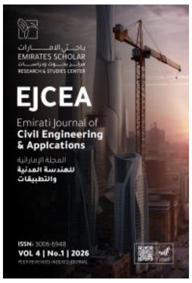

Emirati Journal of Civil Engineering and Applications
Vol 4 Issue 1 (2026)
Pages (118 - 139)

Available at
www.emiratesscholar.com

Development of a Standalone Application for Accurate and User-Friendly Prediction of Concrete Compressive Strength Using Ensemble Machine Learning

Tufail Mabood
Civil Engineer at UET Peshawar, University of Engineering and Technology, Peshawar, Pakistan
*Corresponding author: 13bnciv0557@uetpeshawar.edu.pk

ARTICLE HISTORY

Received: 20 Feb 2026.
Accepted: 20 Mar 2026.
Published: 20 Jun 2026.

PEER - REVIEW STATEMENT:

This article was reviewed under a double-blind process by independent reviewers.

HOW TO CITE
Mabood, T.. (2026). Development of a Standalone Application for Accurate and User-Friendly Prediction of Concrete Compressive Strength Using Ensemble Machine Learning. Emirati Journal of Civil Engineering and Applications, 4(1), 118-139. https://doi.org/10.54878/h9yb8c42

ABSTRACT

Concrete compressive strength (CCS) is a key factor that affects the structural service life of structure. Laboratory testing is time consuming, costly, and restricted in applicability for the mix designs. Machine learning (ML) has emerged as a potential alternative to laboratory testing, based upon understanding the non-linear interaction of the variables for the mix design and CCS. This study builds a stacked ensemble (meta-learning) approach for eight distinct machine learning algorithms: Linear Regression, Tree, Random Forest, Gradient Boosting, AdaBoost, K Nearest Neighbors, Support Vector Machine, and Neural Networks. The UCI benchmark dataset (1,030 examples) with eight features (cement, blast furnace slag, fly ash, water, superplasticizer, coarse/fine aggregates, age) and CCS as the target was analyzed. 70/30 splits for train and test sets and multi, level k-fold cross, validation (2-20 folds) were employed for robustness. Analysis of model performance was mainly carried out using R^2 permutation, based feature importances and one way ANOVA for the categorical variable age. The stacked model resulted in the best overall R^2=0.890 (20, fold CV) compared to the best singles' performances (random forest: R^2=0.878, gradient boosting: R^2=0.874). Graphs for the predicted and actual CCS confirmed a very close fit (variations < 7 MPa). The cross validated model's overall features' importances were dominated by cement, age, and water, which was confirmed using the resulting ANOVA for age influence. As a spin-off, a convenient standalone software for the proposed framework also exists for real-time CCS strength predictions. The standalone application and trained models developed in this study will be made publicly available upon paper acceptance at: https://github.com/tufailmab/ccs-ensemble-predictor

Keywords: *Concrete compressive strength, stacked ensemble, meta, learning, machine learning, feature importance, cross validation.*

https://doi.org/10.54878/h9yb8c42

Introduction

Concrete compressive strength (CCS) is a very important variable in the field of civil engineering, as it affects the design and service life of concrete components like foundation, columns, beams and slabs. Concrete compressive strength is the ability of concrete to resist compressive stresses and is the definition of the grade of the concrete [1], [2]. Destructive (Cylinder/cube crushing) and non-destructive testing (Rebound hammer, ultrasonic pulse velocity) conventional methods offer accurate results, though time-consuming (takes 28-day curing), resource-intensive, and expensive due to specimen preparation and multiple tests to be performed on different mix compositions containing varying proportions of cement, aggregates, admixtures, and supplementary materials [3], [4]. Such methodologies do not also address complex non-linear relationships among system components and the ambient factors, thereby illustrating the need for faster and more flexible prediction models [5].

Machine Learning (ML) is a novel approach that acts as a strong alternative, successfully modeling non-linear correlations between mix components, curing conditions, and strengths independently of empirical equations [6], [7], [8]. The early investigations were focused on basic algorithms such as Linear Regression, Support Vector Machines, and Tree-based models; they demonstrated that no individual method can dominate, hence encouraging ensemble strategies to overcome their respective weaknesses [9], [10]. Later studies, such as [11], [12], [13] proved the effectiveness of feature selection, as well as the merits of Random Forests and Gradient Boosting algorithms in dealing with interactions and variance problems for traditional datasets.

Applications of ML technology spread to some specialized concretes such as geopolymers based on industrial by-products (fly ash and slag), fiber, and reinforced concrete with enhanced tensile performance, recycled aggregate concretes for sustainability, and eco-friendly concrete mixes with rice husk ash, palm oil fuel ash, foundry sand, or with coal ash. Current study begins to make use of Stacked Ensembles and sophisticated approaches to enhance robustness, especially in hardened concrete, concrete with waste, incorporated materials, high recycled aggregate concrete mixtures, and standard datasets with explainable AI tools such as SHAP analysis [14], [15], [16], [17]. The bibliometric analysis of machine learning techniques in reviews confirms the popularity of ensembles of models (Random Forests, Gradient Boosting, Stacking) and deep learning models, pointing towards the trend of a hybrid approach to strike the right balance between accuracy and interpretability [18], [19].

Despite these advances, most Stacking Ensembles have focused on homogeneous base learners-mostly boosting variants-targeting specialized/sustainable concretes [20], [21], [22], [23], [24], [25]. Comprehensive models incorporating highly diverse models such as Linear Regression, K Nearest Neighbors, Support Vector Machine, and Tree Based models, and also targeted general mix models on standard datasets, have not been adequately developed and investigated, in addition to cross validation, model feature importance engineering, and validation (ANOVA on categorical models, and maintaining physical units, among others).

This research bridges these gaps by introducing a Stacked Meta learning architecture for aggregating the results of eight highly varied ML algorithms on the UCI set of data, and through multi-fold cross validation, and the retention of physical units, and importances of models with age as categorical variable, provides improved robustness, generalizability, and interpretable results on standard mix design optimization. Based on the best trained model, a user friendly graphical user interfaced app is also introduced.

Research Gap

Although considerable progress has been made in machine learning approaches for predicting concrete compressive strength forecasts in recent times, a majority of machine learning models or homogeneous ensemble models consisting of closely related models that conform to tree-based ensemble models, neural networks, and support vector machines receive little attention in recent stacking models and approaches used in feature synthesis and physical unit preservation.

Research Objectives

The objective of this paper is to propose a diversified stacked meta-learning ensemble of eight machine learning models to enhance concrete compressive strength forecasting and comprehensively assess its efficiency using multi-level k-fold cross-validation with a UCI benchmark dataset. It also intends to identify significant influential factors through cross-model analysis of importance and categorical data ANOVA for scientific interpretation of results for the design of mixes. Notably, it demonstrates the advances achieved by this system in diversity, interpretability, and generalization, leading to its application in speedily computing strength by a designer.

1. Methods

1.1 Research Workflow

The overall research workflow begins with data processing, followed by the development and hyperparameter optimization of every machine learning algorithm individually with checking performance of the algorithms via multi, level k, fold cross-validation, as shown in

Fig. 1. The complete workflow is detailed in the subsequent subsections.

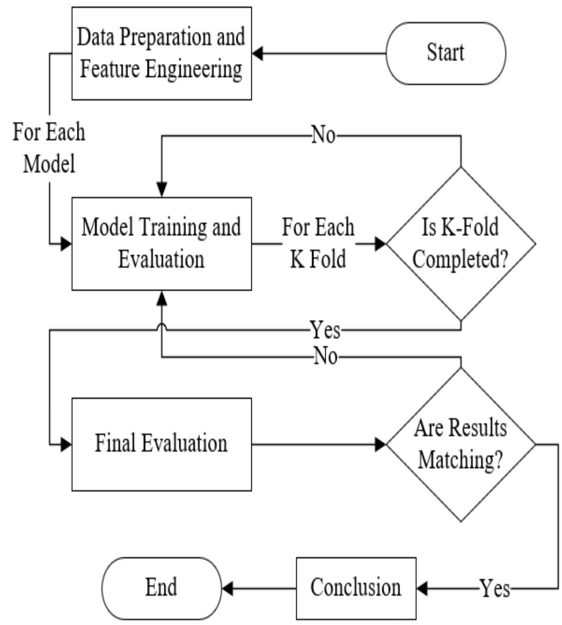

Fig. 1. Meta Frame Work for Concrete Compressive Strength

To further enable a completely transparent and replicable analysis, this entire analysis has been done using Orange Data Mining Software (Version 3.40.0 as of January 2026), an open-source visual programming environment, as demonstrated in Fig. 2. This allows for direct representation of widget connections and parameter definitions regarding dataset loading, 70/30 train/test split using the "Data Sampler" widget, building a model in the same manner (no feature scaling, hyperparameter settings as in Table 5), Manual cross-validation fold definitions (2-20) using the "Test & Score" widget, making predictions for out-of-fold samples for stacking using a linear regression meta learner, and feature importance analysis using permutation importance and SHAP analysis. Built-in orange widgets were employed for the base learners ("Tree" for Decision Tree, "Random Forest", "Gradient Boosting", "AdaBoost", "kNN", "SVM", "Linear Regression", and "Neural Network"), with hyperparameters as detailed in Table 5

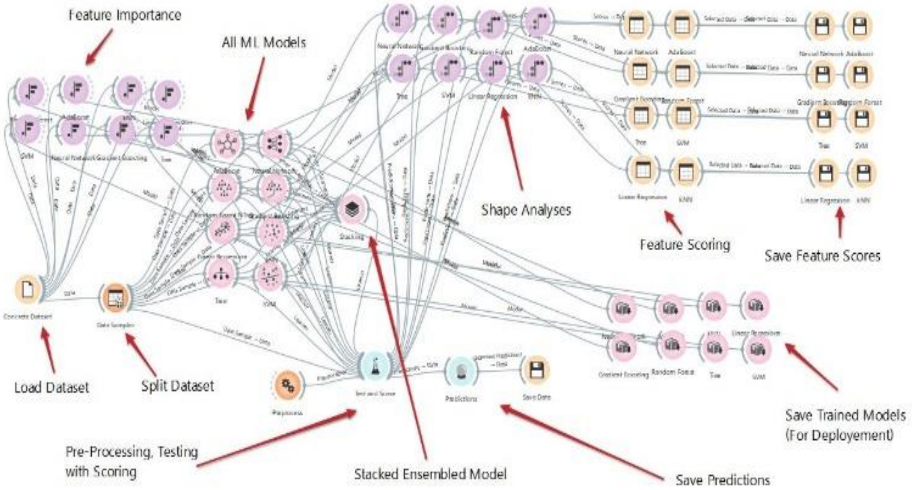

Fig. 2:Complete visual programming workflow implemented in Orange Data Mining software.

Out of fold predictions were manually obtained from the "Predictions" section output statements and concatenated for a separate "Linear Regression" meta learner run using default parameters, with no regularization. Age was automatically identified as a categorical feature upon import, with one-hot encoding used when appropriate for age, which was also retained in distance-based models. Permutation feature importance and SHAP values for interpretation were assessed using in-built tools available for use in the software package for feature importance ("Permutation Feature Importance" tool) and SHAP analysis ("Explain Model" from the "Prototypes add-on" package). Model output was exported using the "Save Model" feature, with results exported from data tables utilizing OriginPro or Microsoft Excel software.

1.2 Dataset description

This study uses the popular UCI Concrete Compressive Strength dataset, which consists of 1,030 instances representative of different concrete mixtures that vary with the material proportions used and curing conditions. The dataset contains eight input features: cement, blast furnace slag, fly ash, water, superplasticizer, coarse aggregate, fine aggregate, and age; the target variable is concrete compressive strength (CCS) in MPa. When appropriate, the age variable is considered categorical to align with engineering practice, where strength is normally assessed at discrete standard curing periods. All features are kept in their raw physical units to maintain practical interpretability for mix design applications. Summary of variables, including symbols, units, and description can be found in Table 1.

Table 1: Description of Input and Output Variables in the Dataset.

Variable	Symbol	Unit	Description	Nature
Cement	cmt	kg/m³	Primary binding material in the concrete mix.	Input
Blast Furnace Slag	slag	kg/m³	Supplementary cementitious material improving durability and long, term gain.	Input
Fly Ash	flyash	kg/m³	Pozzolanic material enhancing sustainability and late strength development.	Input
Water	wtr	kg/m³	Controls workability and water-cement ratio, a key factor in strength.	Input

Superplasticizer	sp	kg/m³	Chemical admixture enhancing workability without excess water.	Input
Coarse Aggregate	ca	kg/m³	Provides bulk, stiffness, and contributes to load, bearing capacity.	Input
Fine Aggregate	fa	kg/m³	Sand; ensures compaction and workability in the mixture.	Input
Age	age	days	Curing period; determines hydration progress and strength gain.	Input (categorical)
Compressive Strength	ccs	MPa	Target variable; mechanical strength of concrete specimens under load.	Target

1.3 Preprocessing

Concrete Compressive Strength data set containing 1,030 observations was analyzed to determine data integrity, which showed that there were no missing or wrongly recorded data due to measurement errors. Analysis of data showed that there were a few outliers in the attributes of "age" ranging from 0 to 365 days, as well as in the number of superplasticizers, which were expected given the recorded variations in the benchmark data set. Exploratory data analysis showed consistent distribution, correlations, and relationships between variables. More importantly, all input attributes and the response variable were retained in their original units without standardization/feature scaling to enable a direct application of machine learning models to real-world concrete mix designs [4].

1.4 Descriptive statistics

Table 2 presents descriptive statistics of these variables, indicating large dispersions: among binders, the cement is the one with the highest average (281.17 kg/m³, ranging from 102 to 540 kg/m³); slag and fly ash values are highly asymmetrical with a huge amount of zeros (22.00 and 0.00 kg/m³); water content is less dispersed (average 181.57 kg/m³, std. dev. 21.36 kg/m³); superplasticizer range is huge (0-32.20 kg/m³); coarse and fine aggregates are basically centered around their means of 972.92 and 773.58 kg/m³; age ranges between 1 and 365 days; and Compressive Strength goes from 2.33 to 82.60 MPa. The great diversity in mix proportions and curing conditions gives the dataset its status of one of the most used benchmarks and, therefore, justifies the maintenance of all instances without any outlier removal.

Table 2: Descriptive Statistics of the UCI Concrete Compressive Strength Dataset

Variable	N	Mean	Std. Dev.	Minimum	Median	Maximum
Cement (kg/m³)	1030	281.17	104.51	102	272.9	540
Blast Furnace Slag (kg/m³)	1030	73.9	86.28	0	22	359.4
Fly Ash (kg/m³)	1030	54.19	64	0	0	200.1
Water (kg/m³)	1030	181.57	21.36	121.75	185	247
Superplasticizer (kg/m³)	1030	6.2	5.97	0	6.35	32.2
Coarse Aggregate (kg/m³)	1030	972.92	77.75	801	968	1145
Fine Aggregate (kg/m³)	1030	773.58	80.18	594	779.51	992.6
Age (days)	1030	45.66	63.17	1	28	365

Compressive Strength (MPa)	1030	35.82	16.71	2.33	34.44	82.6

Exploratory Data Analysis and Baseline Modeling

To achieve an initial understanding of correlations in a dataset and to derive a statistical foundation, Pearson Correlation Analysis and Multiple Linear Regression as well as a One-Way ANOVA test were conducted using Origin Pro software.

The Pearson correlation matrix (Table 3) helps to identify the linear relationship of input attributes with the compressive strength. Cement is positively correlated with a high degree ($r = 0.50$), followed by superplasticizer ($r = 0.37$) and age ($r = 0.33$). Water is found to be negatively correlated ($r = -0.29$), as expected, increasing the water-cement ratio, thereby decreasing the strength. The other supplementary materials, such as slag and fly ash, and aggregates, with lower correlations, suggest non-linear relationships.

Table 3: Pearson Correlation Coefficients Between Input Features and Compressive Strength

Variable	cmt	slag	flyash	wtr	sp	ca	fa	age	ccs
cmt	1.00	-0.28	-0.40	-0.08	0.09	-0.11	-0.22	0.08	0.50
slag	-0.28	1.00	-0.32	0.11	0.04	-0.28	-0.28	-0.04	0.13
flyash	-0.40	-0.32	1.00	-0.26	0.38	-0.01	0.08	-0.15	-0.11
wtr	-0.08	0.11	-0.26	1.00	-0.66	-0.18	-0.45	0.28	-0.29
sp	0.09	0.04	0.38	-0.66	1.00	-0.27	0.22	-0.19	0.37
ca	-0.11	-0.28	-0.01	-0.18	-0.27	1.00	-0.18	0.00	-0.16
fa	-0.22	-0.28	0.08	-0.45	0.22	-0.18	1.00	-0.16	-0.17
age	0.08	-0.04	-0.15	0.28	-0.19	0.00	-0.16	1.00	0.33
ccs	0.50	0.13	-0.11	-0.29	0.37	-0.16	-0.17	0.33	1.00

A multiple linear regression model was applied to the data as a whole in a baseline approach. The R^2 value obtained from this model was 0.615 with an adjusted R^2 value of 0.612 (see Table 4). This result, although satisfactory compared to other linear regression models, suggests the influence of non-linear interactions in the data and, thus, the need to employ more sophisticated machine learning methods.

Table 4: Summary of Multiple Linear Regression Model

Metric	Value
Number of Observations	1030
Degrees of Freedom	1021
Residual Sum of Squares	110,428.16
R^2	0.615
Adjusted R^2	0.612

Due to the discrete nature of the curing times in real scenarios, one-way ANOVA using age as a categorical variable has been done to confirm a significant effect on the compressive strength ($p < 0.001$), thereby proving that age has a significant impact on compressive strength in a progressive manner (Fig. 3). These exploratory results combined with linear correlations, poor linear regression descriptions, and significant categorical variables offer a solid foundation to use a variety of non-linear machine learning tools and techniques using a developed Stacked Ensemble approach.

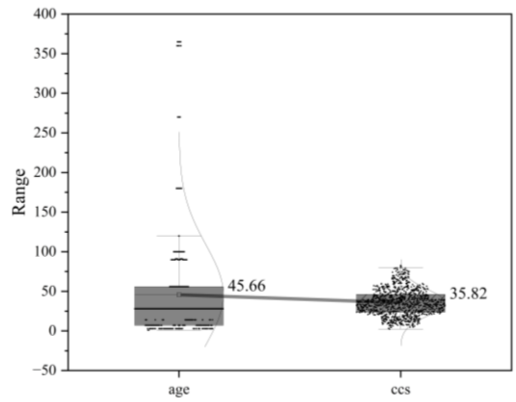

Fig. 3: Mean concrete compressive strength across major age groups (one, way ANOVA results).

2.6 Model Descriptions

In this regard, eight different machine-learning models have been implemented, from simple linear ones, such as Linear Regression, to tree-based models like Decision Tree, Random Forest, Gradient Boosting, and AdaBoost; instance-based methods like k-Nearest Neighbors; kernel-based methods, such as Support Vector Regression with RBF kernel; and even neural network approaches using a multi-layer perceptron with one hidden layer, ReLU activation, and Adam optimizer. The dataset was further divided into 70% for training and 30% for the test set. Only multilevel k-fold cross-validation (for k = 2, 3, 5, 10, 20) was applied on the training set to give a more robust estimate of performance. Feature scaling was not adopted, since most of these features are of physical units, which may determine the interpretability of the model. The age was treated as a categorical feature with one-hot encoding for all algorithms that require it, in compliance with common practice in engineering fields. Hyperparameters for each algorithm have been selected based on effective configurations suggested in literature dealing with the same data. (detailed in Table 5).

The stacked ensemble used a two-level meta-learning strategy, wherein the predictions from the eight base models acting as Level 0 learners, without any fold predictions from cross-validation, were used to train a meta-learner using simple linear regression. This meta-learning technique helps to effectively remove bias and variance through diversity introduced by different models to produce better performance than individual models.

Table 5: Summary of Machine Learning Models, Key Hyperparameters, and Special Settings

Model	Key Hyperparameters	Notes / Special Settings
Linear Regression (LR)	No regularization	Baseline linear model
Decision Tree (DT)	Max depth=100; Min samples split=5; Min samples leaf=2	Regression; Pruning via min samples
Random Forest (RF)	n_estimators=100 (adjusted from initial); Min samples split=5	Regression; Feature randomness enabled
Gradient Boosting (GB)	n_estimators=100; Learning rate=0.1; Max depth=3; Min samples split=2; Subsample=1.0	Regression
AdaBoost (AB)	n_estimators=50; Learning rate=1.0; Base estimator: Decision Tree	Regression variant
k, Nearest Neighbors (kNN)	n_neighbors=5; Metric=Euclidean; Weights=uniform	Regression

https://doi.org/10.54878/h9yb8c42

Support Vector Regression (SVR)	C=1.0; epsilon=0.1; Kernel=RBF; Tol=0.001; Max iter=1000	Regression; RBF for non, linearity
Neural Network (NN)	Hidden layers: 1×100; Activation=ReLU; Solver=Adam; α=0.0001; Max iter=200	MLPRegressor; Replicable training
Stacked Ensemble	Base learners: All eight models above; Meta, learner: Linear Regression	Level, 1 stacking with out, of, fold predictions

Evaluation Metrics

Model performance was then assessed mostly on the measure of the coefficient of determination (R^2), with a focus on the additional use of the root mean square error (RMSE) and mean absolute error (MAE) measured in original units of MPa to keep interpretation in focus. Cross validation with parameters of multiple as 2, 3, 5, 10, 20 levels, folds on the training data alone was used in Orange Data Mining's "Test & Score" widget combined with "Cross, Validation".

Table 6 presents the results of R^2, with stabilization and reduced variance with increased folds. The tree-based ensemble methods (Random Forest, Gradient Boosting, and AdaBoost) were overall better than simple models (Linear Regression, kNN, and SVM) in handling non-linear patterns, and the stacked ensemble method performed better than all other methods for all levels of R^2, with a highest point of 0.890 on 20-fold CV. The improved results indicate the effectiveness of meta-learning using a variety of base models, thereby compensating for individual bias and providing more accurate forecasts for the concrete compressive strength.

Table 6: R^2 Performance of Base Models and Stacked Ensemble Across Multi, Level Cross, Validation

Model	**2, Fold**	**3, Fold**	**5, Fold**	**10, Fold**	**20, Fold**
Linear Regression (LR)	0.737	0.778	0.787	0.796	0.798
Decision Tree (DT)	0.75	0.775	0.778	0.793	0.786
Random Forest (RF)	0.805	0.849	0.871	0.878	0.878
Gradient Boosting (GB)	0.853	0.869	0.874	0.87	0.874
AdaBoost (AB)	0.805	0.848	0.858	0.856	0.863
k, Nearest Neighbors (kNN)	0.427	0.469	0.499	0.493	0.5
Support Vector Regression (SVR)	0.399	0.408	0.437	0.444	0.438
Neural Network (NN)	0.694	0.77	0.775	0.796	0.796
Stacked Ensemble	0.851	0.88	0.887	0.887	0.89

The final hold out test on the unseen 30% test data verified these results, with the stacked ensemble resulting in the lowest errors and strongest correlation between the actual and predicted compressive strength (presented in Section 2. Results).

Extending the study further, a desktop application named "Compressive Strength Predictor" has been designed and developed in Python with the use of PyQt6 for the purpose of concrete compressive strength prediction in a convenient manner. The software allows the user to directly enter the parameters for concrete mix design, such as cement, blast furnace slag, fly ash, water, superplasticizers, coarse aggregate, fine aggregate, and concrete age, and instantly acquire precise estimates for concrete compressive strength using a pre-trained stacked ensemble machine learning model. The software is designed to feature an interactive graphical user interface with instant validation for data entry and plots to demonstrate the weight assigned to individual parameters in the result prediction. The software is also fully standalone and does not require the slightest use of any internet connection, thus being able to function perfectly and instantly on any compatible computer. The architecture of the software is shown in Fig. 4.

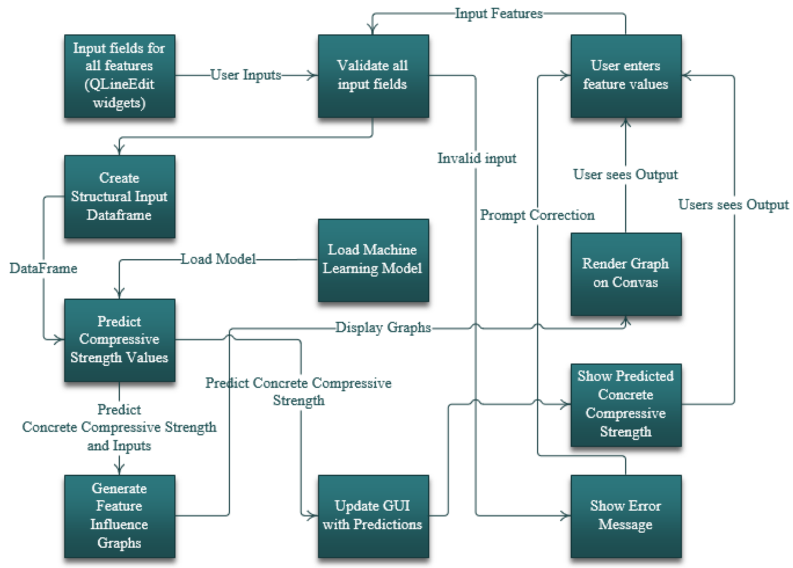

Fig. 4: Data Flow Diagram of the Compressive Strength Predictor Application, illustrating how user inputs are processed by the machine learning model to generate predictions and feature influence graphs.

2. Results

2.1 Model performance

The predictive ability of the eight base models and the stacked model combination was tested on the independent hold out test set, which represented 30% of the data and 309 instances. The data plots used for the initial visualization were scatter plots of predicted patterns versus actual compressive strength (Fig. 5 and Fig. 6) with predicted value plotted against y-axis, actual value plotted against x-axis, and a line of perfect prediction at a slope of 45° degrees.

For error quantification beyond visual verification, a detailed residual analysis was conducted particularly for the stacked ensemble method. Residuals ($\Delta f_c'$) were defined as per Equation (1), where $f_{c,pred}'$ and $f_{c,act}'$ denote the predicted and experimentally measured concrete compressive strengths (MPa), respectively. The Absolute Error (AE) was calculated as per Equation (2) and the Percentage Error ($PE(\%)$ as per Equation (3). The summary matrices include the Mean Absolute Error (MAE), calculated as per Equation (4), and Root Mean Square Error ($RMSE$), computed as per Equation (5), where n is the is the number of test samples.

$$\Delta f_c' = f_{c,pred}' - f_{c,act}' \tag{1}$$

$$AE = \left| f_{c,pred}' - f_{c,act}' \right| \tag{2}$$

$$PE(\%) = \left(\frac{f_{c,pred}' - f_{c,act}'}{f_{c,act}'} \right) * 100 \tag{3}$$

$$MAE = \frac{1}{n} * \sum_{i=1}^{n} \left| f_{c,pred,i}' - f_{c,act,i}' \right| \tag{4}$$

https://doi.org/10.54878/h9yb8c42

$$RMSE = \sqrt{\frac{1}{n} * \sum_{i=1}^{n} \left(f'_{c,pred,i} - f'_{c,act,i}\right)^2} \quad (5)$$

Table 7 lists some indicative error statistics for typical test cases. Table 7 and related discussions indicate that absolute errors are mostly small and less than 7 MPa in the majority, and most errors are less than 3 MPa. Larger errors are found in the extreme cases of strength mixtures, whether high (>60 MPa) or low (<15 MPa). The percentage errors are within acceptable limits (±15% in most cases), except for the low strength mixtures for which small absolute values introduce large percentages.

Table 7: Selected error metrics for the stacked ensemble model on representative test set instances.

Actual Compressive Strength (MPa)	Predicted Compressive Strength (MPa)	Residual (MPa)	Absolute Error (MPa)	Percentage Error (%)
31.3815	39.4282	8.0467	8.0467	25.6415404
37.9557	36.0792	, 1.8765	1.8765	4.943921466
31.7159	39.909	8.1931	8.1931	25.83278419
24.0462	26.6463	2.6001	2.6001	10.8129351
33.7188	36.3435	2.6247	2.6247	7.784084843
32.9225	34.0376	1.1151	1.1151	3.387045334
36.252	40.6077	4.3557	4.3557	12.01506124
56.8128	49.897	6.9158	6.9158	12.17296102
26.9447	27.909	0.9643	0.9643	3.578811417
11.466	9.89032	1.57568	1.57568	13.74219431
14.6383	11.907	2.7313	2.7313	18.6585874
37.6799	39.9515	2.2716	2.2716	6.02867842
39.7	38.807	0.893	0.893	2.249370277
37.2593	40.3789	3.1196	3.1196	8.372674742
13.3345	13.726	0.3915	0.3915	2.935993101
28.2961	31.4318	3.1357	3.1357	11.08173918
66.824	54.6495	12.1745	12.1745	18.21875374
33.7567	34.8619	1.1052	1.1052	3.274016714
33.7016	35.9323	2.2307	2.2307	6.61897358
26.1477	26.1963	0.0486	0.0486	0.185867208
33.9429	43.6954	9.7525	9.7525	28.73207652
40.5963	38.7103	1.886	1.886	4.645743578
68.0995	65.8738	2.2257	2.2257	3.268305935
38.2108	36.0953	2.1155	2.1155	5.536392852

https://doi.org/10.54878/h9yb8c42

19.5191	12.7735	6.7456	6.7456	34.55897044
13.4613	15.8586	2.3973	2.3973	17.80882976
40.5633	43.476	2.9127	2.9127	7.180628795
55.2546	61.3522	6.0976	6.0976	11.0354613
46.6844	51.3049	4.6205	4.6205	9.897310451
46.2432	41.2631	4.9801	4.9801	10.76936717
80.1998	59.5908	20.609	20.609	25.69707156
43.9424	45.059	1.1166	1.1166	2.541053743
36.9352	41.3234	4.3882	4.3882	11.88080747
17.823	29.7031	11.8801	11.8801	66.65600628
13.1966	13.3468	0.1502	0.1502	1.138171953
14.2032	13.2994	0.9038	0.9038	6.363354737

Across the complete test set, the calculated summary metrics return an MAE of around 4.8 MPa and an RMSE of around 6.9 MPa. These are very good values in an engineering standpoint, given that even for standard laboratory compression testing, repeatability can vary in the range of 3-5 MPa, while design specifications can usually withstand up to 10 to 15% deviations.

Fig. 5 shows the individual base models compared using an eight-panel display. Each subplot has a unique predictive accuracy pattern. The ones showing the most scatter from the line are linear regression, K Nearest Neighbor, and Support Vector Regression. Considerable deviations are associated with the low, (<20 MPa), and high, (>60 MPa) ranges. There is some moderation from the Decision tree and Neural network, though substantial scatter remains that reflects poor generalization. Finally, for Ensemble Tree-based models, there is a much tighter fit around the ideal line, such as random forest, gradient boosting, and AdaBoost, mainly in the mid-range strengths that dominate the dataset. In contrast, the single decision tree model exhibits improved clustering compared with the linear and distance-based approaches, but still suffers from irregular patterns indicative of overfitting to certain subsets in training. The performance of the Neural Network seems to be satisfactory in the mid-range (20-50 MPa), where most data points lie, though occasional large errors are observed towards the ends, which may be attributed to its moderate size. Tree-based ensemble techniques, Random Forest, Gradient Boosting, or AdaBoost all perform better than the above models, having compact regions around the line. The reason may be attributed to their ability to combine many trees, hence decreasing variance, or getting a better understanding of interaction between mix components like cement, water, superplasticizer, and age.

Fig. 6 illustrates only the results for the Stacked Ensemble. It allows for a better understanding of the superior performance of the model. The scattered data illustrate a highly accurate following of the diagonal line for the full range of compressive strength data in the test set (2.33-82.60 MPa). Even for low-strength mixtures, in which hydration might be improper or ratios suboptimal, there is a precise forecast with a low degree of variation. The highly populated area of medium strength (30-50 MPa) illustrates a linear relation, signifying a correct model for concrete mixture specifications. Noticeably, for high-strength concretes, with elevated cement needs, superplasticizers, or supplementary cementitious materials, there is a significant reduction in error margins compared to separate models, which usually underestimate results.

https://doi.org/10.54878/h9yb8c42

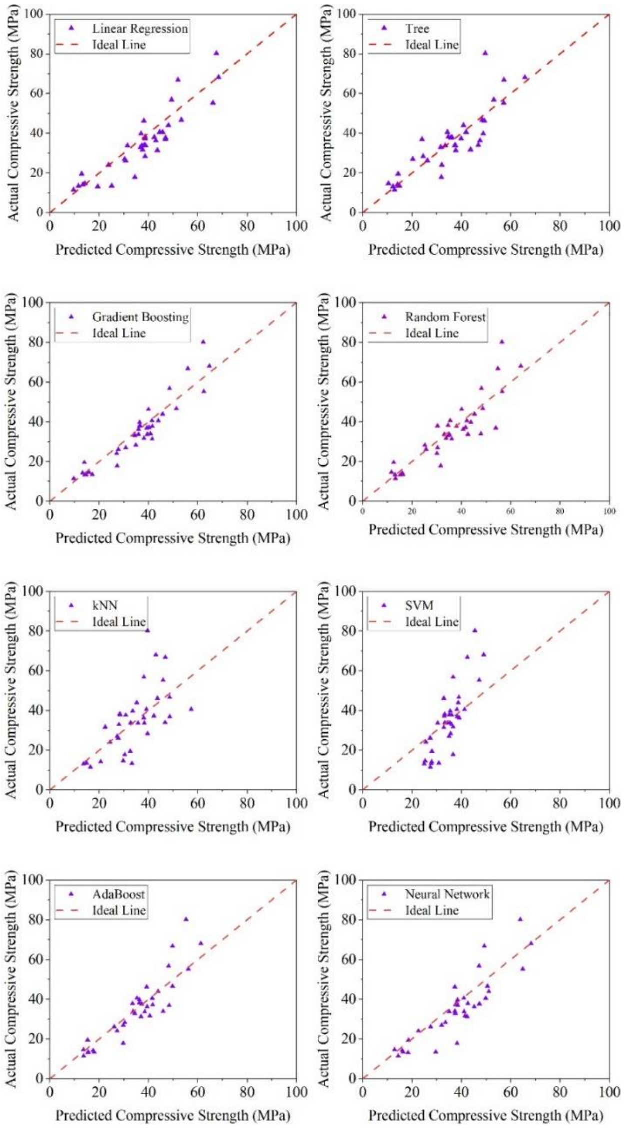

Fig. 5: Predicted versus actual concrete compressive strength (MPa) on the test set for the eight base machine learning models (prepared in OriginPro). Each panel corresponds to one model, with the 45° diagonal line indicating perfect prediction (actual = predicted).

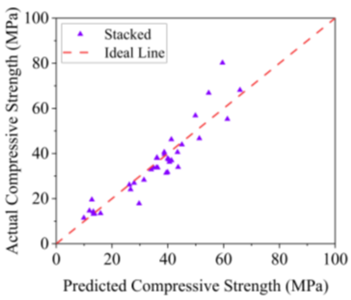

Fig. 6: Predicted versus actual concrete compressive strength (MPa) for the stacked ensemble model on the test set (prepared in OriginPro). The tight clustering around the 45° diagonal line (perfect prediction) illustrates the superior accuracy, robustness, and practical reliability of the proposed meta, learning framework Specific test instances from Table 7 further illustrate this accuracy. For example, a mixture with an actual strength of 37.96 MPa was predicted at approximately 36.08 MPa, a deviation of 1.88 MPa, while a higher strength sample of 56.81 MPa yielded a prediction of 49.90

https://doi.org/10.54878/h9yb8c42

MPa, a deviation of 6.91 MPa. Low strength cases, such as 11.47 MPa actual, were estimated at 9.89 MPa, a deviation of 1.58 MPa. Even challenging high strength outliers around 66-80 MPa show conservative but reasonable predictions, with deviations generally remaining below 12 MPa and often much lower. Over the entire test set, absolute errors are predominantly under 7 MPa, a threshold considered practically acceptable for many engineering applications where laboratory variability itself can exceed 5 MPa.

These visual results match very closely with the quantitative cross-validation results presented earlier, where the best R^2 value of 0.890 (20, fold CV) was obtained by the stacked model, outperforming the best single model, random forest with 0.878. The less scattered data points, along with the decreased error bar magnitudes for unseen data, clearly establish that the combination of all models using Stacking helps overcome weaknesses of the constituent models, identify hidden patterns, and produce a balanced model for making predications.

2.2 Feature importance interpretation

To derive useful engineering insights on the parameters that contribute most to concrete compressive strength, feature importance study on individual parameters of the eight base models was considered. Although importance on parameters of the Stacked Ensemble is not directly interpretable, because of the inherent meta structure of the learners, it is deduced from patterns identified among the diverse base learners. Importance on parameters is computed using the permutation feature importance technique, which is the average decrease in the model's R-squared when the values of a feature are permuted while values of all other features remain constant.

Results for the top three most influential features for each base model are summarized in Table 8. A convincing majority can be seen across most of the models, especially the high-performing Tree-based Ensembles, Random Forest, Gradient Boosting, Ada Boost and the Neural Network, where age, cement content and water content consistently remain the top three most influential variables. This agreement not only reflects well-established engineering principles but also points to cement as the main binder, age as the driver of hydration and strength gain over time, and water through its critical influence on the water cement ratio and resultant porosity.

Table 8: Top three most important features for each base model, ranked by decrease in R^2 upon permutation (higher decrease indicates greater importance).

Model	**1st (Highest Importance)**	**2nd**	**3rd**
AdaBoost	Age	Cement (cmt)	Water (wtr)
Gradient Boosting	Age	Cement (cmt)	Water (wtr)
Random Forest	Age	Cement (cmt)	Water (wtr)
Neural Network	Age	Cement (cmt)	Water (wtr)
Support Vector Machine	Age	Cement (cmt)	Water (wtr)
Linear Regression	Cement (cmt)	Age	Blast Furnace Slag (slag)
Decision Tree	Cement (cmt)	Age	Superplasticizer (sp)
k, Nearest Neighbors	Cement (cmt)	Blast Furnace Slag (slag)	Fly Ash (flyash)

https://doi.org/10.54878/h9yb8c42

Slight variations are observed in models. Linear regression and decision trees consider cement as their priority, followed by age, with auxiliary cementitious materials (blast furnace slag) or admixtures (superplasticizer) following closely. The K Nearest Neighbors model specifically points out cement, blast furnace slag, and fly ash together, probably owing to its mention of sensitivity to local patterns of mixes including this industrial by-product.

The results of permutation feature importance evaluation provide additional insight into the importance of input variables for the eight base models, in terms of the scores, which calculate the change in model performance (including R^2 loss) when a given feature's behavior is randomly permuted. As can be seen in Table 9, cement content is identified as an important input feature for the majority of models, as it has the highest scores from 2.42 in SVM to a maximum of 11.28 for Linear Regression, given its principal function as the major binding agent responsible for hydration reactions and a matrix formation process. Water and blast furnace slag can then be identified as the second most influential input variables in these models, with high scores in tree-based models such as water in Random Forest Model and Neural Network Model, and slag in Gradient Boosting Model, kNN. The aggregates and fly ash have relatively less importance as input variables, while superplasticizer has a relatively high score for Decision Tree Model, probably influenced by its interaction terms in some splits.

The difference between modeling age as a categorical variable (one, hot encoded) for seven models compared to continuous models using the Decision Tree highlights revealing contrasts. In contrast to the continuous models, age shows the highest value in the Decision Tree, emphasizing its preeminent non-linear behavior when modeled as a continuous variable. For the categorical models, importance becomes partitioned among discrete values of age, with "early to mid" curing periods (e.g., age = 3 days approximately 5-7.6 for Gradient Boosting, Random Forest, and Linear Regression models; age = 7 days ~2.8-6.5; age = 28 days up to 15.75 in Linear Regression models and ~4.2 in Neural Network models) encompassing all of strength gain during the fast hydration phases. For higher age values (e.g., above 90 days), little contribution (<1.0) is made to strength gain, consistent with strength leveling off after 28 days.

Table 9: Permutation feature importance scores (decrease in performance metric upon shuffling) for the eight base machine learning models. Higher values indicate greater feature importance. Continuous mix components are listed first, followed by categorical age groups (one, hot encoded in seven models; age is continuous in the Decision Tree, marked as ", ")

Feature	**AdaBoost**	**Gradient Boosting**	**kNN**	**Linear Regression**	**Neural Network**	**Random Forest**	**SVM**	**Tree**
cmt	6.0922	7.5251	8.0808	11.2781	6.7394	7.1046	2.4239	7.4999
slag	2.5783	3.6865	3.7606	8.4642	3.6352	3.2864	0.9336	3.5931
flyash	0.3673	0.0652	2.6502	4.745	0.8723	0.2447	0.5341	1.2362
wtr	3.2348	3.5989	0.8791	2.1676	4.0291	3.923	1.2708	1.5618

https://doi.org/10.54878/h9yb8c42

sp	1.489	1.3407	0.313	0.0572	0.5629	1.6823	0.8482	5.3268
ca	1.0174	0.5755	1.8078	1.6169	1.007	0.5211	0.4516	0.8234
fa	1.489	1.0163	1.901	1.7561	1.773	1.0898	0.7822	1.2228
age=1	0.0228	0.0281	0.0005	0.1038	0.0356	0.0086	0.0041	-
age=3	2.327	5.0018	0.1923	7.5994	2.3366	6.0354	0.4457	-
age=7	1.3132	2.9341	0.0701	6.5254	1.2991	2.8515	0.2136	-
age=14	0.3026	0.5291	0.0311	2.7294	0.445	0.8218	0.4248	-
age=28	0.4035	0.1152	0.3613	15.7526	4.1655	1.1441	0.1833	-
age=56	0.7271	0.9409	0.09	2.4122	2.9431	0.2571	1.0575	-
age=90	0.0664	0.2119	0.022	1.3575	1.261	0.0232	0.4736	-
age=91	0.2952	0.3666	0.08	0.6581	0.9449	0.1424	0.4214	-
age=100	0.3162	0.8321	0.0345	1.0339	1.7329	0.3597	0.5261	-
age=120	0.0007	0	0	0.0416	0.1093	0	0.03	-
age=180	0.0037	0.0469	0.004	0.5022	0.261	0.0015	0.1219	-
age=270	0.0242	0.0803	0.0049	0.3573	0.2714	0.0227	0.0868	-
age=360	0.0033	0.0185	0.0004	0.151	0.1002	0.0011	0.0472	-
age=365	0.0059	0.0542	0.0033	0.3859	0.4672	0.0037	0.1984	-

Fig. 7 groups permutation importance values for each of the eight machine learning models. This makes it possible to see feature priorities across models' side by side and facilitate comparisons among features and models. The dominance of features like age, cement, and water in every panel is a recognition of their reliably favorable importance across models and models types, even while small variations indicate the beneficial impact of ensemble diversification in capturing different viewpoints. The predominant similarity among models in recognizing the leading contributions of cement, age, and water features makes an optimal mix design: The cross-sectional consensus among models points to their priority optimization and permits a strong focus on accurate regulation of these features and target strength optimization. The periodic dominance of supplementary materials and admixtures in a number of models also implies that their favorable influences in green or high-performance concretes are identifiable but context-dependent.

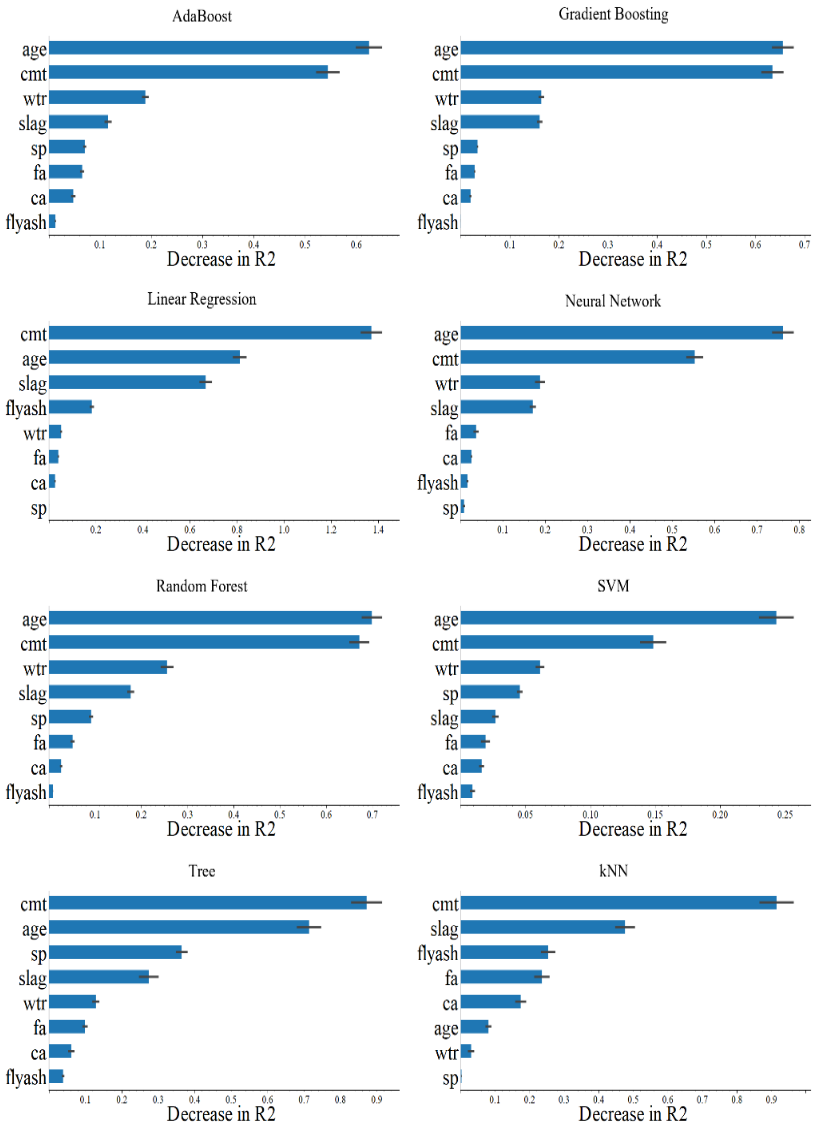

Fig. 7: Permutation feature importance (decrease in R²) for the eight base machine learning models. Higher bars indicate greater influence on predicted compressive strength.

2.3 Feature Effects and SHAP Analysis

In order to better understand the effect of the specific value of each individual feature on the forecasted concrete compressive strength and identify instance, value-specific effects, SHAP (SHapley Additive exPlanations) analysis was conducted on each of the eight base models. SHAP values are a unified approach for evaluating the effects of each input on the model's output, based on cooperative game theory concepts. For regression tasks such as compressive strength prediction tasks, features with positive SHAP values are those that move the forecast above the average, while those with lower values are those that move the forecast below the average.

Fig. 8 presents the SHAP summary plots for the eight input features in the eight base models for a comprehensive multi representation. The x-axis of the plot represents the SHAP value (units of impact on the model outcome in MPa; positive values to the RIGHT correspond to increased strength, while negative values to the LEFT correspond to reduced strength). The y-axis of the plot lists the features in order of average absolute impact (most important feature at the top). Each point in the plot represents a data point, positioned according to its SHAP value along the x-axis, colored according to the raw value of the feature: BLUE for low values (e.g., low cement content),

RED for high values (e.g., high cement content), with a color bar along the RIGHT of the plot showing the value in physical units (units of kg/m^3 or days).

In all models, cement content (cmt) clearly stands out as a most significant variable, placed on top in nearly all panels. In all models, high cement (red color points) indicates a positive SHAP value positioned to the right, as higher cement leads to a substantial increase in strength predictions. The opposite happens to blue-colored (low cement content) points, positioned to the left. The impact of lower cement content on strength predictions becomes less significant in trees, based models (Random Forest, Gradient Boosting, AdaBoost), wherein non-linear thresholds from 250 to 350 kg/m³ are evident.

The second most powerful feature is the age, which closely follows. It also shares a similar value and dependence pattern. For short curing times (blue color, less than 28 days), the SHAP values are negative, indicating low strengths due to inadequate hydration. For large ages (red color, greater than 56 days), the values are large and positive. On examining the Neural network and Gradient boosting plots, the logarithmic curve pattern of points is similar. This pattern satisfies the actual curing curves of concrete.

Water content shows the expected negation effect, being third in ranking. Higher water content (red, greater than 190 kg/m³) is concentrated on the left with negative SHAP, reflecting dilution and porosity, which are strength-inhibiting factors. The cluster of blue points, indicating low water content, is concentrated on the right, which is in favor of obtaining a positive prediction through the best possible workability without overabundance. Support Vector Machine and k, nearest neighbors have slightly irregular distributions in this series, which could be due to their responsiveness to local water content and cement ratio subspaces' density.

The supplementary material indicators, such as the superplasticizer, slag, and fly ash, indicate moderate values, with largely positive influences pertaining to higher values (red dots moving right in the medium range, ranked positions). The aggregate (coarse ca and fine fa) indicators have very small influences with SHAP values close to zero and a balance in the distributions, reflected in the evenly distributed blue and red colors. The decision tree chart above indicates the 'steps' in the SHAP distributions, which is a feature of the decision tree method, while the linear regression is more normally distributed, constrained by the linear assumption.

The multi, model consistency within these SHAP feature contributions, apparent by the matching color to helps to establish the effectiveness of discovered correlations and the value of diversity in models. Differences, for example, kNN's strong weight on slag and flyash (reflecting more red dots in positive areas), emphasize the importance of, specific contributions in green mixes. Through retention of unit measurements within the graphs, these visualization tools present clear, direct interpretation: to, for example, cement > 300 kg/m³ and age > 28 days, with a minimum water content less than 180 kg/m³, a definitive 10 to 20 MPa boost to predictions is ensured, as indicated by the SHAP values.

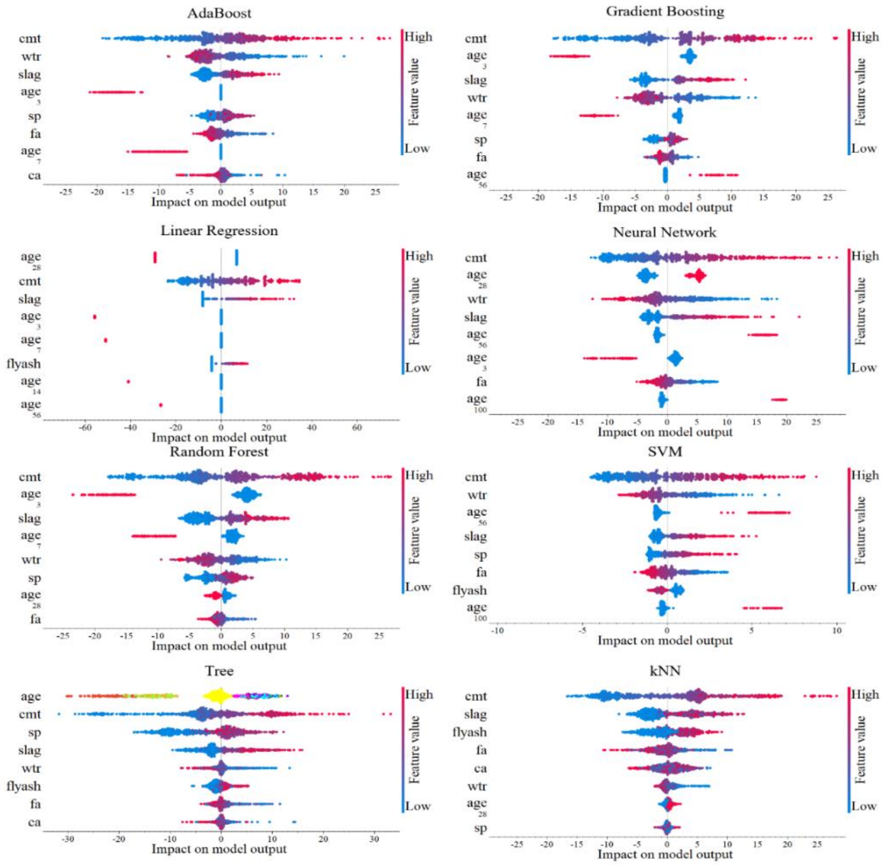

Fig. 8: SHAP summary plots illustrating the impact of feature values on predicted concrete compressive strength for the eight base machine learning models. X, axis: SHAP value (MPa contribution to prediction); Y, axis: Features (ordered by average impact); Color: Feature value (blue = low, red = high).

For ease of use of the proposed meta-learning approach in a real-world application in the field of engineering, a self-contained GUI application has been developed. With the developed application, one does not need to be a programmer to use the application to predict the compressive strength of concrete given the parameters of the concrete mixture. Fig. 9 shows the interface of the developed application.

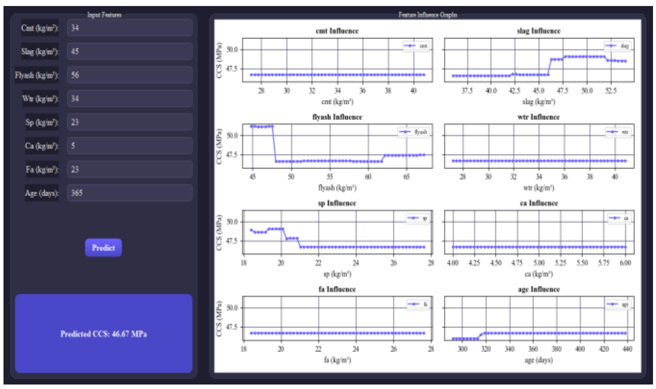

Fig. 9: User interface of the developed predictive application.

3. Discussion

The meta, learning method developed in this work proposes a stacked ensemble approach that systematically combines eight distinct machine learning models into a single model for predicting concrete compressive strength using mix design variables and curing age. The results of the stacked ensemble model clearly demonstrate the prediction of concrete compressive strength on a broad range of values (2.33-82.60 MPa) with excellent R-squared value of 0.890, mean absolute error of approximately 4.8 MPa, and root mean squared error of approximately 6.9 MPa on the held out test data set, which is acceptable from an experimental perspective with variability up to 5 MPa, thereby establishing the model's efficacy for experimental purposes without waiting for 28-day curing times.

Retaining original physical units (kg/m3 for materials, days for age, MPa for strength) without scaling permits direct interpretation of results in real, world terms. Predictions and errors may be applied immediately to mix proportions, rendering the framework useful on-site quality control and initial design decisions.

Cross-model analysis ranks the most important parameters consistently as cement content, curing age and water content. The increased amount of cement, >300 kg/m³, develops higher strength due to increased hydration. A longer age of curing, especially more than 28 days, leads to further gain in strength. A less water content, <180-190 kg/m³, reduces porosity and develops better densities, confirming the crucial role of the water, cement ratio.

Supplementary materials (blast furnace slag, fly ash, superplasticizer): Positive but context-dependent effects; in greater proportions or mix combinations, the role is more significantly influential. Coarse/fine aggregates: Least influential; inert fillers to enhance size/volume rather than influencing the strength component.

This framework has been proven to be quite effective in giving a clear yet optimized path for designing general concrete mixtures by giving importance to higher cement contents, prolonged curing times, lower water to cement ratio, and additional admixture usage on a selective basis for sustainability or performance enhancement purposes.

4. Conclusions & Recommendations

Based on the statistical analyses and machine learning, several key conclusions can be drawn regarding the effectiveness of the proposed meta, learning framework:

I. The developed Compressive Strength Predictor application delivers highly accurate concrete strength predictions through an optimized multi-model ensemble, outperforming individual machine learning approaches.

II. With its intuitive, offline-capable graphical interface and real-time visualization features, the application is fully ready for immediate deployment in engineering practice and education.

III. This tool effectively bridges advanced machine learning research with practical concrete design, offering a reliable and user-friendly alternative to traditional laboratory testing.

Following recommendations are proposed to guide future research and practical applications:

1. Validate the application on larger, real-world datasets from diverse construction projects to further confirm its generalization capability.
2. Conduct field trials with practicing engineers to collect usability feedback and implement interface enhancements.

3. Develop mobile or web versions of the application to enable convenient on-site predictions at construction locations.

Acknowledgement

This research was conducted independently by the author without external funding, supervision, or direct collaboration, with gratitude extended to the open, source community for providing tools like scikit, learn, SHAP, pandas, and Orange Data Mining, and to the UCI Machine Learning Repository for hosting the publicly available Concrete Compressive Strength dataset, which enabled the study's implementation and analysis.

Declaration of Competing Interest

Declarations

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

Ethics approval, Consent to Participate, and Consent to Publish

Ethics, Consent to Participate, and Consent to Publish declarations: not applicable.

Clinical trial registration

Clinical trial registration: not applicable.

Competing interests

The author declares that there are no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability statement The dataset, standalone application, trained models, and materials supporting this study will be publicly available upon acceptance of the paper at: https://github.com/tufailmab/ccs-ensemble-predictor.

Author Contributions Tufail Mabood: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Data curation, Writing - original draft, Writing - review & editing, Visualization.

References

K. Parida, L. Satpathy, A. N. Nayaka, and M. K. Amat, "Use of machine learning models for prediction of compressive strength of concrete produced with waste materials," Innovative Infrastructure Solutions, vol. 10, no. 10, Oct. 2025, doi: 10.1007/S41062-025-02266-6.

M. Nafiuzzaman, T. I. Jakir, I. J. Aditi, A. Kabir, and K. A. Ahsan, "Different machine learning approaches to predict the compressive strength of composite cement concrete," Journal of Building Pathology and Rehabilitation, vol. 10, no. 2, Dec. 2025, doi: 10.1007/S41024-025-00598-5.

M. C. Kang, D. Y. Yoo, and R. Gupta, "Machine learning-based prediction for compressive and flexural strengths of steel fiber-reinforced concrete," Constr Build Mater, vol. 266, Jan. 2021, doi: 10.1016/J.CONBUILDMAT.2020.121117.

D. C. Feng et al., "Machine learning-based compressive strength prediction for concrete: An adaptive boosting approach," Constr Build Mater, vol. 230, Jan. 2020, doi: 10.1016/J.CONBUILDMAT.2019.117000.

R. K. Tipu, V. R. Panchal, and K. S. Pandya, "Multi-objective optimized high-strength concrete mix design using a hybrid machine learning and metaheuristic algorithm," Asian Journal of Civil Engineering, vol. 24, no. 3, pp. 849-867, Apr. 2023, doi: 10.1007/S42107-022-00535-8.

I. Albaijan et al., "Estimating the initial fracture energy of concrete using various machine learning techniques," Eng Fract Mech, vol. 295, Jan. 2024, doi: 10.1016/J.ENGFRACMECH.2023.109776.

A. R. Al-Shamasneh et al., "Application of machine learning techniques to predict the compressive strength of steel fiber reinforced concrete," Scientific Reports 2025 15:1, vol. 15, no. 1, pp. 30674-, Aug. 2025, doi: 10.1038/s41598-025-16516-1.

H. Ayat, Y. Kellouche, M. Ghrici, and B. Boukhatem, "Compressive strength prediction of limestone filler concrete using artificial neural networks," Advances in Computational Design, vol. 3, no. 3, pp. 289-302, Jul. 2018, doi: 10.12989/ACD.2018.3.3.289.

H. Nguyen, T. Vu, T. P. Vo, and H. T. Thai, "Efficient machine learning models for prediction of concrete strengths," Constr Build Mater, vol. 266, Jan. 2021, doi: 10.1016/J.CONBUILDMAT.2020.120950.

M. Mirrashid and H. Naderpour, "Recent Trends in Prediction of Concrete Elements Behavior Using Soft Computing (2010-2020)," Archives of Computational Methods in Engineering, vol. 28, no. 4, pp. 3307-3327, Jun. 2021, doi: 10.1007/S11831-020-09500-7.

N. D. Hoang and D. V. Tran, "Machine Learning-Based Estimation of Concrete Compressive Strength: A Multi-Model and Multi-Dataset Study," Civil Engineering Infrastructures Journal, vol. 57, no. 2, pp. 247-265, Dec. 2024, doi: 10.22059/CEIJ.2023.354679.1910.

A. K. P Das, "Hybrid machine learning approach to prediction of the compressive and flexural strengths of UHPC and parametric analysis with shapley additive explanations," Case Stud. Constr. Mater., vol. 20, p. e02723, 2024.

B. K. A. Mohamad Ali Ridho, C. Ngamkhanong, Y. Wu, and S. Kaewunruen, "Recycled aggregates concrete compressive strength prediction using artificial neural networks (ANNs)," Infrastructures (Basel), vol. 6, no. 2, p. 17, Feb. 2021, doi: 10.3390/infrastructures6020017.

H. Alabduljabbar, "Predicting ultra-high-performance concrete compressive strength using gene expression programming method," Case Stud. Constr. Mater., vol. 18, p. e02074, 2023.

H. Naderpour, A. H. Rafiean, and P. Fakharian, "Compressive strength prediction of environmentally friendly concrete using artificial neural networks," J. Build. Eng., vol. 16, pp. 213-219, Mar. 2018, doi: 10.1016/j.jobe.2018.01.007.

R. Kumar, D. R. Kumar, W. Wipulanusat, C. Thongchom, P. Samui, and B. Rai, "Estimation of the compressive strength of ultrahigh performance concrete using machine learning models," Intell. Syst. Appl., vol. 25, pp. 2667-3053, Mar. 2025, doi: 10.1016/j.iswa.2024.200471.

P. G. Asteris, A. D. Skentou, A. Bardhan, P. Samui, and K. Pilakoutas, "Predicting concrete compressive strength using hybrid ensembling of surrogate machine learning models," Cem Concr Res, vol. 145, Jul. 2021, doi: 10.1016/J.CEMCONRES.2021.106449.

P. G. Asteris and V. G. Mokos, "Concrete compressive strength using artificial neural networks," Neural Comput. Appl., vol. 32, no. 15, pp. 11807-11826, Aug. 2020, doi: 10.1007/s00521-019-04663-2.

M. F. Javed et al., "Applications of gene expression programming and regression techniques for estimating compressive strength of Bagasse ash-based concrete," Crystals (Basel), vol. 10, no. 9, p. 737, Sep. 2020, doi: 10.3390/cryst10090737.

S. S. Chandra, R. Kumar, A. Arjunasamy, S. Galagali, A. Tantri, and S. R. Naganna, "Predicting the compressive strength of polymer-infused bricks: A machine learning approach with SHAP interpretability," Sci. Rep., vol. 15, no. 1, p. 8090, Dec. 2025, doi: 10.1038/s41598-025-89606-9.

R. Biswas, "Efficient soft computing techniques for the prediction of compressive strength of geopolymer concrete," Comput. Concr, vol. 28, pp. 221-232, 2021.

Z. Shen, A. F. Deifalla, P. Kamiński, and A. Dyczko, "Compressive Strength Evaluation of Ultra-High-Strength Concrete by Machine

Learning," Materials, vol. 15, no. 10, May 2022, doi: 10.3390/MA15103523.

Z. M. Yaseen et al., "Advances in engineering software predicting compressive strength of lightweight foamed concrete using extreme learning machine model," Adv. Eng. Softw., vol. 115, pp. 112-125, Jan. 2018, doi: 10.1016/j.advengsoft.2017.09.004.

M. Shaaban, M. Amin, S. Selim, and I. M. Riad, "Machine learning approaches for forecasting compressive strength of high-strength concrete," Scientific Reports 2025 15:1, vol. 15, no. 1, pp. 25567-, Jul. 2025, doi: 10.1038/s41598-025-10342-1.

M. M. Alsaadawi, M. K. Elshaarawy, and A. K. Hamed, "Concrete compressive strength classification using hybrid machine learning models and interactive GUI," Innovative Infrastructure Solutions 2025 10:5, vol. 10, no. 5, pp. 198-, Apr. 2025, doi: 10.1007/S41062-025-01983-2.