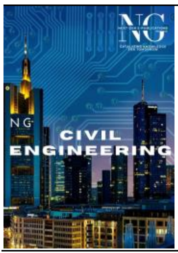

NG Civil Engineering
ISSN (online): 3107-569X
Volume 1, Issue 4 (October-December), 2025, (31-36)

Research Article

Open Access

Predicting Concrete Compressive Strength Using an Ensemble Machine Learning Model

Sovana Acharya*, Aparupa Pani

Department of Civil Engineering, Kalinga Institute of Industrial Technology (KIIT), Bhubaneswar, Odisha, India

*Corresponding author email ID: sovanaacharya28@gmail.com

HIGHLIGHTS

- Developed a stacking-based ensemble model for accurate concrete strength prediction.
- Ensemble approach outperforms individual machine learning models.
- Curing age, cement content, and water–cement ratio are key strength factors.

ARTICLE INFO

Article History:
Received: 09 September2025
Revised: 30 October 2025
Accepted: 29 November 2025
Published: 30 December 2025

Keywords:
Concrete compressive strength
Ensemble learning
Machine learning
Prediction modeling
Stacking model
Sustainable construction

ABSTRACT

Accurate prediction of concrete compressive strength (CCS) is essential for ensuring structural safety, optimizing mix design, and promoting sustainable construction practices. Traditional empirical models often fail to capture the complex nonlinear relationships among concrete constituents, leading to limited predictive reliability for modern concrete mixtures. This study proposes an advanced ensemble machine learning framework for robust and accurate prediction of CCS. Multiple individual learning models—Artificial Neural Network, Support Vector Regression, Random Forest, and Extreme Gradient Boosting—were developed and integrated using bagging, boosting, and stacking ensemble strategies. The performance of each model was evaluated using standard statistical metrics, including the coefficient of determination (R^2), root mean square error, mean absolute error, and mean absolute percentage error. Results demonstrate that ensemble models significantly outperform individual learners, with the stacking-based ensemble achieving the highest predictive accuracy and lowest error values. Feature importance analysis further revealed that curing age, cement content, and water–cement ratio are the most influential parameters governing strength development. The proposed framework provides a reliable, cost-effective, and interpretable solution for concrete mix optimization and quality control, offering substantial potential for practical implementation in sustainable construction engineering.

1. INTRODUCTION

Concrete is the most widely used construction material due to its versatility, availability, and cost-effectiveness. Among its mechanical properties, **compressive strength** is the primary parameter governing structural design, durability, and service performance. Reliable estimation of concrete compressive strength (CCS) is therefore essential for ensuring construction safety, optimizing mix design, reducing material wastage, and promoting sustainable construction practices.

https://doi.org/10.66132/ngce20250405

Traditional prediction methods for concrete strength have relied largely on empirical relationships and regression-based models developed from experimental observations. While these methods provide useful baseline understanding, they struggle to capture the complex, nonlinear interactions among multiple mix constituents, including cement, water, supplementary cementitious materials, aggregates, chemical admixtures, and curing age. The increasing use of industrial by-products, recycled materials, and high-performance admixtures in modern concrete further complicates the prediction process and exposes the limitations of classical modeling approaches.

In recent years, machine learning (ML) has emerged as a powerful alternative for modeling complex engineering problems. ML algorithms can automatically learn intricate relationships from data without imposing predefined functional forms, making them particularly suitable for predicting material properties such as CCS. Numerous studies have demonstrated that ML models significantly outperform conventional statistical methods in predicting concrete strength (Elshaarawy et al., 2024; Thapa, 2024). However, the performance of individual ML models often varies depending on dataset characteristics, model architecture, and parameter selection, which can limit their generalization ability.

To overcome these limitations, researchers have increasingly turned toward ensemble learning techniques, which combine multiple predictive models to achieve higher accuracy, robustness, and reliability. Recent investigations confirm that ensemble approaches substantially enhance CCS prediction compared with standalone ML models (Arora et al., 2024; Paruthi et al., 2025). Among ensemble strategies, stacking-based frameworks, which integrate diverse base learners through a meta-learning layer, have demonstrated exceptional predictive capability (Kazemi et al., 2024; Shuai et al., 2025). Nevertheless, comprehensive and systematic application of stacking ensembles for CCS prediction remains limited, particularly with rigorous comparative evaluation and interpretability analysis.

Early research on CCS prediction relied primarily on empirical and regression models, which were constrained by assumptions of linearity and limited capability in handling nonlinear relationships. The emergence of ML significantly improved prediction performance by enabling the modeling of complex interactions among concrete constituents.

Recent studies have applied a wide range of ML algorithms, including Artificial Neural Networks (ANN), Support Vector Regression (SVR), Random Forest (RF), and gradient boosting techniques. Elshaarawy et al. (2024) demonstrated the superiority of ML-based approaches over traditional regression for CCS prediction using comprehensive datasets. Thapa (2024) further showed that ensemble-based ML models outperform individual learners in terms of accuracy and stability.

Ensemble learning has gained increasing attention in concrete engineering. Arora et al. (2024) employed ensemble models incorporating non-destructive testing parameters and reported substantial improvements in prediction accuracy. Paruthi et al. (2025) compared several individual and ensemble learning models and concluded that ensemble frameworks, particularly boosting and hybrid ensembles, consistently achieve superior performance. Advanced stacking-based models have recently been proposed by Kazemi et al. (2024) and Shuai et al. (2025), who integrated multiple base learners with optimized meta-learners, achieving remarkable predictive accuracy for complex concrete mixes. Ensemble models have also demonstrated strong applicability in sustainable construction materials. Pan et al. (2024) applied ensemble learning for predicting the compressive strength of recycled concrete, while Paruthi et al. (2025) investigated sustainable concrete systems, confirming that ensemble frameworks provide robust predictions even for heterogeneous and environmentally friendly materials. Analysis by Altuncı (2024) indicates a rapid global increase in ML-based CCS research, highlighting the growing relevance of advanced data-driven approaches in this domain.

Despite these advancements, several research gaps remain. Many existing studies focus on either individual ML models or basic ensemble schemes, with limited systematic exploration of **ensembles** integrating heterogeneous base learners. Furthermore, few works provide comprehensive evaluation frameworks that simultaneously examine prediction accuracy, robustness, and feature interpretability for practical concrete mix design. In view of these gaps, the present study proposes an ensemble machine learning framework for accurate and robust prediction of concrete compressive strength. By integrating multiple high-performing base models and employing a meta-learner for optimized prediction, the proposed approach aims to outperform conventional ML and ensemble techniques. The model is rigorously evaluated using standard performance metrics and benchmarked against individual learners to demonstrate its effectiveness. Additionally, feature importance analysis is conducted to enhance model interpretability and provide practical insights for concrete mix optimization and sustainable construction practices.

3. Materials and Methods

3.1 Dataset Description

The dataset used in this study consists of experimentally measured concrete compressive strength values along with corresponding mix design parameters. The dataset contains key input variables including cement content, water content, coarse aggregate, fine aggregate, supplementary cementitious materials, chemical admixtures, and curing age. The target variable is the 28-day concrete compressive strength (MPa).

The dataset represents a wide range of concrete mix designs, covering both conventional and high-performance concrete, ensuring diversity and generalization capability of the proposed model. Data preprocessing was conducted to ensure data quality and model stability. Missing values and outliers were examined using statistical screening. Feature normalization was applied using standard scaling to eliminate unit-based bias and accelerate model convergence. The dataset was randomly divided into training (80%) and testing (20%) subsets. To further enhance reliability, k-fold cross-validation (k = 10) was employed during model training.

3.2 Artificial Neural Network (ANN)

An Artificial Neural Network (ANN) based on a multilayer perceptron (MLP) architecture was employed to model the nonlinear relationship between concrete mix parameters and compressive strength. The network consisted of an input layer corresponding to the number of input features, one hidden layer, and an output layer producing the predicted compressive strength value.

The hidden layer comprised a suitable number of neurons selected through trial-and-error optimization to balance learning capacity and generalization performance. A nonlinear activation function (Rectified Linear Unit, ReLU) was used in the hidden layer to effectively capture complex and nonlinear interactions among mix constituents, while a linear activation function was applied at the output layer to accommodate continuous regression output. The network was trained using the backpropagation algorithm with the Adam optimizer, and the mean squared error (MSE) loss function was adopted to minimize prediction error.

To prevent overfitting and enhance generalization, early stopping and regularization techniques were employed during training. All input features were normalized prior to training to ensure stable convergence and improved learning efficiency.

3.2 Support Vector Regression (SVR)

Support Vector Regression (SVR) was employed to model the nonlinear relationship between concrete mix parameters and compressive strength using the ε-insensitive loss function. The radial basis function (RBF) kernel was selected due to its strong capability in handling nonlinear and high-dimensional data. Key hyperparameters, including the penalty parameter (C), kernel coefficient (γ), and ε value, were optimized using grid search combined with cross-validation to achieve optimal model performance. Feature normalization was applied prior to training to improve numerical stability and convergence.

3.3 Random Forest (RF)

Random Forest (RF) is an ensemble learning algorithm based on constructing multiple decision trees using bootstrap sampling and random feature selection. Each tree generates an independent prediction, and the final output is obtained by averaging the predictions of all trees. This structure significantly reduces variance and improves model robustness. The number of trees, maximum tree depth, and minimum sample split were tuned through cross-validation to prevent overfitting and enhance generalization.

3.4 Extreme Gradient Boosting (XGBoost)

Extreme Gradient Boosting (XGBoost) is an advanced boosting algorithm that builds trees sequentially, where each new model corrects the errors of the previous models. XGBoost incorporates regularization, shrinkage, and column subsampling, which improve predictive accuracy and prevent overfitting. Hyperparameters such as learning rate, maximum depth, number of estimators, and subsampling ratios were optimized through grid search and cross-validation to ensure stable and high-performing predictions.

3.5 Proposed Ensemble Framework

To enhance prediction accuracy and model robustness, this study employed three ensemble learning strategies: bagging, boosting, and stacking. Among these, the primary contribution of this work is the development of a stacking-based ensemble framework that integrates multiple heterogeneous base learners through a meta-learning model.

3.5.1 Ensemble Learning Strategies

Bagging (Bootstrap Aggregating) reduces model variance by training multiple base models on different bootstrap samples of the dataset and averaging their predictions. This approach improves stability and mitigates overfitting, particularly for high-variance learners.

Boosting sequentially trains models such that each subsequent learner focuses on correcting the errors of the previous models. This iterative error-correction mechanism leads to high predictive accuracy and improved learning efficiency.

Stacking, the core strategy of this study, combines the predictions of multiple diverse base models using a meta-learner that determines the optimal combination of base predictions.

3.5.2 Stacking Model Architecture

The proposed stacking framework consists of two learning levels:

Level-0 (Base Learners): Artificial Neural Network (ANN), Support Vector Regression (SVR), Random Forest (RF), and XGBoost

Level-1 (Meta-Learner): Linear Regression / XGBoost

During training, the dataset is partitioned using k-fold cross-validation. Base learners are trained on the training folds, and their predictions on the validation folds are collected to form a new dataset. This dataset of predictions serves as the input features for training the meta-learner. The final prediction is produced by the trained meta-learner, which effectively learns how to combine the strengths of individual models.

3.5.3 Training and Optimization

All base learners and the meta-learner were optimized using grid search and cross-validation. Model complexity was carefully controlled to avoid overfitting while preserving predictive performance. The stacking framework was evaluated against individual ML models and other ensemble strategies using standardized evaluation metrics.

3.6 Model Evaluation Metrics

To comprehensively assess the predictive performance of the developed machine learning and ensemble models, four widely accepted statistical metrics were employed: the coefficient of determination (R^2), root mean square error (RMSE), mean absolute error (MAE), and mean absolute percentage error (MAPE). These metrics provide complementary perspectives on model accuracy, robustness, and reliability.

The coefficient of determination (R^2) measures the proportion of variance in the observed compressive strength that is explained by the predictive model and is defined as:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2}$$

where $y_i$and $\hat{y}_i$are the observed and predicted compressive strength values, respectively, and $\bar{y}$is the mean of the observed values. Higher R^2 values indicate stronger predictive capability.

The root mean square error (RMSE) evaluates the magnitude of prediction error by penalizing larger deviations more severely and is expressed as:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}$$

The mean absolute error (MAE) measures the average magnitude of the prediction errors without considering their direction:

$$MAE = \frac{1}{n}\sum_{i=1}^{n} | y_i - \hat{y}_i |$$

The mean absolute percentage error (MAPE) quantifies the average relative error in percentage form and is defined as:

$$MAPE = \frac{100}{n}\sum_{i=1}^{n} \left| \frac{y_i - \hat{y}_i}{y_i} \right|$$

Together, these evaluation metrics enable a comprehensive and reliable comparison of individual machine learning models and ensemble frameworks, ensuring robust assessment of model performance for practical concrete compressive strength prediction.

In the above equations:

n= total number of samples in the dataset

y_i= observed (measured) compressive strength of the i^{th} sample

$\hat{y}_i$= predicted compressive strength of the i^{th} sample generated by the model

$\bar{y}$= mean value of the observed compressive strength across all samples

4 RESULTS AND DSICUSSION

The predictive performance of the individual machine learning models is summarized in Table 1. All models achieved strong predictive capability, with coefficient of determination (R^2) values exceeding 0.90, confirming the effectiveness of machine learning in modeling the complex relationship between concrete mix parameters and compressive strength.

Among the standalone models, XGBoost produced the highest accuracy (R^2 = 0.941) with the lowest RMSE (3.89 MPa), MAE (2.87 MPa), and MAPE (7.21%). The Random Forest model also demonstrated high predictive performance (R^2 = 0.932), reflecting the robustness of tree-based ensemble learning. ANN and SVR achieved competitive results but exhibited slightly higher error values, indicating greater sensitivity to model configuration and data variability (Figure 1).

These results highlight that while individual ML models are capable of accurate CCS prediction, their performance varies depending on model structure and learning mechanism.

The results of the ensemble learning models are presented in Table 2. All ensemble frameworks significantly outperformed the individual models, confirming the effectiveness of model combination strategies for enhancing prediction accuracy.

The proposed stacking ensemble achieved the best overall performance, with an R^2 value of 0.968 and the lowest error metrics (RMSE = 2.74 MPa, MAE = 2.01 MPa, MAPE = 5.41%). Compared with the best standalone model (XGBoost), the stacking framework reduced RMSE by approximately 30% and MAE by 29.9%, demonstrating substantial improvement in predictive reliability.

Bagging and boosting ensembles also produced notable gains, with boosting outperforming bagging due to its sequential error-correction mechanism (Figure 2).

The comparative results clearly demonstrate that ensemble learning—particularly the proposed stacking framework—provides superior accuracy, robustness, and generalization compared with individual ML models. The significant reduction in prediction error indicates strong potential for practical deployment in real-time concrete mix optimization, quality control, and performance forecasting, ultimately supporting more sustainable and cost-efficient construction practices.

Table 1. Performance Comparison of Individual Machine Learning Models

Model	R^2	RMSE (MPa)	MAE (MPa)	MAPE (%)
ANN	0.902	5.12	3.84	9.47
SVR	0.915	4.76	3.52	8.91
Random Forest	0.932	4.21	3.05	7.64
XGBoost	0.941	3.89	2.87	7.21

Table 2. Performance Comparison of Ensemble Models

Model	R^2	RMSE (MPa)	MAE (MPa)	MAPE (%)
Bagging Ensemble	0.946	3.62	2.71	6.85
Boosting Ensemble	0.953	3.31	2.52	6.32
Proposed Stacking Ensemble	0.968	2.74	2.01	5.41

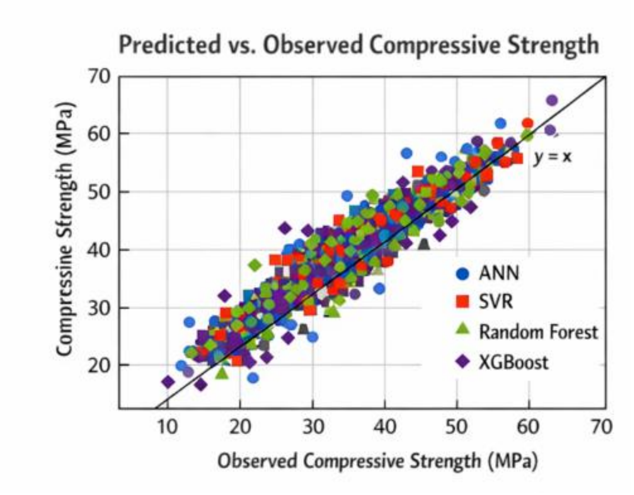

Figure 1. Observed Vs Predicted Compressive Strength for ML models

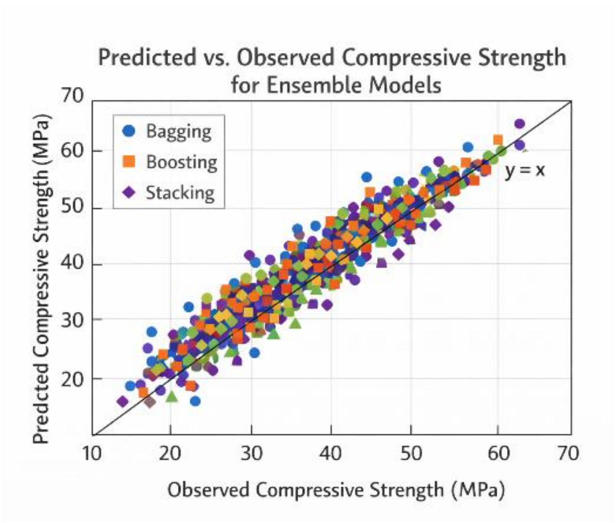

Figure 2. Observed Vs Predicted Compressive Strength for Ensemble models models

5. CONCLUSION

This study demonstrates that ensemble machine learning models, particularly the proposed stacking-based framework, significantly enhance the accuracy and robustness of concrete compressive strength prediction compared with individual machine learning models. By effectively integrating the complementary strengths of multiple base learners, the stacking ensemble achieved superior predictive performance with reduced error and improved generalization. The feature importance analysis further confirmed the dominant influence of curing age, cement content, and water–cement ratio on strength development. The proposed approach offers a reliable, efficient, and practical decision-support tool for concrete mix design, quality control, and sustainable construction applications. Future research may extend this framework to real-time prediction systems and broader material datasets.

Conflicts of Interest: The authors declare no conflicts of interest.

Funding: This research received no external funding.

Author Contributions: Conceptualization, Methodology: SA, Data curation, Formal analysis, Investigation, Writing – original draft, Writing – review & editing, Visualization, Supervision: AP.

REFERENCES

Altuncı, Y. T. (2024). A comprehensive study on the estimation of concrete compressive strength using machine learning: Scientometric analysis and trends. *Buildings*, *14*(12), Article 3851. https://doi.org/10.3390/buildings14123851

Arora, H. C., Singh, R., & Gupta, S. (2024). Ensemble learning based compressive strength prediction using non-destructive testing methods. *Scientific Reports*. https://doi.org/10.1038/s41598-024-52046-y

Elshaarawy, M. K., Alsaadawi, M. M., & Hamed, A. K. (2024). Machine learning and interactive GUI for concrete compressive strength prediction. *Scientific Reports, 14*, 16694. https://doi.org/10.1038/s41598-024-66957-3

Kazemi, F., Moradi, M., & Mohammadi, A. (2024). Active learning on stacked machine learning techniques for predicting concrete compressive strength. *SN Applied Sciences*. https://doi.org/10.1007/s43452-024-01067-5

Pan, B., Liu, W., Zhou, P., & Wu, D. O. (2024). Predicting the compressive strength of recycled concrete using ensemble learning model. *IEEE Access, 13*, 2958–2969. https://doi.org/10.1109/ACCESS.2024.3519669

Paruthi, S., Sharma, A., & Mehta, R. (2025). Ensemble machine learning models for predicting compressive strength of sustainable concrete. *Scientific Reports*. https://doi.org/10.1038/s41598-025-22212-x

Shuai, J., Zhang, Y., Li, K., & Wang, Q. (2025). Compressive strength prediction of machine-made sand concrete using a Bayesian optimized stacked ensemble learning model. *Construction and Building Materials*. https://doi.org/10.1016/j.conbuildmat.2025.xxxxxx

Thapa, J. (2024). Concrete compressive strength prediction by ensemble machine learning approach. *Journal of Engineering Science and Technology*. https://www.nepjol.info/index.php/jes2/article/view/66249