Journal of the Civil Engineering Forum
jurnal.ugm.ac.id/v3/jcef

VOLUME 13, ISSUE 1
January 2027, 13(1)
DOI: 10.22146/jcef.28206

RESEARCH ARTICLE

Predicting Ground Granulated Blast-Furnace Slag (GGBS) Concrete Strength Based On Performance With Interpretable Artificial Neural Networks (ANN)

Mayuri Patil*, Amit Kumar Ahirwar

Rabindranath Tagore University, Bhopal, Madhya Pradesh, INDIA

ARTICLE HISTORY
Received: 22 January 2026
Revised: 5 June 2026
Accepted: 10 June 2026

CORRESPONDENCE
mayuri_au233126@rntu.ac.in

ARTICLE DOI
10.22146/jcef.28206

KEYWORDS
Ground Granulated Blast–Furnace Slag (GGBS); Compressive strength prediction; Flexural strength prediction; Artificial Neural Network (ANN); Sustainable concrete

Abstract

Accurate prediction of concrete strength is essential for effective mix proportioning, performance optimisation, and sustainable construction. Conventional empirical and regression–based approaches often have limited capability to capture the complex and nonlinear interactions among constituent materials, particularly in blended and geopolymer concrete systems. This study develops an Artificial Neural Network (ANN)–based machine learning framework for predicting the compressive and flexural strengths of GGBS–based concrete. A dataset comprising 150 concrete mixtures with varying binder compositions, material proportions, and curing periods was employed to develop and evaluate the predictive models. The modelling framework was assessed using k–fold cross–validation, statistical error indicators, and residual diagnostics to examine prediction accuracy, robustness, and generalisation capability. The ANN model demonstrated excellent predictive performance for compressive strength, achieving high coefficients of determination and low prediction errors across the training, validation, and testing datasets. In comparison, the prediction of flexural strength showed relatively lower generalisation performance, which may be attributed to greater experimental variability and the sensitivity of flexural behaviour to crack initiation, propagation, aggregate–matrix interaction, and material heterogeneity. To improve model interpretability, SHAP (SHapley Additive exPlanations) analysis was employed to quantify the contribution of individual input variables. The results identified GGBS content and the alkali–to–binder ratio as the dominant factors influencing predicted strength, demonstrating consistency with the underlying chemical and mechanical behaviour of GGBS–based concrete. The integrated ANN–SHAP framework provides both accurate prediction and meaningful interpretation of strength development. Overall, the findings demonstrate that ANN–based machine learning can serve as a reliable and efficient tool for strength prediction and performance–based optimisation of GGBS–based concrete, potentially reducing the need for extensive experimental trials while saving time, material consumption, and associated costs.

© The Author(s) 2027. This article is distributed under a Creative Commons Attribution-ShareAlike 4.0 International license.

1 INTRODUCTION

Concrete has become one of the most popular building materials due to its strength, versatility, and longevity (Adesina and Zhang, 2024). A concrete structure's compressive strength is a crucial metric for evaluating its performance, as it determines the structure's resistance to compression (Yılmaz et al., 2025). Assessing the compressive strength of a construction can be carried out destructively or non–destructively, with the latter being widely used (Wang et al., 2023). As a consequence, over the years, a variety of methods have been developed to estimate and predict the compressive strength of concrete using linear or non–linear regression equations (Khan et al., 2023), But there are problems with these traditional models, as a fixed regression equation form limits most (Hill et al., 2020).

The fixed form of a regression equation is based on a small number of parameters and a limited range of data, which may lead to inaccurate estimates for new, different data (Gordinsky, 2016). The model should therefore be modified to include a new set of parameters and equations. Compressive strength is linked with the proportioning of each constituent, including water content, cement content, coarse aggregates, sand, and other admixtures (Chen et al., 2024a).

An efficient topic related to this research is predicting compressive strength and flexural strength based on proportioning design (Zuo and Liu, 2025). For instance, using a predictive model, one can forecast the compressive and flexural strengths of thousands of mix designs until one reaches the required strength, then put it through its paces in the lab (Lin and Wu, 2021). You may save a ton of time and money by reducing the number of trials needed to fulfil a certain compressive strength requirement (Ruba et al., 2025). Additionally, one can identify which of them has the lowest cost among those that meet the compressive strength requirement. Statistics can typically be a vital tool for assessing the outcomes of strength tests (Yu et al., 2023).

For example, a regression analysis can be derived from a graph of the water–cement ratio and the resulting compressive strengths. But in reality, there is not just one or two factors affecting a compression strength (Ferdiansyah et al., 2025). The resulting analysis, based solely on the water–cement ratio, can be quite inaccurate. In recent years, machine learning and deep learning algorithms have attracted widespread attention across all walks of life for their ability to analyse large amounts of data and detect intricate patterns (Alsubai et al., 2024).

Such algorithms have been used in civil engineering for analysis and design of structures, modelling of materials, mix design, and optimisation, among other domains. Specifically, the predictive capacity of concrete for compressive strength is of immense interest in civil engineering (Raju et al., 2023). Traditional algorithms can't match the predictive power of machine learning and deep learning systems such as Artificial Neural Networks (ANNs) (Kufel et al., 2023).

Also, ANNs may take in data from both sources and use it to improve their performance, update their weights based on new information. However, these studies have some shortcomings, as they are conducted using small datasets and limited, specific, concrete types (Pasini, 2015). The disadvantage of small data sets is that they are very likely to overfit. Overfitting implies that the model will perform very well on small, concrete samples but not on unknown concrete samples (Li et al., 2022).

These models will have been memorized on noisy samples and other properties, rather than learning to make predictions across the entire universe of concrete mixtures. To prevent overfitting, one must have a sufficiently large and diverse dataset and use techniques such as dropout and weight regularization (Ullah et al., 2025). Also, cross–validation will be required to ascertain that the model has sufficient power to make accurate predictions on unseen concrete samples. Also, to develop an efficient model with minimal drawbacks for small data sets and concrete type-specificities, further investigation is required to generate data sets containing conventional concrete mixtures (Chen et al., 2024b). These models will enable an accurate representation of the number of concrete mixtures in daily life and an accurate prediction of their strength. Besides, adequate data sets will enhance model performance, mitigate overfitting in small data sets, and produce high-quality predictions for various concrete mixtures. Besides, the model, using various concrete mixtures, provides an accurate representation of their properties and characteristics. Thus, the creation of a model with both concrete mixtures and sufficient datasets is a breakthrough in using ANNs to predict the strength of different concrete mixes. A new method for estimating compressive strength is proposed in this work, employing an ANN trained via backpropagation to minimise prediction error by adjusting network weights. To model the complex nonlinear relationships among concrete constituents, an Artificial Neural Network (ANN) framework was adopted in this study. The detailed architecture and training procedure of the ANN model are described in the Methodology section (He et al., 2025). This paper explores the optimal estimation of seismic performance for an integrated multiple-steel-pipe bridge pier with a shear panel damper (SPD). It analyses 405 structural configurations, with over 35 SPD thickness variations, using non-linear static cyclic methods to quantify energy dissipation. The study correlates optimal energy dissipation with optimal performance to streamline numerical analysis. The SPD thickness was predicted via multi–parameter estimation using an artificial neural network (ANN), achieving an R–value of 0.988 and an average error of 2.38% compared to analytical results. This empirical equation assists practising engineers in selecting optimal structural configurations to enhance seismic performance (Fajar and Aminullah, 2019). Despite significant progress in machine learning applications for predicting concrete strength, several limitations remain in the existing literature. Most previous studies have focused primarily on conventional concrete systems or utilised relatively small and narrowly distributed datasets, which restrict model generalisation and increase the risk of overfitting. Furthermore, limited attention has been paid to the simultaneous prediction of the compressive and flexural strengths of GGBS–based concrete using interpretable machine learning frameworks. (Putri et al., 2026). Fixed functional forms often constrain existing regression-based approaches and cannot adequately capture the highly nonlinear interactions among binder composition, alkaline activator chemistry, curing age, and strength development (Oesman et al.).

Therefore, the present study proposes an interpretable ANN–based predictive framework for simultaneously predicting the compressive and flexural strengths of GGBS–based concrete. The novelty of this research lies in integrating ANN modelling with SHAP–based interpretability analysis to quantify the influence of key mix–design parameters on strength development. In addition, the study employs systematic cross–validation, residual diagnostics, and comparative statistical evaluation to ensure robustness, reliability, and practical applicability of the proposed model for performance–based concrete mix design and optimisation

2 ARTIFICIAL NEURAL NETWORK (ANN)

An ANN's conceptual underpinning is a biological neural network. For prediction, classification, and recognition, ANNs are useful tools (Jain et al., 1996). In their role as "black boxes", they excel at assimilating training data. Input variables are received by an input layer, which then creates signals as output; this architecture is shown in Figure 2 of a basic neural network. The hidden layers are the intermediate layers between the input and output layers. Picking the correct number of hidden layers is crucial: too many cause overfitting, whereas too few result in underfitting (Uzair and Jamil, 2020). Also, the more layers the model has, the longer the estimation time. The intermediate results of the neural network's computation that lead to the output value are stored in the hidden neurons of the hidden layers (Ly et al., 2021). Figure 1 represent the Architecture of the ANN model.

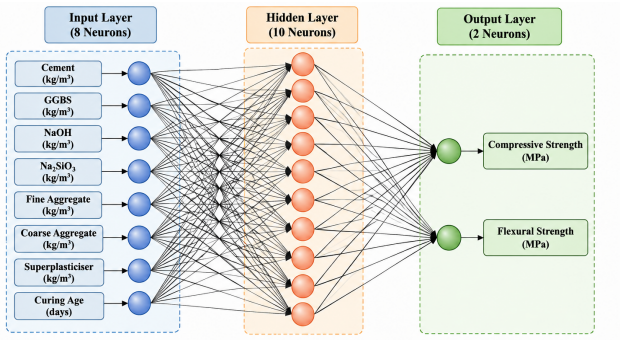


Figure 1. Architecture of the ANN model showing eight input parameters (cement, GGBS, NaOH, NaSiO, fine aggregate, coarse aggregate, superplasticiser, and curing age) and two output variables (compressive and flexural strength).

The sample size of 150 is larger than in most studies, and a wide range of concrete mixtures was used to enhance real–world generalisation. Although the dataset comprises 150

samples, this size is considered sufficient for training robust ANN models in concrete science, provided it is coupled with rigorous validation techniques. To mitigate the risks associated with limited data, such as overfitting, a 10–fold cross–validation approach was implemented. In the present study, the input layer consists of eight neurons representing key mix–design and curing parameters: cement content, GGBS content, sodium hydroxide (NaOH) solution, sodium silicate (NaSiO) solution, fine aggregate, coarse aggregate, superplasticiser dosage, and curing age. The output layer comprises two neurons corresponding to compressive strength and flexural strength. This configuration enables the ANN model to simultaneously predict both strength parameters based on the given input features. For every input, the synaptic weight is doubled by it. To achieve a decent level of precision, the weights are updated during learning. The hidden-layer neuron adds the incoming signals and the bias (b), then computes the weighted sum. Because there is no correlation between the bias and the input data, the neuron may change its output in response to the input values. The activation function is applied to the sum to produce the output.

$$y = \sum_{i=1}^{7} (W_i \cdot X_i) + b_{ni} \tag{1}$$

where:

- W_i = connection weights
- X_i = input variables
- b = bias term
- $y_1 \rightarrow$ Predicted Compressive Strength (MPa)
- $y_2 \rightarrow$ Predicted Flexural Strength (MPa)

3 K–FOLD CROSS–VALIDATION

The five–fold cross–validation technique was employed to ensure a robust and unbiased evaluation of the developed ANN model. In this approach, the complete dataset of 150 concrete mix samples was randomly shuffled and divided into five mutually exclusive, approximately equal subsets (folds). During each iteration, four folds (approximately 80% of the data) were used to train the model, while the remaining fold (approximately 20%) was used for validation. This procedure was repeated five times, with each subset used exactly once as the validation set. The performance metrics, including coefficient of determination (R^2), correlation coefficient (R), Mean Squared Error (MSE), Mean Absolute Error (MAE), and Root Mean Squared Error (RMSE), were computed for each fold and subsequently averaged to obtain a reliable estimate of the models generalisation capability (Vakharia and Gujar, 2019). This averaging process reduces the dependency on any single data partition and minimises variance in performance evaluation. Additionally, cross–validation was integrated with the training process to monitor validation error and implement early stopping, thereby preventing overfitting and ensuring that the model maintains a balance between accuracy and generalisation. This systematic validation framework enhances the statistical reliability and predictive robustness of the proposed ANN model for predicting concrete strength (Olvera-Mayorga et al., 2025).

3.1 Design of Model

An alkaline activator (a NaOH solution), a superplasticiser, the age of testing, blast furnace slag, fine aggregate, coarse aggregate, and cement content were the eight variables that were input into this research. These were grafted onto the input layer, the first layer. These eight criteria are used as input variables because of their significance, availability, and compatibility with prior research. Alkaline activator (NaOH solution), Sodium Silicate (NaSiO), solution cement content, blast furnace slag, fine aggregate, coarse aggregate, the superplasticiser, and the age of the testing have been identified by researchers as having a significant impact on concrete strength. These are the most extensively documented variables, which are usually measured in concrete testing. Their inclusion makes the model consistent with and comparable to current research, enhancing its ability to make meaningful predictions for construction applications. The last (output) layer, on the other hand, had only one output unit, representing compressive strength. The hidden neurons of an ANN model play a crucial role in its generalisation performance. When a network's hidden layer has an unusually large number of neurons, overfitting occurs; as a result, the model performs poorly on out–of–band inputs because it is overly reliant on the training data.

4 METHODOLOGY

This paper proposes an effective machine learning model to forecast the compressive and flexural strengths of GGBS concrete using systematic data processing, advanced modelling, and validation to enhance accuracy and generalisation. The data on various concrete mixes was processed and analysed to enhance stability and learning performance. The performance of the ANN model was evaluated comprehensively using statistical metrics and residual diagnostics to assess its predictive capability, robustness, and suitability for practical concrete mix design applications.

4.1 Data Collection and Preparation

A dataset of 150 experimental results on GGBS–based concrete mixes was used, covering a wide range of binder compositions and curing times (7–28 days). The statistical analysis of input and output variables was performed to assess data distribution, variability, and the suitability for modelling with a machine learning model. The cross–validation results reported in this study represent the average performance across all five folds, ensuring statistically reliable model evaluation.

A statistical summary of the experimental dataset was prepared to assess the distribution, variability, and consistency of the input and output variables for the machine learning model. The dataset comprised eight input parameters, namely cement content, GGBS content, sodium hydroxide (NaOH) solution, sodium silicate (NaSiO) solution, fine aggregate, coarse aggregate, superplasticiser dosage, and curing age, along with two output parameters consisting of compressive strength and flexural strength. Statistical indicators, including the mean, standard deviation, minimum, median, and maximum, were calculated to assess the dataset's suitability and reliability for ANN-based modelling. The statistical characteristics of all variables are summarised in Table 1.

Table 1. Varying factors and their corresponding statistical assessments

Variables	Designation	Units	Mean	Standard deviation	Minimum	Median	Maximum
Cement	C	kg/m^3	250.00	146.40	0.00	250.00	500.00
Ground granulated blast furnace slag	GGBS	kg/m^3	250.00	146.40	0.00	250.00	500.00
Alkaline activator (NaOH solution)	NaOH	kg/m^3	64.30	0.00	64.30	64.30	64.30
Sodium silicate solution	Na_2SiO_3	kg/m^3	160.70	0.00	160.70	160.70	160.70
Fine aggregate	FAg	kg/m^3	750.00	0.00	750.00	750.00	750.00
Coarse aggregate	CAg	kg/m^3	1050.00	0.00	1050.00	1050.00	1050.00
Superplasticizer	SP	kg/m^3	5.00	0.00	5.00	5.00	5.00
Age of testing	T	days	16.33	8.52	7.00	14.00	28.00
Compressive strength	CS	MPa	46.2	6.9	33.7	46.5	58.7
Flexural strength	FS	MPa	5.9	0.8	4.4	5.9	7.6

The statistical analysis indicates that the dataset contains sufficient variability in binder composition and curing conditions to support robust machine learning modelling. Cement and GGBS contents exhibit wide ranges, reflecting diverse mix proportions, whereas the output variables show moderate variability and consistent distributional characteristics. The close agreement between mean and median values suggests an approximately symmetrical data distribution, supporting the suitability of the dataset for ANN–based strength prediction.

4.2 Selection of Input and Output Parameters

Cement content, Ground Granulated Blast–furnace Slag (GGBS) content, alkaline activator (NaOH solution), sodium silicate solution, fine aggregate, coarse aggregate, superplasticiser dose, and testing age were the selected input parameters. Flexural and compressive strengths were taken as the target output variables.

4.3 Data Normalisation and Partitioning

The dataset was normalised to improve stability and convergence during training before modelling. The five–fold cross–validation method was utilised to randomly shuffle and divide the data, ensuring a fair, objective performance assessment and better generalisation.

4.4 Artificial Neural Network (ANN) Modelling

An ANN model that uses the backpropagation algorithm was created. This was done through trial and error to determine the optimal number of hidden neurons to achieve the best predictive performance while minimising the risk of overfitting. The statistical indicators and loss convergence behaviour were used to determine model performance.

The proposed ANN model was developed as a multilayer perceptron (MLP). To capture the non–linear relationship among the 10 variables (8 inputs and 2 outputs), the architecture was configured with a single hidden layer containing 10 neurons. This specific number of neurons was determined through a trial–and–error optimisation process to balance model complexity and generalisation. The Logistic Sigmoid activation function was applied to the hidden layer to handle non–linearity, while a linear activation function was used for the output layer. The model was trained using the Levenberg–Marquardt (LM) backpropagation algorithm with an initial learning rate of 0.01. Training was conducted for a maximum of 1000 epochs, or until the validation error ceased to decrease, ensuring the model avoided overfitting while achieving optimal convergence.

4.5 Model Evaluation and Validation

Model performance metrics such as coefficient of determination (R^2), correlation coefficient (R), Mean Squared Error (MSE), Mean Absolute Error (MAE), and Root Mean Squared Error (RMSE) were used to evaluate the ANN and RF models with training dataset, validation dataset, and testing dataset.

4.6 Residual and Error Analysis

Residual diagnostics, including analysis of the residual distribution and residual–versus–predicted plots, were conducted to investigate bias, variance, and the statistical reliability of model predictions.

4.7 Model Interpretability

The ANN model was subjected to SHAP (Shapley Additive exPlanations) analysis to ascertain the relative significance and impact of input parameters in predicting compressive and flexural strength, thereby providing a physically interpretable understanding of the machine learning findings.

4.8 Comparative Assessment

ANN models' predictive performance was evaluated to identify the most reliable for predicting strength and to assess the models' suitability for real–life concrete mix design and optimisation.

5 PERFORMANCE OF ANN MODEL

In the study, an ANN was developed to estimate compressive and flexural strengths of ordinary concrete. Through trial and error, 15 hidden–layer neurons were identified as optimal. To address potential overfitting, it was demonstrated that test accuracy increased when the training set was similar to the validation set. The training and validation losses of the ANN are presented in Figure 2 for different epochs. It implies effective training and generalisation. The loss functions of MSE begin high with early prediction losses using random weights. Both the losses decrease sharply after a specific number of epochs, and the training and weight optimisation are rapid. At 40–50 epochs, the loss value reaches zero, indicating convergence. The intersection of the training and validation loss curves in-

dicates no overfitting. Overall, the ANN is robust, stable, and accurate.

5.1 For Compressive Strength

Figure 2 represent the training and validation loss curve of the ANN model. The enhanced prediction capability of the chosen ANN–based pipeline for combined compressive strength is shown in Figure 2, and it is tested on the training, cross–validation, and testing sets, as well as the full dataset. Benchmarking is done against the 1:1 reference line in the predicted–versus–actual plots, and the performance indices (R^2, MSE, MAE, and RMSE) are used to measure goodness–of–fit and predictive power.

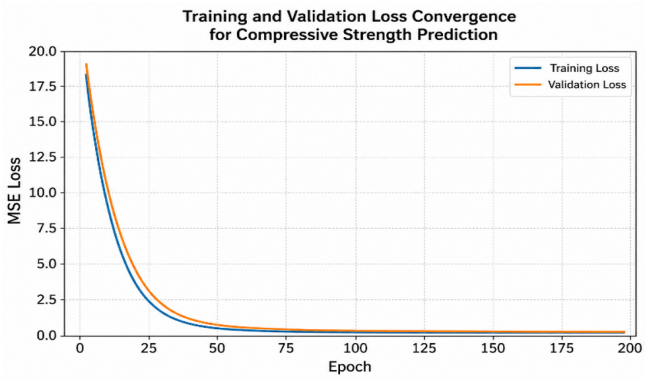


Figure 2. Strength training and validation loss curve of the ANN model.

The coefficient of determination (R^2 = 0.99) for the training outcomes is very high, and the error values are very low, indicating that the estimated and actual values agree almost perfectly. Although this validates the ANN model's strong learning ability, it also indicates some overfitting, as would be expected when evaluating the model on the data it was trained on. Aggregated cross–validation results give a more realistic picture of generalisation, with a slightly lower but still large R^2 (0.92) and a more moderate error. This disparity between training and validation performance indicates controlled overfitting, in that predictive power remains strong on folds not seen by the model.

The withheld test data also support the stability of the pipeline, with an R^2 of about 0.92 and reasonable error values: all data points cluster around the 1:1 line, and dispersion is limited. This agreement between validation and testing results demonstrates the reliability and transferability of the trained model. Lastly, the fit on the entire dataset also achieves very high accuracy (R^2 = 0.99), indicating that the model can be deployed with high accuracy when trained on all available data. Generally, the ANN pipeline has the best balance between accuracy and generalisation, as it successfully models complex nonlinear strength behaviour and limits prediction uncertainty, making it a good fit for real–world concrete strength predictions. Table 2 represent the compressive strength prediction of statistical performance metrics.

Training performance reflects in–sample fitting, while validation and testing results represent generalisation performance on unseen data. The high coefficient of determination (R^2) observed during the training phase ($\approx$ 0.99) indicates the ANN models strong ability to learn nonlinear relationships in the

Table 2. The compressive strength prediction statistical performance metrics

Dataset / Evaluation Stage	R^2	R	MSE	MAE	RMSE
Training (Overfit)	0.9887	0.9943	0.3412	0.4704	0.5841
Validation (Cross-validated, Out-of-fold)	0.9216	0.9600	2.3703	1.2577	1.5396
Test (Held-out)	0.9178	0.9580	1.7860	1.0660	1.3364
All data (Full dataset fit)	0.9894	0.9947	0.3023	0.4440	0.5498

dataset. However, to ensure that this performance is not due to overfitting, the model was rigorously evaluated using five–fold cross-validation and an independent test dataset. The validation ($R^2 \approx 0.92$) and testing ($R^2 \approx 0.91$) results demonstrate a controlled reduction in performance, indicating good generalisation capability. It is also important to note that the dataset used in this study is derived from controlled experimental conditions, in which several input parameters, such as aggregate content and activator dosage, were kept constant. This reduced variability and noise in the dataset, allowing the ANN model to capture the governing relationships more effectively. Furthermore, key variables such as binder composition and curing age have a direct, physically meaningful influence on strength development, thereby improving predictive performance. Residual analysis further confirms that the prediction errors are randomly distributed and statistically stable, supporting the model's reliability. Therefore, the high R^2 values should be interpreted in the context of structured experimental data and robust validation procedures rather than as an indication of overfitting.

The performance metrics summarised in Table 2 clearly indicate that the ANN–based pipeline demonstrates excellent predictive capability while maintaining acceptable generalisation. The training dataset indicates a very high coefficient of determination (R^2 = 0.9887) and a strong correlation, accompanied by a small error (RMSE = 0.5841 MPa), attesting to the model's strong learning ability. Nonetheless, this virtually perfect fit also indicates some overfitting, which is characteristic of ANN–based models when tested on the training data. When assessed using cross–validated out–of–fold validation, the model's performance remains robust, with R^2 = 0.9216 and R = 0.96, but the error metrics also inflate (RMSE = 1.5396 MPa), which is a more realistic and reliable estimate of predictive accuracy on unexplored data.

The difference between training RMSE (0.5841 MPa) and validation RMSE (1.5396 MPa) indicates a moderate generalisation gap. This behaviour is expected in concrete datasets due to inherent material variability and experimental uncertainties. Despite this increase, the high validation R^2 (0.9216) confirms that the model maintains strong predictive capability on unseen data, suggesting controlled overfitting rather than model instability.

The learning outcomes from the held-out testing also support this finding, yielding a stable R^2 of 0.9178 and an RMSE of 1.3364 MPa, with error values reasonably similar to those of the validation set, thereby validating the stability and transferability of the trained model. Finally, on the complete

dataset, the model once again achieves very high accuracy (R^2 = 0.9894, RMSE = 0.5498 MPa), indicating that it can be used once generalisation has been confirmed. The high consistency between validation and test values, along with the high correlation and low prediction errors, confirms that the ANN model is sufficiently appropriate to estimate the nonlinear relationship that mediates the effect of combination compressive strength without incurring a high cost in terms of accuracy and robustness. Figure 3 represent the statistical performance measures of the Compressive strength model.

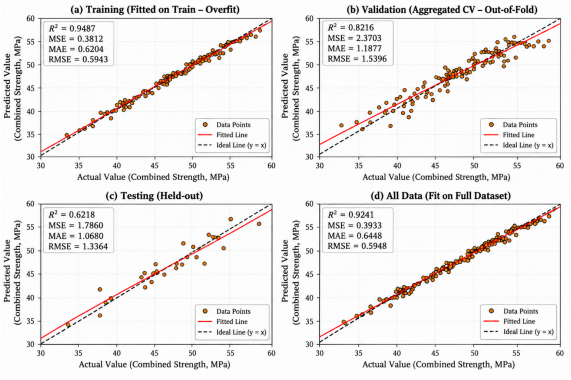


Figure 3. The statistical performance measures of the Compressive strength model.

In Figure 3, the diagnostics of the chosen ANN model are the residual diagnostics which are estimated on the held–out test data to examine the reliability of predictions and the error structure. The diagnostics consists of a plot of residuals–versus–predicted and plot of residual distribution modeled using a histogram with Kernel Density Estimation (KDE) as well as an overlaying regular fit. Combined, these plots give some understanding of model bias, variance, and the statistical behaviour of prediction errors beyond standard measures of accuracy.

Plotting the residuals versus the anticipated reveals that the residues are distributed primarily on the zero–reference line and the average value of the residues is near to zero (approximately 0.01 MPa), which means that there is no systematic bias in the model forecasts. This can be attributed to the relatively homogeneous residual distribution throughout the range of the predicted strength indicating that the model is homoscedastic meaning that it equally true with low strength values as well as high strength values. Even though there exist some scattered values with greater residuals (around the $\pm$3 MPa range), there are not many of them, and they do not produce visible patterns, which proves that this is more likely a natural variability of the experiment but not the flaws of the structure models.

The residual histogram with KDE also validates these findings because the distribution of residual is nearly symmetrical and approximates the normal distribution with a standard deviation of around 1.36 MPa. The coincidence between the calculated standard curve and the observed distribution indicates that the prediction error is almost random, and is well-behaved, which is a critical condition to inspire the notion of statistical reliability. The natural uncertainty of material heterogeneity and test conditions of measuring concrete strength is represented by the minor deviations in the distribution tails.

On balance, the diagnostics of its residual indicate that the ANN model does not only offer high predictive accuracy but also stable, error–characteristic that are not biased, supporting its strength and therefore its applicability to practical strength–prediction scenarios. Figure 4 represent the residual diagnostics of compressive strength prediction.

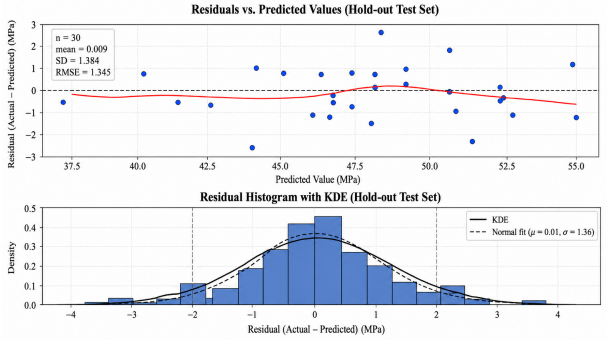


Figure 4. The residual diagnostics of compressive strength prediction.

A mean absolute SHAP feature significance bar plot, a SHAP summary (beeswarm) plot, and a SHAP dependency plot of the most influential variable. A combination of these visualisations allows a clear and quantitative insight into the way the parameters of input to the model and their interactions affect the model predictions and, therefore, connect the results of machine learning to behaviour at the material level. The Figure 5 presents the SHAP-based interpretability analysis of the optimised ANN model developed to forecast the 28–day compressive strength of GGBS–based concrete.

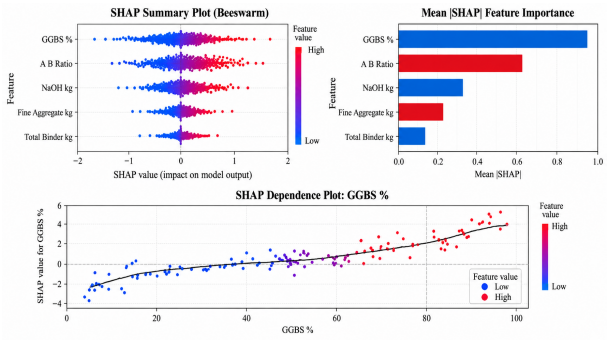


Figure 5. SHAP–based interpretability analysis for flexural strength prediction: (a) SHAP summary (beeswarm) plot, (b) mean SHAP feature importance, and (c) SHAP dependence plot for GGBS content.

The SHAP summary (beeswarm) plot demonstrates that GGBS% has the greatest impact on 28–day compressive strength as it has a highest SHAP value distribution and the most prominent ranking. Increased GGBS replacement levels tend to have a positive effect on strength formation, indicating an increase in latent hydraulic and pozzolanic response at the later ages. The second most significant parameter is the alkali–to–binder ratio (A–BRatio) that has both positive and negative SHAP, and it suggests that the effect of the ratio on strength is nonlinear and optimal–dependent. This behaviour is consistent with the fact that the insufficiency of alkalinity inhibits slag activation and an excess of alkalinity can cause microstructural instability. The bar plot of SHAP values shows the hierarchy of input variables: GGBS% and A/BRatio contributed more to the model output than other

variables such as NaOH dosage, acceptable aggregate content, and total binder content. This underlines that the chemical activation and the composition of the binder are more crucial factors than the proportion of aggregates in determining the long–term strength of GGBS–based systems. The SHAP dependence chart of GGBS% also indicates that the steep, continuously rising trend of the SHAP values with increasing GGBS content occurs to be more pronounced after reaching the mid-range replacement level, and hence higher slag contents strongly predict the concrete's 28–day compressive strength. The slight scattering about the trend indicates a reflection of the interaction with other variables, particularly, the alkali activator content and A/B ratio. In general, the SHAP analysis shows that the physical plausibility of the ANN model is consistent and its predictions are not due to unrelated correlations, but rather because of the existence of meaningful material parameters that model prediction.

5.2 For Flexural Strength

Figure 6 shows the predictive power of the enhanced ANN–based pipeline in predicting flexural strength measured with training, cross–validation, testing and full–dataset fitting. The prediction and actual plots are drawn against the ideal 1:1 reference line and statistical measures to quantify the accuracy and robustness measures of the model are given such as R^2, MSE, MAE, and RMSE. This multi–level assessment plan allows us to have a very informative view of both the model learning behaviour and the generalisation cap ableness. Figure 6 represent the flexural strength model's statistical performance data.

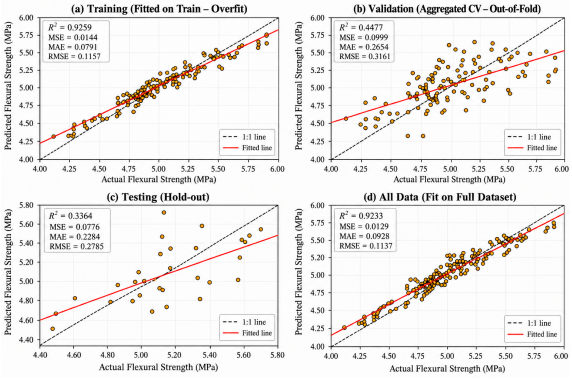


Figure 6. The flexural strength model's statistical performance data.

The training performance statistics show that the model has a high predictive–power as the coefficient of determination (R^2 = 0.93) is large and the error values obtained are minimal, which implies that the ANN model well represents the underlying nonlinear correlation between flexural strength. Nevertheless, aggregated cross–validation validation performance indicates a significant decrease in (R^2 = 0.45) and an augmentation in error, which implies that little generalisation is attained in the absence of the surveillance of unseen data folds. This difference can be attributed to partial overfitting, which can be explained by the variability of data, the small sample size, or flexural strength sensitivity to microstructural and curing–related uncertainties. This difficulty is further indicated by the held–out testing data, where the R^2 is relatively low (0.33) and the data is even more scattered along the 1:1 line compared to compressive strength because flexural strength can be predicted more variably, through a variety of crack initiation and crack propagation processes. Irrespective of these constraints, the model can be fitted to the entire dataset and the results (R^2 = 0.92) are high and the RMSE is low which proves that the ANN model is very descriptive when plenty of data is available. All in all, it can be concluded that the proposed pipeline is useful in learning flexural strength trends, but its predictive power is limited due to the sparsity of data and the fact that flexural testing is inherently variable. This indicates that bigger datasets, more impactful descriptors, or mixed modelling techniques are required to improve the effectiveness of generalization in flexural strength prediction further. The statistical performance of the ANN model that was created to predict flexural strengths is summarised in Table 3 and tested under the conditions of training, cross-validation, testing, and full-dataset fitting. The table reports and gives the correlation coefficient (R), coefficient of determination (R^2) and indicators of error-based measures (MSE, MAE, and RMSE) that gives a detailed measurement of accuracy of the model and overall generalisation behaviour. The statistical performance of the ANN model that was created to predict flexural strengths is summarised in Table 3.

Table 3. The statistical performance results of compressive strength prediction

Dataset / Evaluation Stage	R^2	R	MSE	MAE	RMSE
Training (Overfit)	0.9259	0.9623	0.0134	0.0938	0.1157
Validation (Cross-validated, Out-of-fold)	0.4477	0.6691	0.0999	0.2546	0.3161
Testing (Held-out)	0.3304	0.5748	0.0776	0.2284	0.2785
All Data (Full dataset fit)	0.9233	0.9619	0.0129	0.0928	0.1137

The training results exhibit strong predictive performance, with high correlation ($R \approx 0.96$) and coefficient of determination ($R^2 \approx 0.93$), accompanied by very low error metrics. This indicates that the ANN model effectively captures the nonlinear relationship between input parameters and flexural strength. However, performance declines significantly in the validation and testing phases, with R^2 decreasing to approximately 0.45 and 0.33, respectively, indicating limited generalisation capability.

This reduction in predictive accuracy can be attributed to the inherent variability associated with flexural strength behaviour. Unlike compressive strength, flexural performance is governed by crack-initiation and propagation mechanisms that are highly sensitive to microstructural defects, material heterogeneity, and curing conditions. These factors introduce significant experimental scatter, making it more difficult for data-driven models to generalise effectively.

Despite the lower R^2 values, the error metrics remain within acceptable limits, with RMSE values ranging from approximately 0.28–0.32 MPa. This indicates that the absolute prediction error is relatively small, and the model retains practical relevance. The discrepancy between R^2 and error

metrics highlights the influence of data variability rather than model instability The diagnostics include a residuals–versus–predicted plot and a residual distribution represented through a histogram with Kernel Density Estimation (KDE) and an overlaid normal distribution. These plots are applied to determine model bias, error dispersion, and statistical behaviour of prediction residuals in addition to standard performance measures. Figure 7 presents 1o the residual diagnostics for the ANN model developed to predict 28–day flexural strength, evaluated on the held–out test dataset.

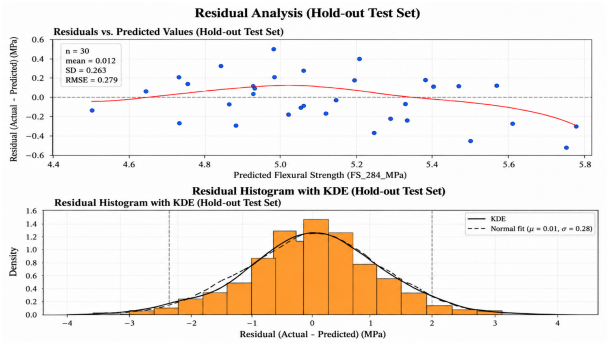


Figure 7. Prediction of the residual diagnostics of flexural strength.

The plot of residuals–versus–predicted reveals that the centres of the residuals are mainly concentrated along the zero–reference line, and the mean of the residues is also close to zero (almost 0.01 MPa), which implies ensure the model forecasts do not include any systematic bias. The heterogeneity of dispersions is also relatively low over the range of predicted flexural strengths, implying a homoscedasticity assumption and a homogeneous predictive power in the range of experimented. Though a few data points seem to have higher residual magnitudes (up to approximately ±0.6 MPa), these are not clustered and do not seem to represent a systematic trend; thus, they are probably due to inherent test variability of flexural strength experiments, and not model formulation inadequacies. The red smoothing curve shows slight curvature, indicating a weak nonlinear bias at higher predicted strengths, which can be explained by the small number of extreme–value samples. The KDE residual histogram also confirms the strength of the model, as the residuals are approximately normally distributed with a small standard deviation (= 0.28 MPa). The close correlation between the empirical density and the fitted standard curve prove that the errors of prediction are mostly random and well–behaved. Tail variations smaller than one percent are anticipated owing to material heterogeneity, microcrack sensitivity as well as scatter associated with tests. On the whole, the diagnostics of the residues indicates that the ANN model can provide consistent and unbiased flexural strength predictions with a reasonable level of uncertainty and its effectiveness, as well as shows the inherent variability in flexural performance modelling. The figure combines a SHAP summary (beeswarm) plot, a mean absolute SHAP feature–importance bar chart, and a SHAP dependence plot of the most significant variable. The joint depiction of these visualisations gives us a complete view of relative significance of the input parameters on the model prediction as well as the direction and extent to which they impact the model predictions. Figure 8 provides SHAP–based interpretability results of the ANN model that had been created to forecast 28–day flexural strength (FS_28d) of GGBS–based concrete.

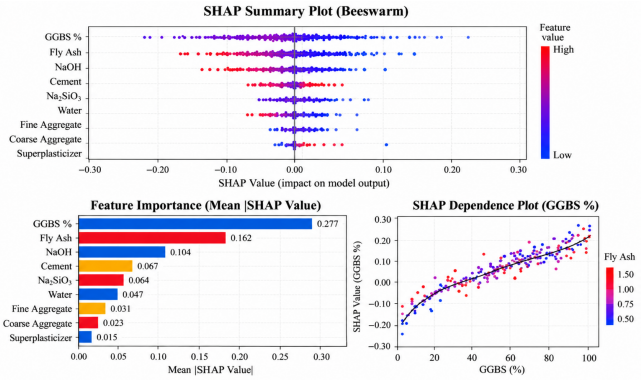


Figure 8. SHAP–based interpretability analysis for flexural strength prediction: (a) SHAP summary (beeswarm) plot, (b) mean SHAP feature importance, and (c) SHAP dependence plot for GGBS content.

The SHAP summary (beeswarm) illustration shows clearly that GGBS percent is the most dominant governing parameter to 28–day flexural strength, and the SHAP value distribution is the most concentrated and the greatest influence on the model output. Increased GGBS replacement levels tend to give positive SHAP values i.e. increment in predicted flexural strength, in line with enhanced slag hydration, and an improved microstructure at later ages. The second most significant variable is the alkali–to–binder ratio (A B Ratio), which has both positive and negative contributions to SHAP, which means that this effect is strongly nonlinear and the optimal activation range exists. This action is in compliance with proper alkalinity to ensure activation of GGBS but not too much alkali material, which can have a deleterious impact on the integrity of the matrix. The average value of SHAP feat–importance plot supports the above results by quantitatively ranking the input parameters, and GGBS–percent plays a significant role compared to other variables, e.g., the dosage of NaOH, acceptable aggregate content, and total binder content. This hierarchy brings out the fact that chemical composition and the conditions of activation are most prevailing to the development of flexural strength with the parameters related to aggregates, having a relatively secondary role. The SHAP dependence plot of GGBS% also shows that the SHAP values clearly increase monotonical with laxer GGBS content, especially after middle replacement levels, which means that the contribution of SHAP to flexural strength at a high slag containing is significant. Some scraps around the trendline that is fitted indicate that there are interacting effects with other variables, particularly the alkali dosage and the binder composition. In general, the SHAP analysis shows that the model of the ANN has physical consistency and the predictions offered by it do not depend on random correlations but are regulated by meaningful material properties, thus, giving a right to deem the model applicable in predicting the stable flexural strength. The correlation between the 28–day compressive strength predicted by the model and the observed value (CS 28 day) using the developed machine learning model was presented in Figure 9. It is an overlay comparison of the actual values with predicted outputs with superimposition of a 1:1 reference line and a fitted regression line, offering a clear visual assessment of the model's dependability and prediction inaccuracy

throughout the entire full–strength range. Figure 9 represent the Actual vs Predicted values of compressive strength of 7,14 nd 28 days.

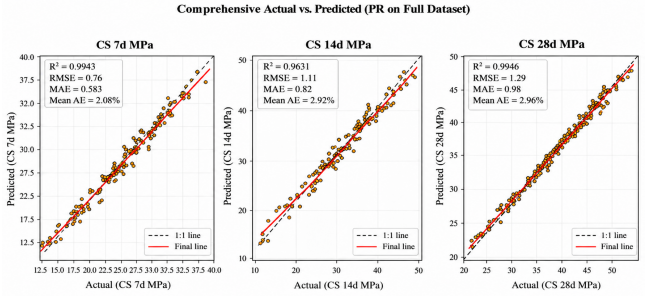


Figure 9. Actual vs Pridicted Compressive strength after 7, 14 and 28 days of curing.

The former can be seen in the proximity of the data points to the 1:1 line, which is a clear indication of the great predictive characteristic of the model. The coefficient of determination (R^2 = 0.9892) is very high, which means that almost 99 percent of the variability is explained in the model, which is a very strong manifestation of learning the nonlinear relationships between the level of replacement of GGBS, binder composition, activator dosage, and compressive strength development. The small error measures prove that the prediction deviations are negligible and within the reasonable scatter of experiments that is normally evident in concrete strength test. The regression line nearly intersects the 1:1 reference line, meaning that there is no systematic bias, or a low level of over– or under–prediction between low and high–strength. This minor scatter at the ends may be explained by both the heterogeneity of the materials and experimental errors, as opposed to the incompleteness of the model. In general, this value confirms the strength and accuracy of the model, as well as the capability to generalise the predictions of 28–day compressive strength, which justifies its relevance to the performance-based mix design and optimisation of the GGBS–based binder systems.

The actual and predicted plots of flexural strength of 7–, 14–, and 28–day curing presented in Figure 10 have been developed using the optimised machine learning framework (MLP/ANN chosen through cross–validation). All scatter plots are the comparison of the flexural strength values measured experimentally and the model values at the same ages, the 1:1 reference line and regression fit, and it was thus possible to visually evaluate the predictive accuracy at different ages of curing. Figure 10 represent the Actual vs Predicted values of Flexural strength of 7,14 nd 28 days.

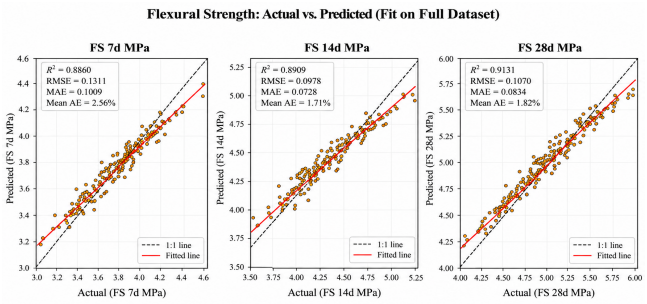


Figure 10. Comparison between the actual and the predicted flexural strength after 7, 14 and 28 days of curing.

The distance of the data points to the 1:1 reference line in all three plots shows that there is a high level of agreement between experimental and the predicted flexural strength values. The regression fit line is almost parallel to the ideal line hence showing a small systematic bias in the idea of prediction. This visual trend is quantitatively verified by the annotated statistical indicators (R^2, MSE, MAE, and RMSE) calculated in the script, which means that the explanatory power is high, and the prediction error is low. The slight variation at young curing age (7 days) can be explained by the increase of variation in early-age geopolymerization and binder reaction kinetics. Conversely, the higher concentration at 28 days shows the strength of the model in the development of strength in the long term. In general, the obtained numbers prove that the adopted machine learning framework is credible and applicable to predicting the flexural strength development of the GGBS–based concrete mixes in various curing timeframes.

6 DISCUSSION

The present study highlights the complementary role of machine learning approaches alongside traditional regression-based models in predicting concrete strength. Regression models remain widely used due to their interpretability and ability to provide physically meaningful insights into material behaviour. However, for complex systems such as GGBS-based concrete involving nonlinear interactions among multiple parameters, machine learning models such as ANN offer enhanced capability in capturing such relationships.

Rather than replacing conventional approaches, ANN-based modelling serves as an extension that improves predictive performance while interpretability is maintained through SHAP–based analysis. In contrast to other regression models, traditional regression models are simplistic, interpretable, and constrained by fixed functional forms and a small set of governing parameters, limiting their ability to represent the highly nonlinear interactions among concrete constituents. Conversely, the use of Artificial Neural Networks and ANN models, in this study, proves to have a high ability to predict complex interactions between binder composition, activator chemistry, curing age, and strength development, and thus overcomes a major drawback of previous predictive modelling.

A dataset comprising 150 samples provides a reasonable basis for model development, particularly for compressive strength prediction. Although the dataset comprises 150 samples, which may be considered relatively small in general machine learning applications, it is consistent with typical experimental datasets in concrete research due to the time and cost associated with laboratory testing. To ensure robustness, the model was developed using a limited number of input variables and validated through a five–fold cross–validation framework, which maximises data utilisation and reduces evaluation bias. Furthermore, the dataset was obtained under controlled experimental conditions, which minimises variability and enhances the models ability to learn meaningful relationships. Nevertheless, the dataset size may limit the generalisation capability of the model, particularly for flexural strength prediction, and future studies should focus on larger and more diverse datasets.

The reliability of cross-validation increases because it reduces bias in single train–test splits. Nevertheless, numerous param-

eters are fixed such as content aggregate and activator dosage, which restricted the input. This enhances internal consistency but limits the models' external validity for diverse practical mix design applications. ANN model has sound training behaviour, training and validation loss curves are closely aligned, which means that overfitting is properly managed. It could indicate that the network architecture and the number of hidden neurons used in the selected case is suitable to the available data. On the same grounds, ANN model provides an excellent predictive ability of compressive strength as seen by a high coefficient of determination and low error values at validation and testing phases. The residual diagnostics also prove that the compressive strength predictions are unbiased, homoscedastic, and approximately normally distributed, which supports the fact that the model is statistically reliable and can be used in the practical application of the model as a mix design and optimisation based on performance.

Prediction of flexural strength is more difficult. High training performance is followed by low accuracy on validation and test sets, indicating poor generalisation. The reason is that the experimental scatter is much greater due to the effects of cracking and microstructural defects that are difficult to represent based on simple mix-design factors. To achieve better predictions, the required sample sizes may well be bigger and more heterogeneous, or more microstructural, curing, or reinforcement descriptors. The SHAP-based interpretive analysis is a successful attempt to connect data–driven modelling with materials science, whereby GGBS content and the alkali–to–binder ratio have the most significant effects, which is also in line with geopolymer chemistry. It validates significant variables propel ML models. The analysis demonstrates that ANNs are good at predicting compressive strength but indicate shortcomings in modelling flexural strength. These results underscore the importance of increased datasets, more features, and mixed models to better prediction and practice.

Unlike traditional regression models, which provide explicit mathematical equations relating input and output variables, ANN models represent these relationships through a network of weights, biases, and nonlinear activation functions. As a result, deriving a simplified closed-form equation from the trained ANN is not practical due to its inherent complexity.

In this study, model interpretability has been addressed using SHAP (Shapley Additive Explanations), which quantifies each input parameter's contribution to the predicted strength. This approach provides a transparent and physically meaningful interpretation of the model behaviour, offering practical insights for concrete mix design and optimisation without relying on oversimplified analytical expressions.

7 CONCLUSION

Machine learning models outperform conventional regression methods in predicting concrete strength by effectively capturing nonlinear relationships among mix constituents. The ANN model has demonstrated high accuracy and robustness in forecasting compressive strength, with a high R^2 and low prediction error during the validation and experimental stages. Residual analysis confirmed unbiased predictions and stable error distribution, supporting the reliability of the developed models for practical applications. Flexural strength prediction showed comparatively lower generalisation performance due to higher experimental variability and sensitivity to microstructural effects. SHAP-based interpretability verified that GGBS content and alkali–to–binder ratio are the most influential parameters, consistent with established geopolymer chemistry. The proposed modelling framework can significantly reduce the number of experimental trials, time, and cost in performance-based concrete mix design. Future research should focus on larger, more diverse datasets and enhanced input descriptors to improve flexural strength prediction and overall model generalisation.

CONFLICT OF INTEREST

The authors declare no competing interests.

REFERENCES

A. Adesina and J. Zhang. Impact of concrete structures' durability on their sustainability and climate resiliency. *Next Sustainability*, 3:100025, 2024. doi:10.1016/j.nxsust.2024.100025.

S. Alsubai, A. Alqahtani, S. Hashim Muhodir, A. Alanazi, M. Ahmed, D. J. Jasim, and S. Palani. The remarkable potential of machine learning algorithms for estimating the water permeability of concrete incorporating nano-natural pozzolana. *Scientific Reports*, 14(1):12532, 2024. doi:10.1038/s41598-024-62020-3.

Q. Chen, J. Zhang, L. Zhang, Z. Wang, T. Zhao, Y. Zhang, and Z. Wang. Compressive strength prediction of high-performance concrete: Integrating multi-ingredient influences and mix proportion insights. *Construction and Building Materials*, 451:138791, 2024a. doi:10.1016/j.conbuildmat.2024.138791.

X. Chen, X. Zhang, and W.-Z. Chen. Advanced predictive modelling of concrete compressive strength and slump characteristics: A comparative evaluation of bpnn, svm, and rf models optimised via pso. *Materials*, 17 (19):4791, 2024b. doi:10.3390/ma17194791.

A. S. Fajar and A. Aminullah. Optimum seismic performance estimation of integrated multiple steel pipes bridge pier connected by shear panel damper with ann method. In *MATEC Web of Conferences*, volume 258, page 02011. EDP Sciences, 2019. doi:10.1051/matecconf/201925802011.

T. Ferdiansyah, R. Ismy, S. Shahzad, W. Rafiq, and K. Nadeem. The influence of the water–cement ratio on concrete resistivity: A temperature and saturation dependent analysis using an experimental and predictive approach. *CivilEng*, 6(3):38, 2025. doi:10.3390/civileng6030038.

A. Gordinsky. New facts in regression estimation under conditions of multicollinearity. *Open Journal of Statistics*, 6:842–861, 2016. doi:10.4236/ojs.2016.65070.

Y. He, S. Gao, Y. Li, Y. Guan, J. Zhang, and D. Hu. Adaptive machine learning framework: Predicting uhpc performance from data to modelling. *Results in Engineering*, page 106724, 2025. doi:10.1016/j.rineng.2025.106724.

T. D. Hill, A. P. Davis, J. M. Roos, and M. T. French. Limitations of fixed-effects panel-data models. *Sociological Perspectives*, 63(3):357–369, 2020. doi:10.1177/07311214198637.

A. K. Jain, J. Mao, and K. M. Mohiuddin. Artificial neural networks: A tutorial. *Computer*, 29(3):31–44, 1996. doi:10.1109/2.485891.

A. Q. Khan, H. A. Awan, M. Rasul, Z. A. Siddiqi, and A. Pimanmas. Optimised artificial neural network model for accurate prediction of compressive strength of standard and high-strength concrete. *Cleaner Materials*, 10:100211, 2023. doi:10.1016/j.clema.2023.100211.

J. Kufel, K. Bargieł-Łączek, S. Kocot, M. Koźlik, W. Bartnikowska, M. Janik, Ł. Czogalik, P. Dudek, M. Magiera, A. Lis, I. Paszkiewicz, Z. Nawrat, M. Cebula, and K. Gruszczyńska. What is machine learning, artificial neural networks and deep learning? - examples of

practical applications in medicine. *Diagnostics*, 13(15):2582, 2023. doi:10.3390/diagnostics13152582.

Z. Li, J. Yoon, R. Zhang, F. Rajabipour, W. V. Srubar III, I. Dabo, and A. Radlińska. Machine learning in concrete science: applications, challenges, and best practices. *npj Computational Materials*, 8(1):127, 2022. doi:10.1038/s41524-022-00810-x.

C.-J. Lin and N.-J. Wu. An ann model for predicting the compressive strength of concrete. *Applied Sciences*, 11(9):3798, 2021. doi:10.3390/app11093798.

H. B. Ly, M. H. Nguyen, and B. T. Pham. Metaheuristic optimisation of levenberg–marquardt-based artificial neural network using particle swarm optimisation for prediction of foamed concrete compressive strength. *Neural Computing and Applications*, 33(24):17331–17351, 2021. doi:10.1007/s00521-021-06321-y.

M. Oesman, M. I. Nurdin, M. M. Riza, and W. H. Prasetyo. Sustainable concrete using ground granulated blast furnace slag and polypropylene fibers: Flexural behavior of rc beams. In *Journal of the Civil Engineering Forum*, pages 40–52. doi:10.22146/jcef.23135.

C. E. Olvera-Mayorga, M. d. J. López-Martínez, J. A. Rodríguez-Rodríguez, S. Vázquez-Reyes, L. O. Solís-Sánchez, J. I. de la Rosa-Vargas, D. Duarte-Correa, J. V. González-Aviña, and C. A. Olvera-Olvera. Ai-based inference system for concrete compressive strength: Multi-dataset analysis of optimised machine learning algorithms. *Applied Sciences*, 15(23):12383, 2025. doi:10.3390/app152312383.

A. Pasini. Artificial neural networks for small dataset analysis. *Journal of Thoracic Disease*, 7(5):953–960, 2015. doi:10.3978/j.issn.2072-1439.2015.04.61.

O. W. Putri, A. F. Setiawan, S. Siswosukarto, M. A. Muflikhun, and N. M. Nor. A flexural behavior of full-scale rc beam strengthened using glass fiber reinforced polymer: Experimental research. In *Journal of the Civil Engineering Forum*, pages 152–169, May 2026. doi:10.22146/jcef.22499.

M. R. Raju, M. Rahman, M. M. Hasan, M. M. Islam, and M. S. Alam. Estimation of concrete materials' uniaxial compressive strength using soft computing techniques. *Heliyon*, 9(11), 2023. doi:10.1016/j.heliyon.2023.e22502.

P. Ruba, S. Aarthi, and P. Bhuvaneshwari. Prediction of ultimate strength for high-strength concrete (hsc) using machine learning approaches - optimised by pso technique. *Discover Computing*, 28:17, 2025. doi:10.1007/s10791-025-09517-1.

M. Ullah, V. Plevris, J. Mir, and A. Ahmad. Image-based prediction of concrete mechanical properties through cnn segmentation and ann modelling. *Results in Engineering*, page 108637, 2025. doi:10.1016/j.rineng.2025.108637.

M. Uzair and N. Jamil. Effects of hidden layers on neural network efficiency. In *2020 IEEE 23rd International Multitopic Conference (INMIC)*, pages 1–6. IEEE, November 2020. doi:10.1109/INMIC50486.2020.9318195.

V. Vakharia and R. Gujar. Prediction of compressive strength and portland cement composition using cross-validation and feature ranking techniques. *Construction and Building Materials*, 225:292–301, 2019. doi:10.1016/j.conbuildmat.2019.07.224.

J. Wang, Y. Xia, C. Zhang, C. Wang, B. Xue, R. Sun, et al. Evaluation of compressive strength of concrete durability degradation in dry and wet environments using destructive and non-destructive testing. *Measurement*, 223:113702, 2023. doi:10.1016/j.measurement.2023.113702.

Y. Yılmaz, S. Nayır, and Ş. Erdoğdu. A comprehensive study on the determination of compressive strength of concrete containing fly ash and blast furnace slag based on machine and ensemble learning. *Journal of Structural Engineering & Applied Mechanics*, 8(1):13–30, 2025. doi:10.31462/jseam.2025.01013030.

Y. Yu, C. Zhang, X. Xie, A. M. Yousefi, G. Zhang, J. Li, and B. Samali. Compressive strength evaluation of cement-based materials in a sulphate environment using optimised deep learning technology. *Developments in the Built Environment*, 16:100298, 2023. doi:10.1016/j.dibe.2023.100298.

S. Zuo and B. Liu. Optimisation design of concrete mix proportion based on support vector machine regression and enhanced genetic algorithm. *Discover Applied Sciences*, 7:200, 2025. doi:10.1007/s42452-025-06603-3.