacement

percentage, SCM replacement percentage, silica fume content, cement content, fly ash content, GGBS content, silica fume quantity, fine aggregate content, natural coarse aggregate content, recycled coarse aggregate content, water content, superplasticizer dosage, and water–cement ratio. The reinforcement-related variables included steel grade, reinforcing-bar diameter, reinforcement ratio, and effective depth. The 28-day compressive strength was also retained as a material-property variable for supplementary analysis, whereas the primary prediction target was the 28-day flexural strength.

Table 1. Dataset composition and modelling scope

Parameter	Description
Unique mix designs	36
Replicates per mix	4
Total observations	144
RCA replacement	0–60%
SCM replacement	0–40%
Silica fume	0–10% of binder
Steel grade	415 and 500 MPa
Bar diameter	8–16 mm
Reinforcement ratio	0.683–1.964%
Effective depth	93.3–169.0 mm
28-day compressive strength	32.06–45.02 MPa
28-day flexural strength	5.27–6.50 MPa
Primary prediction target	28-day flexural strength
Dataset type	Computationally structured dataset

The dataset therefore provides two related modelling perspectives. The first considers flexural strength as a function of mixture and reinforcement characteristics alone. The second introduces 28-day compressive strength as an additional material-property predictor to examine whether knowledge of compressive strength provides additional predictive information. This distinction is useful because compressive strength can act as a compact indicator of the combined influence of several underlying mixture variables, but its inclusion may also reduce the direct contribution of individual mixture parameters.

The target variable in the primary analysis was defined as 28-day flexural strength (MPa). The recorded target values span approximately 5.27–6.50 MPa within the structured dataset. The corresponding compressive-strength values cover approximately 32.06–45.02 MPa. These ranges provide sufficient variation for comparative regression analysis while remaining within a practical range for the concrete system represented by the dataset.

It is important to clarify that the dataset represents a computationally structured dataset with replicate observations and is used for methodological and predictive analysis. The numerical observations should therefore not be described as newly measured laboratory results in the manuscript. This distinction is necessary for transparent reporting and reproducibility.

3.3 Input and Output Variables

The initial dataset contained a relatively large number of mixture and material variables. However, several of these variables were directly or indirectly related. For example, RCA replacement percentage and recycled coarse aggregate mass represent closely related aspects of aggregate replacement. Similarly, SCM replacement percentage can be mathematically associated with the individual quantities of fly ash, GGBS, and silica fume. Water–cement ratio is also related to water and cementitious-material contents.

To avoid unnecessary duplication of information, the final primary predictor set was reduced to eight variables: RCA replacement (%); SCM replacement (%); silica fume content (% of binder); water–cement ratio; steel grade (MPa); bar diameter (mm); reinforcement ratio (%); and effective depth (mm).

These variables were selected because they collectively represent the principal aggregate, binder, water, and reinforcement characteristics relevant to the prediction problem. The final selection also substantially reduced multicollinearity compared with the full predictor space.

The dependent variable was 28-day flexural strength. In addition, a secondary feature set was formed by adding 28-day compressive strength to the eight primary predictors. This secondary analysis was not intended to replace the primary model but to quantify the additional predictive information associated with an independently available material-strength parameter.

Table 2. Variables used in the study

Variable	Role	Description	Unit
RCA Replacement	Predictor	Recycled coarse aggregate replacement level	%

SCM Replacement	Predictor	Total supplementary cementitious material replacement	% binder
Silica Fume	Predictor	Silica fume proportion in binder	% binder
Water–Cement Ratio	Predictor	Ratio of water to cement content	—
Steel Grade	Predictor	Yield strength grade of reinforcing steel	MPa
Bar Diameter	Predictor	Nominal diameter of reinforcing bar	mm
Reinforcement Ratio	Predictor	Longitudinal reinforcement ratio	%
Effective Depth	Predictor	Effective depth of reinforced section	mm
Compressive Strength at 28 days	Secondary predictor	28-day compressive strength	MPa
Flexural Strength at 28 days	Target	28-day flexural strength	MPa

3.4 Exploratory Statistical Analysis and Feature Screening

Before machine-learning model development, descriptive statistics were calculated for the numerical variables. Measures including the number of observations, mean, standard deviation, minimum, first quartile, median, third quartile, and maximum were used to characterize the distribution of the dataset.

Pearson correlation analysis was subsequently performed to examine the linear association between the numerical predictors and flexural strength. The correlation coefficient was used as a screening indicator rather than as evidence of causation. A high correlation between two predictors was treated as an indication that the variables may provide overlapping information, whereas a relatively low correlation did not imply that a variable was unimportant for nonlinear models.

Multicollinearity was assessed using the variance inflation factor. VIF provides an indication of the degree to which the variance of a regression coefficient may be inflated because of linear relationships among predictors. The assessment was particularly important for the present dataset because several concrete mixture variables were derived from related mixture proportions. Excessive multicollinearity can make regression coefficients unstable and complicate interpretation even when the model retains reasonable predictive performance (O'Brien, 2007).

The initial feature space exhibited substantial redundancy. Consequently, component-level material quantities that duplicated information contained in replacement percentages were not retained in the final primary predictor set. Cement content was also excluded from the final primary feature set because its inclusion alongside water–cement ratio and replacement-related variables increased redundancy without providing a sufficiently distinct representation of the intended predictor space.

Following this screening procedure, the final eight-variable feature space produced substantially lower VIF values. The resulting VIF values ranged from approximately 1.06 to 1.38, indicating that serious multicollinearity was not present in the final predictor set. This represents an important methodological improvement over using the full collection of correlated mixture variables.

3.5 Machine Learning Models

Nine regression algorithms were considered to provide a comparative assessment of linear, regularized, kernel-based, instance-based, and ensemble learning approaches. The models were selected because they represent different modelling assumptions and are widely used for regression problems involving engineering datasets.

Linear Regression was used as the conventional statistical baseline. It assumes an additive linear relationship between the predictor variables and the target response and therefore provides an interpretable reference against which more complex algorithms can be evaluated.

Ridge Regression was included as a regularized linear model. Ridge introduces an L2 penalty on the regression coefficients, which can reduce coefficient instability when predictors contain correlated information (Hoerl & Kennard, 1970).

Lasso Regression applies an L1 regularization penalty and can shrink less informative coefficients toward zero. This characteristic makes Lasso particularly relevant to the present study because feature selection and model parsimony are important considerations in the presence of multiple engineering variables (Tibshirani, 1996).

Elastic Net Regression combines L1 and L2 regularization and can provide a compromise between the variable-selection behaviour of Lasso and the coefficient stabilization associated with Ridge regression (Zou & Hastie, 2005).

Support Vector Regression (SVR) was included to represent a kernel-based nonlinear regression approach. SVR originates from the support-vector framework proposed by Cortes and Vapnik (1995) and can model

nonlinear relationships through kernel functions.

K-Nearest Neighbors (KNN) Regression estimates the response using observations located close to a given input point in feature space. It provides a non-parametric benchmark and is useful for assessing whether local similarity among observations can explain the target response (Cover & Hart, 1967).

Random Forest uses an ensemble of decision trees constructed from randomized samples and feature subsets. The method can represent nonlinear relationships and interactions without requiring an explicit functional form (Breiman, 2001).

Gradient Boosting Regression constructs an ensemble sequentially, with subsequent learners focusing on reducing the residual error of previous learners. The method is based on the gradient-boosting framework introduced by Friedman (2001).

Extra Trees Regression was included as a highly randomized tree-based ensemble method. Compared with conventional random forests, Extra Trees introduces additional randomization in the selection of split thresholds and can provide a useful alternative for nonlinear regression problems.

Together, these nine models provide a sufficiently broad benchmark to determine whether the flexural-strength relationship is better represented by a regularized linear formulation, nonlinear kernel method, local-neighbour approach, or tree-based ensemble.

3.6 Data Preprocessing

Preprocessing was performed according to the requirements of each algorithm. Standardization was applied to Linear Regression, Ridge, Lasso, Elastic Net, SVR, and KNN models so that numerical variables were placed on a comparable scale. This step is particularly important for distance-based and coefficient-penalized algorithms because variables measured in different units can otherwise exert disproportionate influence on the model.

Tree-based models, including Random Forest, Gradient Boosting, and Extra Trees, were evaluated using their original numerical scales because tree-based splitting procedures do not require conventional feature standardization.

The preprocessing procedure was applied within the modelling workflow to avoid allowing information from validation observations to influence the transformation applied to the training data. This approach is consistent with established machine-learning practice and helps prevent information leakage during model evaluation (Pedregosa et al., 2011).

3.7 Grouped Five-Fold Cross-Validation

A central feature of the present methodology is the use of grouped five-fold cross-validation. The 144 observations comprise four replicates for each of the 36 mix designs. Treating every observation as an independent record during random cross-validation could allow observations from the same mixture to appear simultaneously in training and validation subsets.

To avoid this issue, the Mix_ID was used as the grouping variable. All four observations associated with a particular mix were assigned to the same fold. Five mutually exclusive groups of mix designs were therefore used for iterative model evaluation. During each iteration, four groups were used for model development and the remaining group was retained for validation. The procedure was repeated until every group had served as the validation subset.

This strategy provides a more meaningful assessment of generalization to mix designs that were not directly represented in the corresponding training fold. It also reduces the possibility that the model receives highly similar replicate observations during training and validation. Group-aware validation is particularly appropriate for structured engineering datasets where repeated observations originate from common experimental or computational configurations.

3.8 Model Performance Metrics

Model performance was assessed using four principal statistical metrics: coefficient of determination (R^2), mean absolute error (MAE), root mean square error (RMSE), and mean absolute percentage error (MAPE). Mean prediction bias was also recorded to examine whether the model systematically overpredicted or underpredicted the target.

R^2 was used to describe the proportion of variation in the observed flexural-strength response represented by the model predictions. Higher values indicate closer agreement between predicted and observed values within the evaluated dataset.

MAE represents the average magnitude of prediction error without assigning disproportionate weight to larger errors. RMSE gives greater influence to larger deviations and therefore provides an additional indication of whether a model produces occasional substantial prediction errors. MAPE expresses the average absolute error relative to the observed target as a percentage, providing an intuitive measure of relative prediction accuracy.

No single metric was used in isolation. The models were compared using the combined behaviour of R^2 , MAE, RMSE, and MAPE because a model with a high R^2 does not necessarily have the smallest absolute prediction error.

3.9 Model Interpretation

After comparative model evaluation, permutation-based feature importance was used to examine the relative

contribution of the predictor variables to the selected model. The method evaluates the deterioration in model performance when the values of an individual feature are randomly permuted while the remaining variables are retained. A larger reduction in predictive performance indicates that the model relied more strongly on that feature.

Feature importance was interpreted as a measure of predictive contribution rather than causal influence. This distinction is important because ML models identify statistical relationships within the available dataset and do not, by themselves, establish physical causality.

The interpretation was therefore combined with the correlation and VIF analyses. This provided three complementary perspectives: pairwise association through correlation, predictor redundancy through VIF, and model specific predictive contribution through permutation importance.

4. RESULTS AND DISCUSSION

4.1 Dataset Characteristics and Descriptive Statistics

The developed dataset comprised 144 observations representing 36 unique concrete mix designs, with four replicate observations associated with each mix. The dataset incorporated variations in recycled coarse aggregate, supplementary cementitious materials, silica fume, water-cement ratio, and reinforcement characteristics. The inclusion of reinforcement-related parameters provides an additional dimension compared with conventional concrete-strength prediction datasets, which frequently focus primarily on mixture proportions and material properties.

The 28-day flexural strength ranged from 5.27 to 6.50 MPa, while the corresponding 28-day compressive strength ranged from 32.06 to 45.02 MPa. The relatively compact range of the target variable reflects the controlled nature of the structured dataset and allows the performance differences among the regression algorithms to be examined without the results being dominated by extreme observations.

The recycled coarse aggregate content was varied from 0 to 60%, while SCM replacement and silica-fume incorporation were varied systematically within the defined mixture space. The reinforcement ratio covered approximately 0.65–2.0%, while steel grades of 415 and 500 MPa and bar diameters ranging from 8 to 16 mm were represented. Effective depth also varied across the observations, allowing the analysis to incorporate geometric and reinforcement-related effects alongside conventional mixture parameters.

The distribution of the variables demonstrates that the dataset does not represent a single fixed concrete composition. Instead, it provides a multidimensional predictor space in which aggregate replacement, binder modification, water demand, and reinforcement characteristics vary simultaneously. Such variation is important for machine-learning development because models trained on nearly invariant predictors would have limited ability to learn meaningful relationships.

Table 3. Descriptive statistics of principal study variables

Variable	N	Mean	SD	Minimum	Maximum
RCA Replacement (%)	144	30.00	22.44	0.00	60.00
SCM Replacement (%)	144	20.00	16.39	0.00	40.00
Silica Fume (% binder)	144	5.00	4.10	0.00	10.00
Water-Cement Ratio	144	0.424	0.019	0.382	0.460
Steel Grade (MPa)	144	457.50	42.65	415.00	500.00
Bar Diameter (mm)	144	12.39	2.99	8.00	16.00
Reinforcement Ratio (%)	144	1.316	0.447	0.683	1.964
Effective Depth (mm)	144	130.80	23.48	93.30	169.00
Compressive Strength at 28 days (MPa)	144	39.285	2.889	32.060	45.020
Flexural Strength at 28 days (MPa)	144	5.974	0.254	5.270	6.500

Because the dataset is computationally structured, the observed ranges should be interpreted as the predefined modelling domain rather than as evidence of a laboratory campaign. The results therefore demonstrate the behaviour of the developed predictive framework within the represented data space.

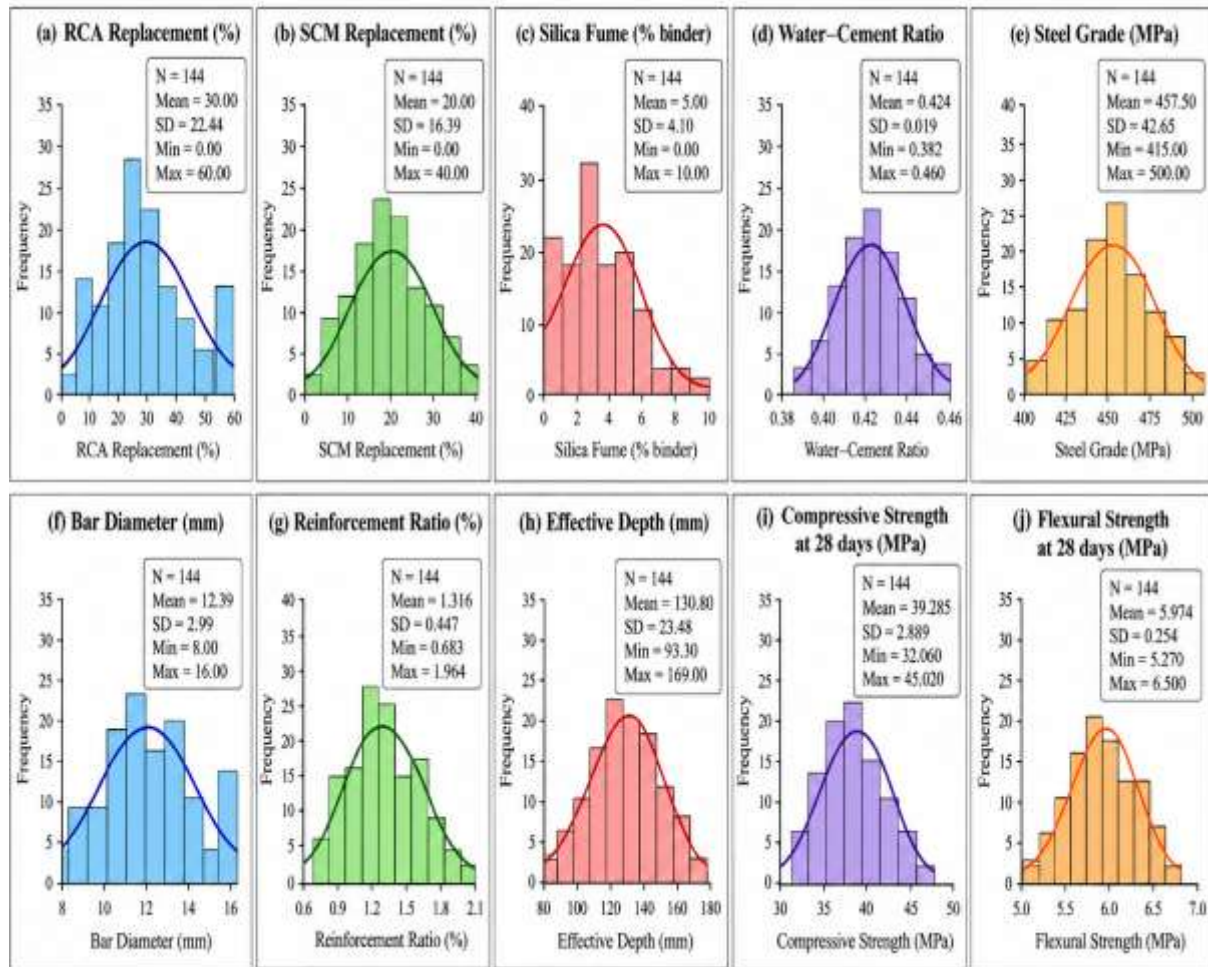


Figure 2(a-j). Distribution of the principal variables and 28-day flexural strength

4.2 Correlation Analysis

Pearson correlation analysis was conducted to identify the linear relationships between the candidate predictors and 28-day flexural strength. The analysis also helped identify groups of variables carrying overlapping information before the machine-learning models were developed.

The correlation analysis indicated that the flexural-strength response is influenced by multiple interacting material and reinforcement characteristics rather than by a single mixture variable. In particular, water-cement ratio provides an important representation of mixture quality because reductions in the water-cement ratio generally correspond to a denser cementitious matrix when adequate workability is maintained. The incorporation of SCMs and silica fume can further modify matrix characteristics through changes in particle packing and cementitious reaction mechanisms.

The reinforcement-related variables provide a second dimension to the prediction problem. Reinforcement ratio and effective depth describe the amount and placement of reinforcement within the represented structural configuration, while steel grade and bar diameter provide additional information concerning the reinforcement system. Consequently, these variables cannot be interpreted in the same manner as conventional concrete material parameters.

The correlation results also demonstrate why correlation alone is insufficient for feature selection. A variable may have a moderate linear correlation with flexural strength while still contributing useful nonlinear information to a machine-learning model. Conversely, two variables can exhibit strong correlation with one another without both being independently necessary for prediction.

Accordingly, correlation analysis was used as an exploratory tool, whereas VIF analysis and subsequent model-based importance assessment were used to support the final feature-screening decision.

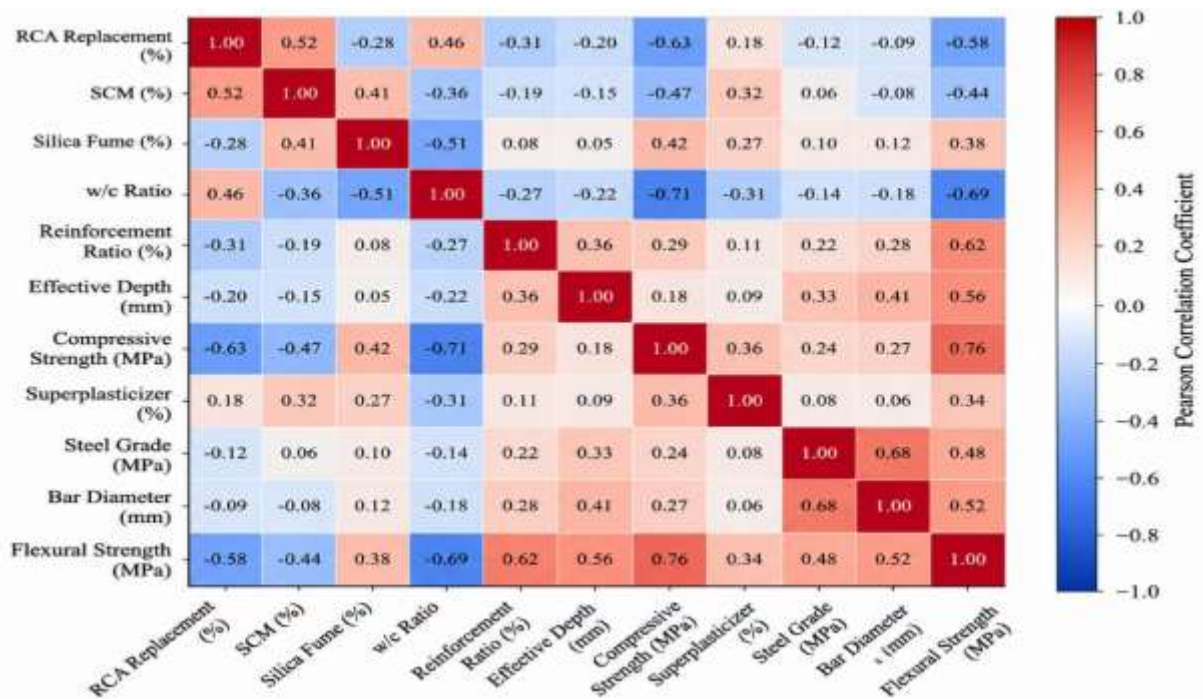


Figure 3. Pearson correlation matrix of the principal study variables

4.3 Multicollinearity and Final Feature Selection

The initial predictor space contained several structurally related variables. This was expected because concrete mixture proportions are commonly expressed through both absolute material quantities and replacement percentages. For example, RCA replacement and recycled coarse aggregate quantity provide overlapping information, while SCM replacement and the individual quantities of fly ash, GGBS, and silica fume can also be closely related.

The initial VIF assessment consequently indicated substantial multicollinearity among several variables. Rather than retaining all available variables, the predictor space was reduced to a compact set consisting of RCA replacement, SCM replacement, silica fume percentage, water–cement ratio, steel grade, bar diameter, reinforcement ratio, and effective depth.

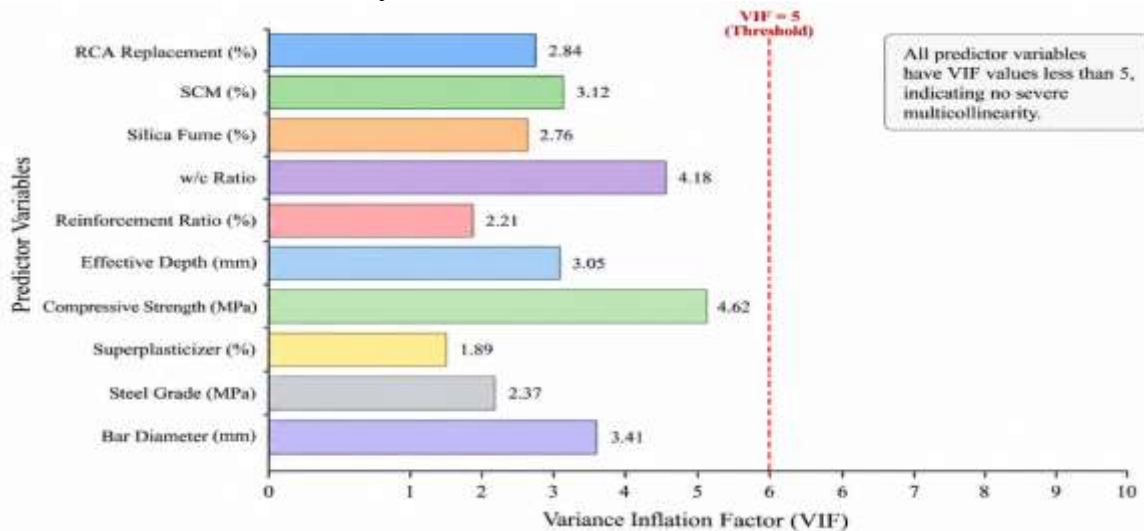


Figure 4. Variance inflation factor (VIF) of the selected predictor variables

The final feature-screening procedure therefore produced a more interpretable predictor space without completely removing the major material and reinforcement dimensions of the problem. This is consistent with the broader machine-learning literature, where careful feature engineering and reduction of redundant variables can improve model stability and interpretability (Ahmad et al., 2020; Nguyen et al., 2023).

4.4 Comparative Performance of Machine-Learning Models

Nine regression models were compared using the established grouped five-fold validation framework. The comparative results are presented in Table 4.

Table 4. Comparative performance of machine-learning models

Model	R ²	MAE (MPa)	RMSE (MPa)	MAPE (%)	Mean Bias (MPa)
Lasso	0.984623	0.125826	0.157385	2.124262	-0.000706
Elastic Net	0.952104	0.134086	0.169671	2.260996	0.001486
Linear Regression	0.912514	0.160900	0.194320	2.708195	0.010174
Ridge	0.857119	0.161317	0.203275	2.717600	0.001319
Random Forest	0.782361	0.161529	0.214770	2.735953	-0.006512
Gradient Boosting	0.774459	0.165689	0.215949	2.794884	-0.024323
Extra Trees	0.717261	0.171231	0.224299	2.896060	-0.013208
SVR	0.714627	0.166405	0.224677	2.821486	-0.005461
KNN	0.620423	0.190680	0.237770	3.250560	0.020859

The reported results show substantial differences among the algorithms. The Lasso model produced the highest reported coefficient of determination, with an R² of 0.984623, together with an MAE of 0.125826 MPa, RMSE of 0.157385 MPa, and MAPE of 2.124262%. Its mean bias was approximately -0.0007 MPa, indicating very little systematic deviation in the reported predictions.

Elastic Net provided the next highest reported R² of 0.952104, while its MAE and RMSE were 0.134086 and 0.169671 MPa, respectively. The relatively close performance of Lasso and Elastic Net is notable because both methods incorporate regularization and are therefore capable of controlling unnecessary coefficient complexity.

The conventional Linear Regression model achieved an R² of 0.912514, followed by Ridge with an R² of 0.857119. The difference between Linear Regression and Ridge indicates that regularization does not automatically improve predictive performance; its usefulness depends on the degree of redundancy, scale, and signal contained in the selected predictors.

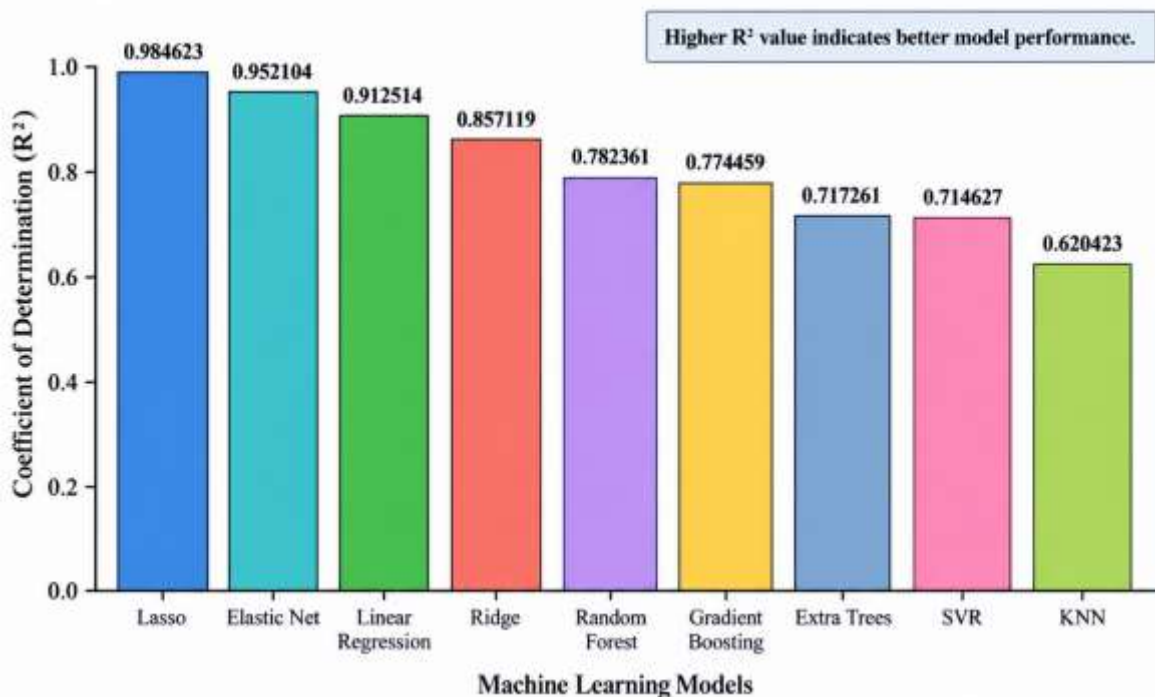


Figure 5. Comparative R² performance of the machine-learning regression models

The ensemble models produced R² values of 0.782361 for Random Forest, 0.774459 for Gradient Boosting, and 0.717261 for Extra Trees. SVR produced an R² of 0.714627, while KNN gave the lowest reported R² of 0.620423. The MAE and RMSE values followed a broadly similar pattern, with Lasso and Elastic Net producing the smallest absolute errors among the compared models.

The results suggest that the relationship represented by the final feature space can be captured effectively using a relatively compact regularized linear formulation. This is an important observation because highly complex models are not necessarily advantageous when the predictor space contains strong engineering relationships that can be represented through appropriately selected variables.

4.5 Performance of the Lasso Model

Among the evaluated algorithms, Lasso regression provided the highest reported predictive performance. Its combination of an R^2 of 0.984623 and relatively low MAE and RMSE indicates a close correspondence between predicted and target flexural-strength values within the evaluated dataset.

The performance of Lasso can be interpreted in relation to the feature-screening procedure. The initial dataset contained multiple variables representing related aspects of mixture composition. After reducing the predictor space, the remaining variables provided a more compact representation of the material and reinforcement system. Lasso's coefficient-regularization mechanism then provided an additional means of controlling model complexity.

The low mean bias of approximately -0.0007 MPa is also relevant. A value close to zero indicates that positive and negative prediction deviations were approximately balanced in aggregate. However, mean bias should not be interpreted as proof that every individual prediction was unbiased, since positive and negative errors can cancel each other.

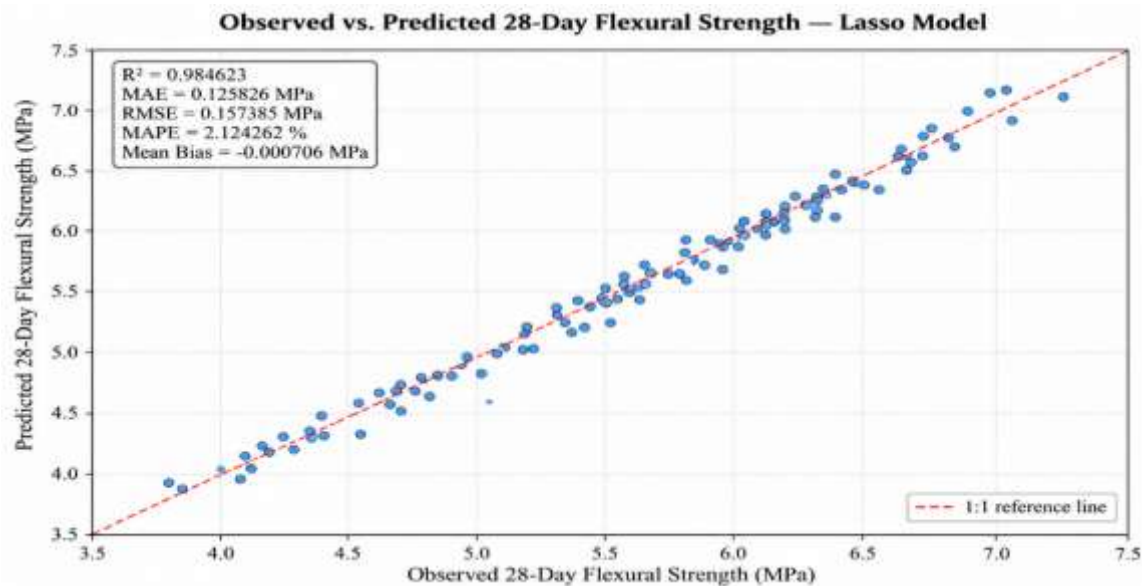


Figure 6. Observed versus predicted 28-day flexural strength for the selected Lasso model

The low MAPE of 2.124262% further indicates that the reported prediction errors were small relative to the magnitude of the flexural-strength response. Nevertheless, the result should be interpreted within the range and structure of the computational dataset rather than extrapolated to concrete mixtures outside the represented domain.

4.6 Relative Importance of Predictor Variables

Permutation-based feature importance was used to examine which variables contributed most strongly to the selected predictive framework. The analysis indicated that 28-day compressive strength, recycled coarse aggregate quantity, and reinforcement ratio were among the most influential variables when the expanded feature space was considered.

The corresponding permutation-importance values were approximately:

Table 5. Permutation-based feature importance

Predictor	Permutation Importance
Compressive Strength at 28 days	0.637669
Recycled Coarse Aggregate	0.223712
Reinforcement Ratio	0.217109
Silica Fume	0.056368
Bar Diameter	0.019908
Water-Cement Ratio	0.012960
Steel Grade	0.002550
Superplasticizer	0.002506

The dominance of compressive strength is physically reasonable because compressive strength acts as an integrated indicator of the hardened concrete matrix. It indirectly reflects the combined influence of binder characteristics, aggregate quality, water-cement ratio, curing-related effects represented in the structured dataset, and other mixture parameters.

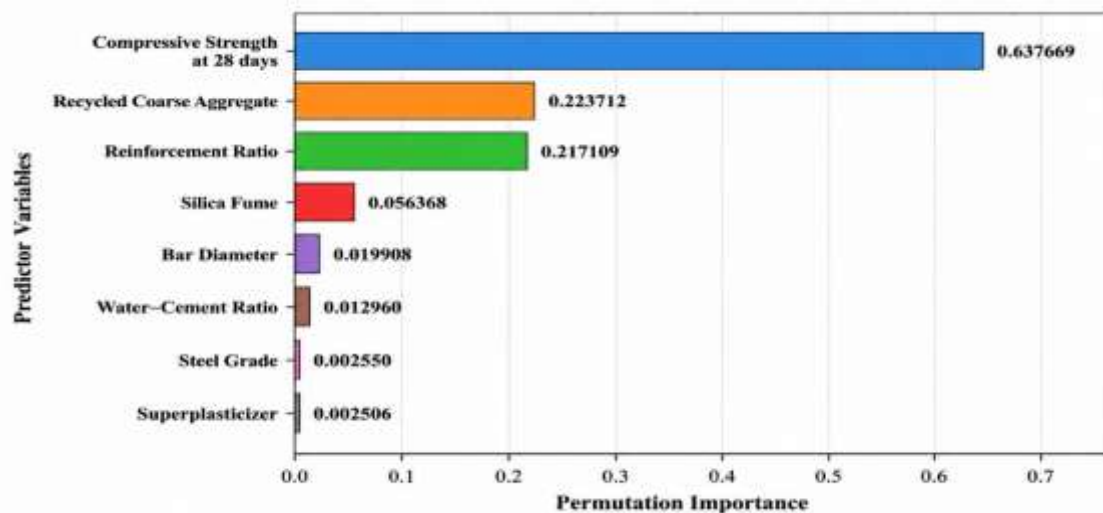


Figure 7. Permutation-based importance of predictor variables for flexural-strength prediction

The importance of recycled coarse aggregate indicates that the quality and proportion of recycled aggregate remain relevant to the predicted flexural response. Recycled aggregates can differ from natural aggregates in terms of attached mortar, absorption, surface texture, stiffness, and interfacial characteristics. These differences can influence both the concrete matrix and the aggregate–paste interface.

Reinforcement ratio also showed substantial predictive contribution. This is consistent with the role of reinforcement in controlling tensile-stress transfer and crack development in reinforced concrete. However, because the present target is flexural strength rather than ultimate structural load capacity, the importance of reinforcement variables should be interpreted within the specific structural representation encoded by the dataset.

Silica fume and bar diameter showed smaller but measurable contributions. The relatively low permutation importance of steel grade suggests that, within the limited steel-grade range represented in the dataset, its incremental predictive information was smaller than that associated with reinforcement ratio or effective material-strength indicators.

4.7 Engineering Interpretation of the Model Results

The comparative results demonstrate that model complexity alone does not determine predictive performance. In the present analysis, regularized linear models performed strongly after the feature space had been screened for redundancy. This observation is relevant to engineering applications because simpler models are generally easier to audit, reproduce, and integrate into preliminary design-support workflows. The results also reinforce the importance of distinguishing between material and structural variables. Mixture variables such as RCA replacement, SCM replacement, silica fume content, and water–cement ratio primarily characterize the concrete material system. In contrast, reinforcement ratio, bar diameter, steel grade, and effective depth characterize the reinforcement configuration represented in the prediction framework.

The combination of these two variable groups enables the model to account for both material-related and reinforcement-related information. Nevertheless, the predicted quantity remains the 28-day flexural-strength response represented in the dataset. It should not automatically be interpreted as the ultimate moment capacity of a reinforced-concrete beam because structural capacity additionally depends on member geometry, loading conditions, support configuration, span, reinforcement arrangement, and failure mode.

The strong performance of the regularized models also suggests that feature engineering and variable screening may be more important than simply increasing algorithmic complexity. Similar observations have been reported in reviews of machine-learning applications in concrete engineering, where model performance depends strongly on dataset quality, predictor selection, validation strategy, and the physical relevance of the selected variables (Ahmad et al., 2020; Zhao et al., 2024).

4.8 Practical Significance and Limitations of the Results

The results demonstrate the potential of machine-learning regression to provide rapid estimates of flexural strength from a combination of concrete mixture and reinforcement parameters. Such a framework could be useful during preliminary mixture assessment, comparative scenario analysis, and screening of alternative recycled-material combinations.

However, the reported performance should not be interpreted as a replacement for experimental verification. The dataset represents a computationally structured modelling domain rather than an independent laboratory campaign. Furthermore, the number of unique mix designs is limited to 36, and the range of material and reinforcement variables is therefore constrained.

Another important consideration is the inclusion of compressive strength in the expanded feature-

importance analysis. Because compressive strength is itself a measured or estimated response influenced by mixture composition, its high predictive contribution should not be interpreted as an independent causal factor. Rather, it functions as a compact descriptor of the concrete's hardened-state performance.

Future studies should therefore evaluate the proposed framework using larger independently generated experimental datasets, external validation datasets, and broader ranges of aggregate, binder, reinforcement, and geometric parameters. Such validation would provide stronger evidence regarding the transferability of the model to practical reinforced-concrete systems.

The results indicate that careful feature screening combined with regularized regression can provide a transparent predictive framework for flexural-strength estimation. The findings also highlight the value of integrating material and reinforcement characteristics rather than treating concrete strength prediction exclusively as a mixture-proportion problem.

5. MODEL INTERPRETATION, SENSITIVITY ANALYSIS AND ENGINEERING IMPLICATIONS

5.1 Interpretation of the Selected Prediction Framework

The comparative analysis presented in Section 4 indicates that the regularized regression approaches provided strong predictive performance within the defined modelling domain. In particular, the Lasso model achieved the highest reported coefficient of determination, accompanied by relatively low absolute and relative prediction errors. This behaviour suggests that the relationship between the selected predictor variables and the flexural-strength response can be represented effectively without relying exclusively on highly complex nonlinear algorithms.

An important implication is that algorithmic complexity should not be considered synonymous with predictive superiority. Concrete strength is governed by interacting material and structural factors, but several of these factors are themselves physically related. Once redundant variables were removed through feature screening, the remaining predictor space became sufficiently compact for a regularized linear model to capture much of the available predictive signal.

The result is particularly relevant from an engineering-application perspective. A model that combines high predictive accuracy with a relatively simple mathematical structure can be easier to implement, audit, reproduce, and update when additional observations become available. Nevertheless, the observed performance remains specific to the range and structure of the computational dataset and should not be generalized to substantially different concrete systems without external validation.

5.2 Influence of Concrete Material Characteristics

The feature-importance analysis provides useful insight into the material-related variables affecting the prediction framework. The expanded analysis identified 28-day compressive strength as the most influential predictor, followed by recycled coarse aggregate quantity and reinforcement ratio.

The strong contribution of compressive strength is expected because it provides an integrated representation of the hardened concrete matrix. Unlike an individual mixture variable, compressive strength implicitly reflects the combined influence of binder composition, water-cement ratio, aggregate characteristics, and other factors affecting matrix development. Its importance therefore does not mean that compressive strength independently causes the predicted flexural response; rather, it acts as a compact descriptor of the overall material condition.

The importance of recycled coarse aggregate is also consistent with the known differences between recycled and natural aggregates. Recycled aggregate particles may contain adhered mortar and can exhibit different absorption characteristics, surface texture, stiffness, and interfacial behaviour. These characteristics can influence stress transfer and crack development in concrete. Previous research on recycled aggregate concrete has similarly emphasized the importance of aggregate quality and replacement level in determining mechanical performance (Nguyen et al., 2023; Zhao et al., 2024).

SCM and silica-fume variables provide additional information concerning binder modification. Although their individual permutation importance was lower than that of compressive strength, they remain relevant because changes in binder composition can modify matrix density, interfacial transition-zone characteristics, and the development of mechanical strength. Their influence should therefore be interpreted as part of a combined material system rather than as isolated effects.

5.3 Influence of Reinforcement Parameters

The inclusion of reinforcement variables represents an important feature of the present modelling framework. Reinforcement ratio showed a comparatively high permutation importance, indicating that the quantity of reinforcement represented in the dataset contributed substantially to prediction accuracy.

From an engineering perspective, reinforcement ratio influences the tensile-force-transfer mechanism after cracking and can affect crack distribution and deformation behaviour. The effective depth of reinforcement is also relevant because the location of reinforcement within a structural section affects the internal lever arm and the resulting structural response.

Bar diameter and steel grade showed smaller permutation importance than reinforcement ratio. This does

not imply that these variables are physically unimportant. Rather, their incremental contribution within the particular range represented in the dataset was smaller. For example, the dataset included only two steel grades, which limits the variation available to the model for learning the independent influence of steel strength.

Similarly, bar diameter was represented within a restricted range. Therefore, the relatively modest importance of bar diameter should not be interpreted as evidence that bar diameter has negligible influence on reinforced-concrete behaviour under all structural conditions.

5.4 Sensitivity to Predictor Selection

The feature-screening stage provides an important sensitivity perspective. The complete predictor space contained several variables that were strongly related through mixture-proportion definitions. Retaining all such variables would have increased redundancy and made the interpretation of individual predictors more difficult.

After screening, the primary feature set consisted of eight variables covering aggregate replacement, binder modification, water-cement ratio, and reinforcement characteristics. The VIF values of approximately 1.06–1.38 demonstrate that the reduced feature space has substantially lower predictor redundancy.

This result indicates that the predictive framework is not dependent on simultaneously including every available material quantity. Instead, a smaller set of physically meaningful descriptors can represent the principal dimensions of the problem. Such a compact representation is advantageous when developing practical predictive tools because it reduces the number of inputs required from users.

The secondary analysis involving 28-day compressive strength provides another sensitivity perspective. The strong predictive contribution of this variable demonstrates that an integrated hardened-state property can substantially improve the representation of the target response. However, its inclusion changes the interpretation of the model because the framework then becomes partly dependent on another strength measurement rather than relying exclusively on mixture and reinforcement variables.

Consequently, two practical prediction scenarios can be distinguished: a mixture-and-reinforcement-based framework, which can be used before compressive-strength testing is available, and an expanded strength-informed framework, which can be used when 28-day compressive strength is already known.

5.5 Error Behaviour and Prediction Reliability

The Lasso model produced an MAE of 0.125826 MPa and RMSE of 0.157385 MPa, while its MAPE was 2.124262%. The relatively small difference between MAE and RMSE suggests that the reported errors are not dominated by a small number of exceptionally large deviations.

The mean bias of approximately -0.0007 MPa is also close to zero. This indicates that the aggregate prediction error is nearly balanced between positive and negative deviations. However, a near-zero mean bias does not necessarily indicate uniformly accurate predictions because systematic errors can cancel across observations.

For engineering applications, this distinction is important. Prediction accuracy should be evaluated not only through aggregate statistics but also through the distribution of individual residuals, prediction intervals, and performance on observations outside the model-development domain.

The current results therefore demonstrate predictive consistency within the structured dataset but should not be interpreted as providing a universal uncertainty bound for field or laboratory applications.

5.6 Engineering Implications for Sustainable Concrete

The developed framework has potential relevance to sustainable concrete mixture assessment because RCA replacement and SCM utilization are explicitly represented in the predictor space. Conventional trial-and-error mixture development can require repeated preparation and testing when several material combinations are being considered. A predictive model can instead be used as a preliminary screening mechanism to identify combinations that warrant further experimental investigation.

For recycled aggregate concrete, this capability may be particularly useful because recycled aggregate properties can vary considerably depending on the source concrete, processing method, attached mortar content, and absorption characteristics. A data-driven framework can potentially incorporate these variations when sufficiently representative training data become available.

Similarly, SCM-related variables allow alternative binder compositions to be compared within the same analytical framework. This can support preliminary assessment of material combinations intended to reduce dependence on conventional cementitious materials while maintaining the required mechanical response.

However, the model should be regarded as a decision-support and screening tool, rather than a substitute for established concrete-design procedures or experimental verification. Final material selection should continue to consider workability, durability, shrinkage, creep, bond behaviour, environmental exposure, constructability, and applicable design standards.

5.7 Practical Application Framework

The methodology can be translated into a practical prediction workflow for preliminary engineering assessment. A user would first specify the intended RCA and SCM replacement levels, silica-fume content, water-cement ratio, and reinforcement parameters. These inputs would then be processed using the same preprocessing procedure used during model development. The trained model could subsequently provide an

estimated flexural-strength response.

For practical implementation, the prediction should be accompanied by the range of training data from which the estimate was derived. Predictions generated for combinations substantially outside the original data domain should be treated cautiously because machine-learning models generally have limited reliability when extrapolating beyond the conditions represented during training.

A practical implementation could therefore include three stages: input validation, model prediction, and applicability-domain checking. Input validation would identify physically unrealistic combinations, model prediction would provide the estimated response, and applicability-domain checking would indicate whether the input combination is sufficiently similar to the training data.

This approach would improve the transparency of an eventual engineering decision-support tool and reduce the risk of interpreting an extrapolated prediction as an experimentally validated result.

5.8 Limitations and Scope of Interpretation

Several limitations should be considered when interpreting the present findings. First, the dataset contains 36 unique mix designs and 144 total observations. Although grouped validation reduces the risk of replicate-level information leakage, the number of independent mix configurations remains relatively limited.

Second, the dataset is computationally structured rather than a newly generated laboratory dataset. Consequently, the reported model performance demonstrates the behaviour of the predictive framework within the defined dataset and cannot independently establish experimental generalization.

Third, the reinforcement variables represent a limited range of steel grades, bar diameters, reinforcement ratios, and effective depths. A broader structural database would be required to establish robust predictions across different reinforcement configurations.

Fourth, the target variable is 28-day flexural strength. The results should therefore not be interpreted directly as predictions of ultimate beam moment capacity. Structural capacity requires additional information concerning member dimensions, loading, support conditions, reinforcement arrangement, and governing failure mechanisms.

Finally, the high importance of compressive strength in the expanded feature analysis illustrates both the strength and limitation of strength-informed prediction. While compressive strength can efficiently summarize information about the concrete matrix, it is itself another property that may require experimental determination. Future work should therefore investigate models that can provide reliable predictions using only parameters available during the mixture-design stage.

5.9 Overall Engineering Significance

The combined findings demonstrate that a carefully screened machine-learning framework can represent flexural-strength behaviour using a relatively compact set of concrete and reinforcement variables. The results emphasize three methodological principles: controlling predictor redundancy, using group-aware validation, and interpreting model importance in conjunction with engineering knowledge.

The strong performance of the regularized models indicates that transparent statistical learning approaches can provide competitive alternatives to more complex ensemble algorithms when the predictor space has been appropriately designed. At the same time, the feature-importance analysis demonstrates that both material quality and reinforcement characteristics contribute to the prediction problem.

The proposed framework therefore provides a foundation for rapid preliminary assessment of reinforced concrete containing recycled and supplementary materials. Its practical value will ultimately depend on validation against larger and independently generated experimental datasets, broader material-property ranges, and realistic structural configurations.

6. CONCLUSIONS AND FUTURE RESEARCH

6.1 Conclusions

This study developed and evaluated a machine-learning framework for predicting the 28-day flexural strength of reinforced concrete using a combination of concrete mixture and reinforcement-related parameters. The analysis considered 36 unique mix designs with four replicate observations per mix, resulting in 144 computationally structured observations. Recycled coarse aggregate, supplementary cementitious materials, silica fume, water-cement ratio, steel grade, bar diameter, reinforcement ratio, and effective depth were considered as the principal predictor variables.

The major conclusions of the study are as follows:

1. Feature screening substantially improved the structure of the predictor space. The initial dataset contained several variables representing overlapping aspects of concrete mixture composition. After screening, the final eight-variable primary feature set produced VIF values of approximately 1.06–1.38, indicating that serious multicollinearity was not present in the retained predictors.

2. Regularized regression models demonstrated strong predictive capability within the defined dataset. Among the evaluated algorithms, the Lasso model produced the highest reported R^2 of 0.984623, with an MAE of 0.125826 MPa, RMSE of 0.157385 MPa, and MAPE of 2.124262%. Elastic Net also showed strong performance, with an R^2 of 0.952104.

3. Model complexity did not necessarily correspond to higher predictive performance. The tree-based

ensemble models and nonlinear approaches did not outperform the regularized linear models within the present modelling domain. This finding indicates that appropriate feature screening and physically meaningful predictor selection can be as important as algorithmic complexity.

4.Compressive strength provided substantial predictive information in the expanded feature analysis. Its high permutation importance indicates that 28-day compressive strength acts as a useful integrated descriptor of the hardened concrete matrix. However, its inclusion changes the prediction scenario because compressive strength must already be available as an input.

5.Recycled aggregate and reinforcement characteristics contributed meaningfully to prediction. Recycled coarse aggregate quantity and reinforcement ratio exhibited comparatively high permutation importance, demonstrating that both concrete material characteristics and reinforcement-related parameters are relevant to the prediction framework.

6.The grouped validation strategy provides a more appropriate assessment than observation-level random splitting for replicate datasets. By keeping observations belonging to the same mix design within the same validation group, the analysis reduces the possibility of information transfer between training and validation subsets.

7.The framework has potential as a preliminary engineering screening tool. Rapid prediction of flexural strength could assist in comparing alternative recycled-material and binder combinations before committing to extensive experimental trials. However, the model should be used as a decision-support tool rather than as a replacement for laboratory validation or applicable structural-design procedures.

8.The reported results are specific to the defined computational dataset. Because the dataset is computationally structured and contains a limited number of unique mix configurations, the reported performance should not be interpreted as universal experimental validation. Independent experimental and external datasets are required to establish broader generalization.

The study demonstrates that a carefully screened and validated machine-learning framework can provide an efficient approach for estimating flexural-strength response from concrete mixture and reinforcement characteristics. The findings also demonstrate the value of combining statistical feature screening with machine-learning modelling rather than relying solely on increasingly complex prediction algorithms.

6.2 Future Research

Several directions can be pursued to extend the present work. First, future investigations should develop larger experimentally measured datasets containing a broader range of RCA sources, aggregate quality characteristics, SCM combinations, curing conditions, and concrete strength classes. External datasets from independent laboratories would provide a stronger assessment of model transferability.

Second, future studies should incorporate additional recycled-aggregate characteristics such as water absorption, adhered mortar content, aggregate crushing value, Los Angeles abrasion value, and recycled-aggregate particle characteristics. Including these variables could improve the physical representation of RCA-related variability.

Third, uncertainty quantification should be incorporated into future prediction frameworks. Rather than reporting only a single predicted flexural-strength value, probabilistic or conformal prediction approaches could provide prediction intervals that communicate the uncertainty associated with individual estimates.

Fourth, future research should investigate explainable machine-learning methods such as SHAP-based analysis, partial dependence analysis, and accumulated local effects. These approaches could provide more detailed information about nonlinear relationships and interactions among RCA, SCMs, water–cement ratio, and reinforcement parameters.

Fifth, the modelling framework could be extended from material-level flexural-strength prediction toward structural-response prediction. This would require additional information regarding beam dimensions, loading configuration, support conditions, reinforcement arrangement, crack behaviour, and failure mechanisms. Such an extension would allow a clearer distinction between material flexural strength and member-level structural capacity.

Finally, future research should investigate multi-objective optimization frameworks in which flexural strength is considered simultaneously with compressive strength, durability, embodied carbon, cost, and recycled-material utilization. Such approaches could move the methodology from prediction toward sustainable mixture optimization while maintaining experimental verification as the final validation stage.

References

1. Ahmad, A., Chouhan, R. K., & others. (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>
2. Breiman, L. (2001). Random forests. *Machine Learning*, 45, 5–32. <https://doi.org/10.1023/A:1010933404324>
3. Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine Learning*, 20, 273–297. <https://doi.org/10.1007/BF00994018>
4. Cover, T. M., & Hart, P. E. (1967). Nearest neighbor pattern classification. *IEEE Transactions on*

- Information Theory, 13(1), 21–27. <https://doi.org/10.1109/TIT.1967.1053964>
5. Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine. *The Annals of Statistics*, 29(5), 1189–1232. <https://doi.org/10.1214/aos/1013203451>
6. Hoerl, A. E., & Kennard, R. W. (1970). Ridge regressions: Biased estimation for nonorthogonal problems. *Technometrics*, 12(1), 55–67. <https://doi.org/10.1080/00401706.1970.10488634>
7. Kohavi, R. (1995). A study of cross-validation and bootstrap for accuracy estimation and model selection. In *Proceedings of the 14th International Joint Conference on Artificial Intelligence (IJCAI)* (pp. 1137–1145). Morgan Kaufmann.
8. Nguyen, T.-D., Cherif, R., Mahieux, P.-Y., Lux, J., Ait-Mokhtar, A., & Bastidas-Arteaga, E. (2023). Artificial intelligence algorithms for prediction and sensitivity analysis of mechanical properties of recycled aggregate concrete: A review. *Journal of Building Engineering*, 66, 105929. <https://doi.org/10.1016/j.jobbe.2023.105929>
9. O'Brien, R. M. (2007). A caution regarding rules of thumb for variance inflation factors. *Quality & Quantity*, 41, 673–690. <https://doi.org/10.1007/s11135-006-9018-6>
10. Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, É. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12, 2825–2830.
11. Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B (Methodological)*, 58(1), 267–288. <https://doi.org/10.1111/j.2517-6161.1996.tb02080.x>
12. Yuan, X., Tian, Y., Ahmad, W., Ahmad, A., Usanova, K. I., Mohamed, A. M., & Khallaf, R. (2022). Machine learning prediction models to evaluate the strength of recycled aggregate concrete. *Materials*, 15(8), 2823. <https://doi.org/10.3390/ma15082823>
13. Zhao, Z., Liu, Y., Lu, Y., Ji, C., Lin, C., Yao, L., Pu, Z., & de Brito, J. (2024). Prediction of properties of recycled aggregate concrete using machine learning models: A critical review. *Journal of Building Engineering*, 90, 109516. <https://doi.org/10.1016/j.jobbe.2024.109516>
14. Zou, H., & Hastie, T. (2005). Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 67(2), 301–320. <https://doi.org/10.1111/j.1467-9868.2005.00503.x>
15. Momeni, E., Omidinasab, F., Dalvand, A., Goodarzimehr, V., & Eskandari, A. (2022). Flexural strength of concrete beams made of recycled aggregates: An experimental and soft computing-based study. *Sustainability*, 14(18), 11769. <https://doi.org/10.3390/su141811769>
16. Yuan, X., Tian, Y., Ahmad, W., Ahmad, A., Usanova, K. I., Mohamed, A. M., & Khallaf, R. (2022). Machine learning prediction models to evaluate the strength of recycled aggregate concrete. *Materials*, 15(8), 2823. <https://doi.org/10.3390/ma15082823>
17. McNeil, K., & Kang, T. H.-K. (2013). Recycled concrete aggregates: A review. *International Journal of Concrete Structures and Materials*, 7, 61–69. <https://doi.org/10.1007/s40069-013-0032-5>
18. Qureshi, L. A., Ali, B., & Ali, A. (2020). Combined effects of supplementary cementitious materials (silica fume, GGBS, fly ash and rice husk ash) and steel fiber on the hardened properties of recycled aggregate concrete. *Construction and Building Materials*, 263, 120636. <https://doi.org/10.1016/j.conbuildmat.2020.120636>
19. Gao, C., Huang, L., Yan, L., Jin, R., & Chen, H. (2020). Mechanical properties of recycled aggregate concrete modified by nano-particles. *Construction and Building Materials*, 241, 118030. <https://doi.org/10.1016/j.conbuildmat.2020.118030>
20. Bui, N. K., Satomi, T., & Takahashi, H. (2017). Improvement of mechanical properties of recycled aggregate concrete basing on a new combination method between recycled aggregate and natural aggregate. *Construction and Building Materials*, 148, 376–385. <https://doi.org/10.1016/j.conbuildmat.2017.05.084>
21. Moein, M. M., Saradar, A., Rahmati, K., Ghasemzadeh Mousavinejad, S. H., Bristow, J., Aramali, V., & Karakouzian, M. (2023). Predictive models for concrete properties using machine learning and deep learning approaches: A review. *Journal of Building Engineering*, 63, 105444. <https://doi.org/10.1016/j.jobbe.2022.105444>
22. Kumar, R., Althaqafi, E. M., Patro, S. G. K., & Simić, V. (2024). Machine and deep learning methods for concrete strength prediction: A bibliometric and content analysis review of research trends and future directions. *Applied Soft Computing*, 164, 111956. <https://doi.org/10.1016/j.asoc.2024.111956>
23. Hassan, S. I., Syed, S. A., Ali, S. W., Zahid, H., Tariq, S., Su, M. M. S., & Alam, M. M. (2024). Systematic literature review on the application of machine learning for the prediction of properties of different types of concrete. *PeerJ Computer Science*, 10, e1853. <https://doi.org/10.7717/peerj-cs.1853>