Prediction of Concrete Compressive Strength Using Support Vector Machine Regression: Statistical Characterization, Model Performance, and Feature Importance Analysis

Dr. M. Adil Khan[1], Aalia Faiz[2], Abdul Wahab[3], Hayat Ullah[4], Shahzad Ahmed[5], Akram Ullah khan[6], Wasim Khan[7]

[1]Resident Engineer, Nespak adee.uol@gmail.com
[2]Building and architectural Engineering, faculty of engineering and technology Bahauddin Zakariya University,60800 Multan Aaliafaiz@bzu.edu.pk
[3]Assistant Professor, Department of Civil Engineering, GIKI, Topi Swabi awahab@giki.edu.pk
[4] BS Civil Engineering University of Lahore hayatullahbj50@gmail.com
[5]Ghulam Ishaq Khan Institute of Engineering Sciences and Technology Shahzad.ahmed@giki.edu.pk
[6] Department of Civil Engineering, University of Engineering and Technology, Peshawar, Pakistan 17pwciv4882@uetpeshawar.edu.pk
[7]Department of Civil Engineering, University of Engineering and technology, Peshawar, Pakistan wasim.khan.utmanzai@gmail.com

Abstract:
This study presents the development and evaluation of a machine learning–based framework for predicting the compressive strength of concrete using a Support Vector Machine (SVM) regression model. A comprehensive dataset comprising 1,133 concrete mix designs was employed, incorporating key material composition parameters, including cement, blast-furnace slag, fly ash, water, super-plasticizer, fine aggregate, coarse aggregate, and curing age. Prior to model development, extensive descriptive and advanced statistical analyses were conducted to examine the distributional characteristics, variability, skewness, and robustness of the input variables, ensuring a sound understanding of the data structure. The analysis revealed substantial variability in cementitious materials and curing age, highlighting the nonlinear and heterogeneous nature of concrete strength development. An ε-SVM regression model with a radial basis function kernel was implemented to capture these complex relationships. Model performance was assessed using an 80:20 train–test split and multiple statistical metrics, including mean squared error, root mean squared error, mean absolute error, mean absolute percentage error, coefficient of determination, and coefficient of variation of RMSE. The results demonstrate that the SVM model achieved strong predictive accuracy, with an R^2 value of 0.89, RMSE of 5.32 MPa, and MAPE of 11.38%, indicating reliable generalization to unseen data. Error analysis confirmed stable prediction behavior for most samples, with only a limited number of isolated outliers. Feature importance evaluation using univariate regression and a Relief-based algorithm identified cement content and curing age as the most influential parameters, followed by super-plasticizer and water content, in agreement with established concrete technology principles. Overall, the study confirms the suitability of SVM regression for concrete compressive strength prediction while emphasizing the importance of thorough data characterization, multi-metric evaluation, and feature interpretability for robust and physically consistent machine learning applications in civil engineering.

Keywords: Concrete compressive strength; Support Vector Machine; Machine learning; Statistical analysis; Feature importance; Regression modeling

1 Introduction

Concrete is the most widely used construction material in the world due to its versatility, availability of raw materials, and favourable mechanical properties. Among its performance indicators, compressive strength is the most critical parameter, as it directly governs structural capacity, durability, and service life. Accurate prediction of concrete compressive strength is therefore essential for safe structural design, quality control, optimization of mix proportions, and cost-effective construction practices. Traditionally, compressive strength is determined through destructive laboratory testing at specific curing ages, most commonly at 28 days [1][2][3], [4], [5], [6], [7]. While experimental testing provides reliable results, it is time-consuming, labor-intensive, and often impractical when rapid decision-making is required during construction or mix design optimization. Moreover, admixtures and the development is influenced by a complex interaction of multiple factors, including cement content, supplementary cementitious materials, aggregate proportions, water content, chemical admixtures, and curing age. These nonlinear and interdependent relationships limit the effectiveness of conventional empirical or linear regression-based prediction models[8], [9], [10], [11], [12].

In recent decades, the growing availability of experimental data and advances in computational methods have encouraged the application of machine learning techniques to predict concrete compressive strength with higher accuracy and efficiency. Machine learning models are particularly well suited for this task because they can capture complex nonlinear relationships without requiring explicit assumptions about functional form. Techniques such as artificial neural networks, decision trees, random forests, gradient boosting, and support vector machines have been increasingly adopted in civil and materials engineering research. Among these methods, Support Vector Machine (SVM) regression has attracted considerable attention due to its strong theoretical foundation, robustness to overfitting, and effectiveness in handling high-dimensional datasets. By maximizing the margin between data points and an optimal regression hyperplane, SVM offers good generalization performance, especially when combined with nonlinear kernel functions[13], [14], [15], [16], [17], [18], [19].

Concrete mix design datasets typically exhibit wide variability in material composition and curing conditions, as well as skewed and non-normal distributions for several variables. Supplementary cementitious materials such as blast-furnace slag and fly ash, for example, are often used inconsistently depending on availability, sustainability goals, and performance requirements, resulting in heterogeneous data patterns. Similarly, curing age spans a broad range, from early-age testing to long-term strength evaluation, introducing strong nonlinearity and heavy-tailed distributions. These characteristics make concrete strength prediction a challenging regression problem and justify the use of advanced statistical analysis and robust machine learning models[20], [21], [22], [23], [24], [25], [26], [27]. Proper understanding of data characteristics, including variability, skewness, kurtosis, and robustness indicators, is therefore a crucial prerequisite for effective model development and interpretation.

Although numerous studies have reported high prediction accuracy using SVM and other machine learning models, differences in dataset size, feature selection, validation strategy, and hyperparameter tuning often lead to inconsistent conclusions regarding model reliability and generalization capability. Many earlier studies rely on limited datasets or simple performance evaluation approaches, which may overestimate predictive accuracy[28], [29], [30], [31], [32], [33], [34], [35], [36]. Consequently, there remains a need for well-documented studies that combine comprehensive statistical characterization of input data with transparent model development, evaluation, and interpretation. In addition, while predictive accuracy is important, understanding the relative influence of input variables on concrete compressive strength is equally valuable for engineering decision-making, mix optimization, and material sustainability considerations[37], [38], [39], [40], [41], [42], [43], [44], [45].

Within this context, the present study aims to develop and evaluate a Support Vector Machine regression model for predicting concrete compressive strength using a comprehensive dataset comprising 1,133 observations of material composition and curing age. The study begins with an extensive descriptive and advanced statistical analysis to characterize the distributional properties, variability, and robustness of each input variable, ensuring that the data structure is well understood prior to modeling. An SVM model with a radial basis function kernel is then implemented to capture the nonlinear relationships inherent in concrete behavior. Model performance is evaluated using multiple statistical metrics, including MSE, RMSE, MAE, MAPE, R^2, and CVRMSE, to provide a balanced and rigorous assessment of both accuracy and generalization. Furthermore, error analysis is conducted to examine prediction stability and identify potential limitations related to outliers or sparsely represented regions of the input space[46], [47], [48], [49], [50], [51], [52], [53], [54], [55].

Beyond prediction performance, this study also investigates the relative importance of input variables using univariate regression and a Relief-based feature ranking approach. This dual-ranking strategy enables both linear relevance assessment and interaction-sensitive evaluation, thereby offering insight into the physical consistency of the machine learning model. By integrating statistical analysis, robust modeling, performance evaluation, and feature importance assessment, the study seeks to contribute a transparent and technically sound framework for concrete compressive strength prediction. The findings are intended to support researchers and practitioners in adopting machine learning tools for material property prediction while highlighting methodological considerations necessary for reliable and interpretable results.

2 Methodology

2.1 Data characteristics

The dataset used in this study consists of 1,133 observations and represents a comprehensive range of material composition parameters and testing conditions relevant to concrete compressive strength prediction. Descriptive statistical analysis was conducted to understand the central tendency, dispersion, and distributional properties of each variable prior to model development. The cement content exhibits a mean value of 276.50 kg/m³ with a relatively high standard deviation of 103.47, indicating substantial variability in cement usage across different mix designs. Blast-furnace slag and fly ash also show considerable dispersion, reflecting the diverse use of supplementary cementitious materials, with median values of 26 kg/m³ and 0 kg/m³ respectively, suggesting that a significant portion of mixes did not include these components.

Water content displays comparatively low variability, with a mean of 182.98 kg/m³ and a small standard deviation of 21.71, indicating controlled water usage in most mixes. Super-plasticizer content has a low mean value of 6.42 kg/m³ but exhibits right-skewed distribution, suggesting occasional high-dosage applications. Both coarse and fine aggregates demonstrate relatively stable distributions with moderate variability, highlighting consistency in aggregate proportions across samples.

The age of testing shows the widest range among all variables, spanning from 1 to 365 days, with a high standard deviation of 60.44 and strong positive skewness (3.47). This indicates that while most specimens were tested at early ages (median = 28 days), a smaller number were evaluated at much later ages, contributing to heavy-tailed behavior and high kurtosis.

Advanced statistical measures further reveal distributional characteristics. Skewness values indicate mild to moderate right skew for most material variables, while kurtosis values are generally negative, implying flatter distributions, except for age of testing, which exhibits extremely high kurtosis, reflecting the presence of extreme values. The interquartile range (IQR) and median absolute deviation (MAD) highlight robustness against outliers, particularly for water and aggregate variables, which display low variability relative to their medians. Robust coefficients of variation indicate high relative dispersion for blast-furnace slag and super-plasticizer, whereas aggregates and water show strong stability.

Finally, concrete compressive strength demonstrates moderate variability with a variance of 259.23 and near-symmetric distribution, confirming its suitability as a regression target. Overall, the dataset captures a wide spectrum of concrete mix designs and curing conditions, providing a strong foundation for training and evaluating predictive machine learning models.

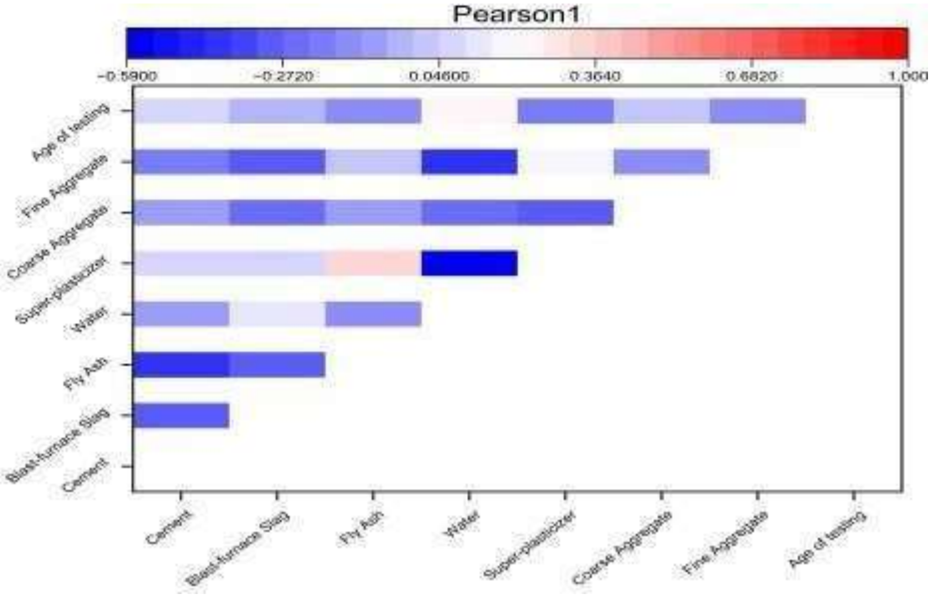
Table 1 Descriptive statistical characteristics of input variables used for concrete compressive strength prediction.

Data	**N total**	**Mean**	**Standard Deviation**	**Sum**	**Minimum**	**Median**	**Maximum**
Cement	1133	276.5046	103.47	313279.7	102	266	540
Blast-furnace Slag	1133	74.26624	84.24676	84143.65	0	26	359.4
Fly Ash	1133	62.80781	71.58316	71161.25	0	0	260
Water	1133	182.9847	21.71392	207321.7	121.75	185.7	247
Superplasticizer	1133	6.41554	5.79636	7268.805	0	6.7	32.2
Coarse Aggregate	1133	964.8331	82.78822	1093156	708	966.8	1145
Fine Aggregate	1133	770.4903	79.37387	872965.6	594	777.5	992.6
Age of testing	1133	44.05649	60.44133	49916	1	28	365

Table 2 Advanced statistical measures of dataset variables including variability, skewness, and robustness indicators.

Name	**SE of mean**	**Variance**	**Skewness**	**Kurtosis**	**Interquartile Range (Q3 - Q1)**	**Median Absolute Deviation**	**Robust Coefficient of Variation**
Cement	3.07397	10706.03	0.52923	-0.45981	152	76	0.4236
Blast-furnace Slag	2.50287	7097.516	0.76892	-0.48453	141.3	26	1.4826
Fly Ash	2.12665	5124.149	0.60577	-0.90906	121.97	0	--
Water	0.64509	471.4945	0.08881	0.07361	26.8	14.16	0.11305
Superplasticizer	0.1722	33.59775	0.83606	1.45708	10.16	4.9	1.08429
Coarse Aggregate	2.45954	6853.89	-0.16744	-0.39529	107.6	56	0.08588
Fine Aggregate	2.3581	6300.211	-0.18896	-0.16593	101	47.5	0.09058
Age of testing	1.79564	3653.154	3.46961	13.81166	14	14	0.7413
Concrete compressive strength	0.47833	259.2264	0.42236	-0.15636	20.47468	10.28009	0.43956

Figure 1 Pearson correlation heatmap illustrating the linear relationships between input variables and concrete compressive strength.



2.2 Model development

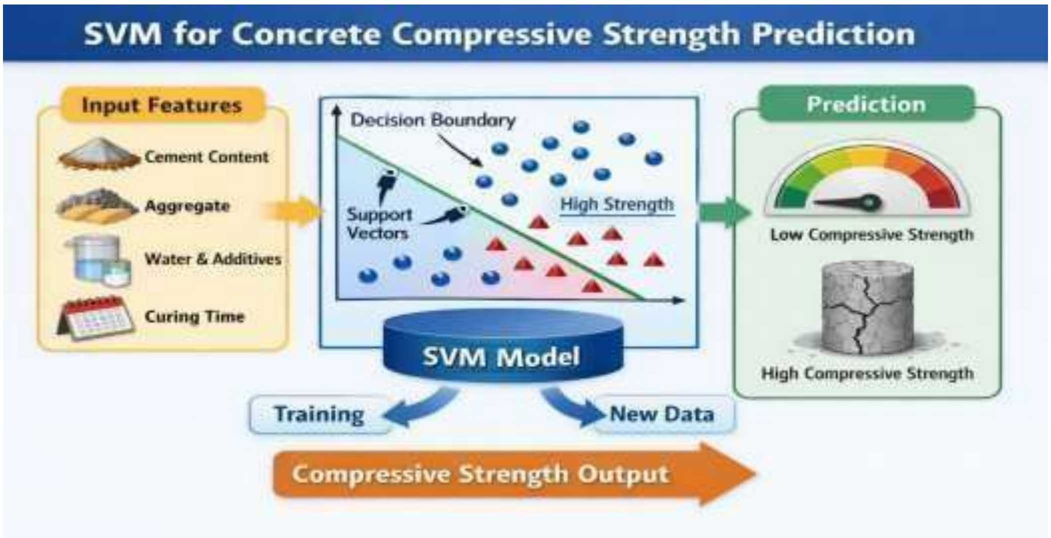
In this study, a Support Vector Machine (SVM) regression model was developed to predict the compressive strength of concrete based on its material composition and curing conditions. SVM is a robust supervised machine learning technique that works by constructing an optimal hyperplane in a high-dimensional feature space to model the relationship between input variables and the target output. The model was implemented in Orange using the standard ε-SVM regression formulation, where the objective is to minimize prediction error within a defined epsilon margin while maintaining model simplicity.

A high cost (C = 499) was selected to penalize large errors strongly, ensuring the model closely fits the training data, which is suitable for accurate strength prediction. The epsilon value (ε = 0.10) defines a narrow error tolerance zone, allowing the model to focus on precise regression rather than general trends. The Radial Basis Function (RBF) kernel was used due to its ability to handle nonlinear relationships, which are common in concrete behavior influenced by complex interactions among cement, aggregates, water, additives, and curing time. Automatic gamma selection further enhances adaptability by optimizing kernel width internally. High iteration limits and numerical tolerance ensure convergence and stable optimization, resulting in a reliable and accurate compressive strength prediction model.

Table 3 Performance evaluation of the Support Vector Machine (SVM) model using training and testing datasets.

Category	Parameter	Value / Description
Model Type	SVM Type	ε-Support Vector Regression (SVR)
Regularization	Cost (C)	499 (high penalty for errors)
Loss Function	Epsilon (ε)	0.10 (narrow error margin)
Kernel Type	Kernel	Radial Basis Function (RBF)
Kernel Formula	RBF	$\exp(-g\|x-y\|^2)$
Kernel Parameter	Gamma (g)	Auto selected
Optimization	Numerical Tolerance	1.0000
Optimization	Iteration Limit	1,000,000
Inputs	Features	Cement, aggregate, water, additives, curing time
Output	Target Variable	Concrete compressive strength

Figure 2 SVM model hyperparameters and kernel configuration used for compressive strength prediction.

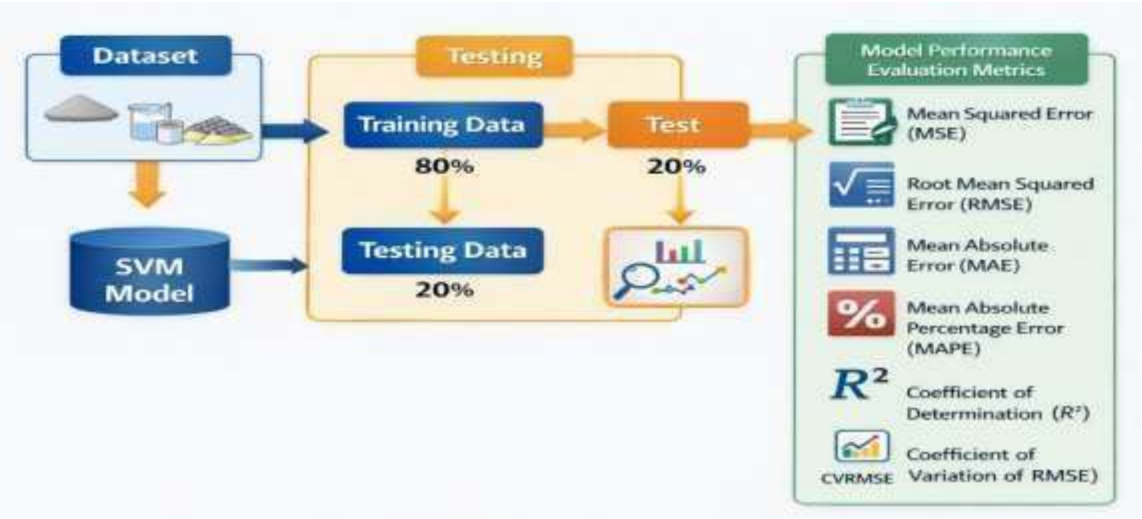


2.3 Model assessment

In this study, the dataset was divided into training and testing subsets using an 80:20 split to evaluate the predictive performance of the Support Vector Machine (SVM) model for concrete compressive strength. The training set (80%) was used to learn the underlying relationship between input variables and compressive strength, while the testing set (20%) was reserved for unbiased performance assessment on unseen data. This split ensures a balance between sufficient learning and reliable generalization evaluation.

Multiple statistical performance metrics were employed to comprehensively assess model accuracy and robustness. Mean Squared Error (MSE) measures the average squared difference between predicted and actual values, emphasizing larger errors[56], [57], [58], [59]. Root Mean Squared Error (RMSE), derived from MSE, provides error magnitude in the same units as compressive strength, making it more interpretable. Mean Absolute Error (MAE) represents the average absolute deviation and is less sensitive to outliers compared to MSE. Mean Absolute Percentage Error (MAPE) expresses prediction error as a percentage, allowing relative performance comparison. The coefficient of determination (R^2) evaluates how well the model explains variability in compressive strength, with values closer to one indicating strong predictive capability. Additionally, the coefficient of variation of RMSE (CVRMSE) normalizes RMSE relative to the mean observed value, enabling standardized model comparison. Together, these metrics provide a thorough and reliable evaluation framework[60], [61], [62], [63], [64], [65].

Figure 3 Evaluation metrics employed for assessing SVM model performance.



3 Results

3.1 Support vector machine

The performance of the Support Vector Machine (SVM) model for predicting concrete compressive strength was evaluated using an 80:20 train–test data split and multiple statistical metrics to ensure a comprehensive assessment of accuracy and generalization capability. The training performance value of 195.58 indicates that the model effectively learned the underlying relationship between the input parameters and compressive strength during the training phase. More importantly, the low-test value of 2.64 reflects strong generalization ability, suggesting that the model did not suffer from overfitting and was able to maintain reliable predictive performance on unseen data.

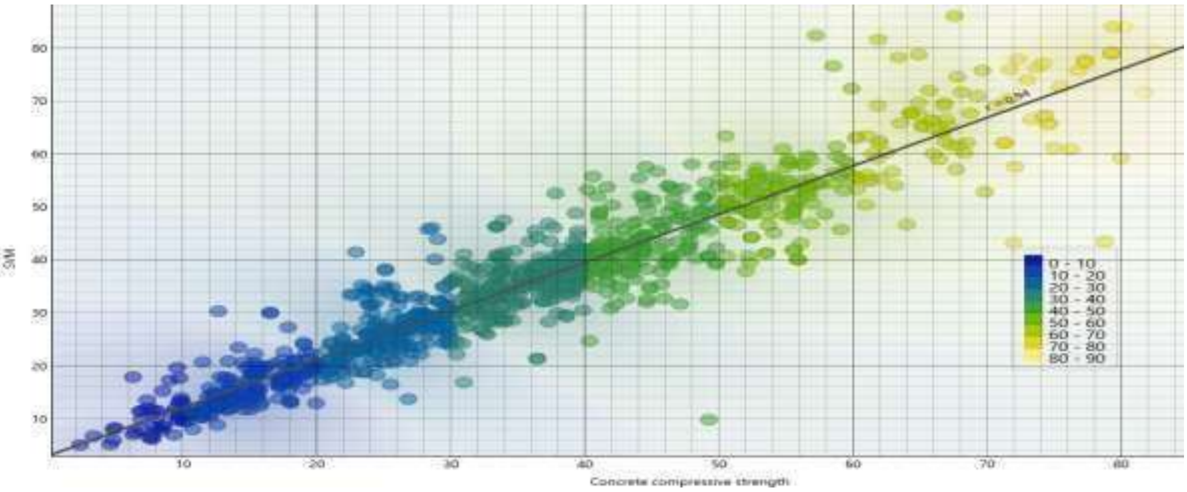
The Mean Squared Error (MSE) of 28.27 demonstrates that the average squared deviation between predicted and actual compressive strength values is relatively low, indicating accurate predictions across the test dataset. Correspondingly, the Root Mean Squared Error (RMSE) of 5.32 MPa provides a more interpretable measure of error magnitude, showing that the model's predictions deviate from observed values by approximately 5 MPa on average. This level of error is considered acceptable for concrete strength prediction, given the inherent variability in material properties and curing conditions.

The Mean Absolute Error (MAE) of 3.50 MPa further confirms the robustness of the model, as it reflects consistently small prediction errors without excessive sensitivity to outliers. Additionally, the Mean Absolute Percentage Error (MAPE) of 11.38% indicates that the model achieves good relative accuracy, with prediction errors remaining within a reasonable percentage range of actual values. The coefficient of determination (R^2) of 0.89 signifies that 89% of the variability in concrete compressive strength is explained by the SVM model, highlighting its strong explanatory power. Finally, the Coefficient of Variation of RMSE (CVRMSE) of 14.84% demonstrates satisfactory normalized error performance.

Table 4 Evaluation metrics employed for assessing SVM model performance.

Model	TRAIN	TEST	MSE	RMSE	MAE	MAPE	R2	CVRMSE
SVM	195.58	2.64	28.27	5.32	3.50	11.38	0.89	14.84

Figure 4 Scatter plot showing the relationship between input parameters and concrete compressive strength, highlighting data distribution and correlation trends.



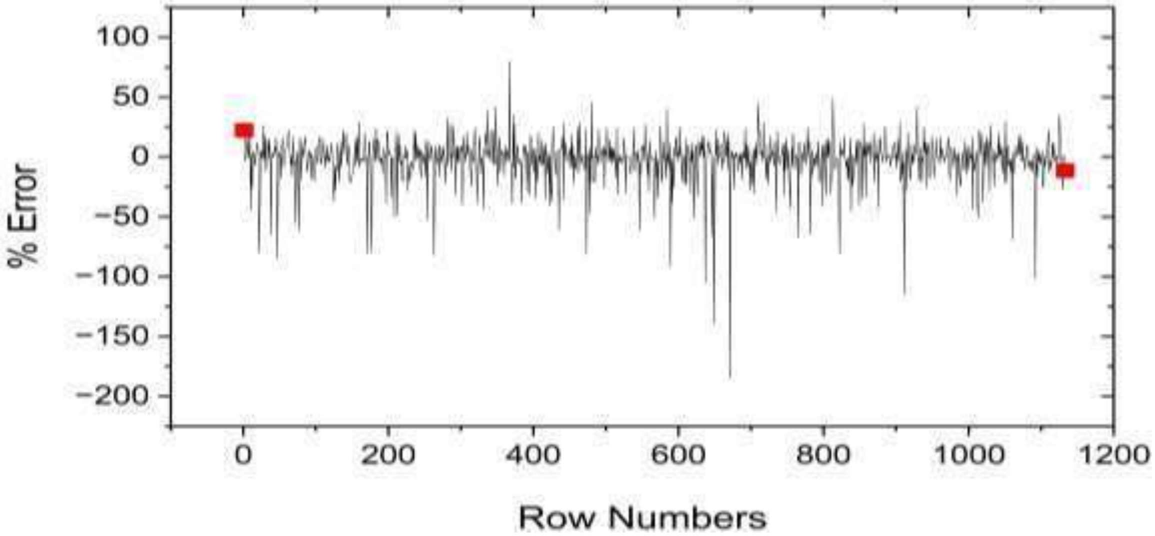
3.2 Error analysis

The error analysis plot presents the percentage error distribution of the Support Vector Machine (SVM) model predictions across all data instances. The horizontal axis represents the row numbers corresponding to individual observations, while the vertical axis shows the percentage error between predicted and actual concrete compressive strength values.

Most of the error values fall within a relatively narrow band of approximately ±30%, demonstrating stable predictive behavior for the majority of samples. This concentration around zero aligns with the low MAPE and RMSE values obtained during model evaluation, confirming the reliability of the SVM model. However, a limited number of data points exhibit large negative percentage errors, reaching values below −150%. These extreme deviations likely correspond to outlier cases with unusual material compositions or long curing ages, where the actual compressive strength values differ significantly from typical patterns captured during training.

The presence of sporadic high-magnitude errors suggests that while the model performs well overall, certain complex or rare combinations of input variables remain challenging to predict accurately. Importantly, these large errors are isolated rather than systematic, indicating that the model's generalization ability remains intact. The absence of visible trends or increasing error magnitude across row numbers further supports model stability.

Figure 5 Error plot percentage difference between experimental compressive strength of concrete and support vector machine results



3.3 Ranking of feature in prediction

Feature ranking was conducted to evaluate the relative importance of input variables in predicting concrete compressive strength using two complementary methods: Univariate Regression and the Relief-based algorithm (RReliefF). The Univariate Regression scores measure the individual predictive strength of each feature with respect to the target variable, while RReliefF captures nonlinear dependencies and feature interactions, providing a more comprehensive importance assessment.

Cement content emerged as the most influential feature, achieving the highest univariate regression score (354.661) and a strong RReliefF weight (0.077), confirming its dominant role in strength development. Age of testing also demonstrated high importance, particularly in the RReliefF ranking (0.097), highlighting the critical influence of curing duration on concrete strength gain. Super-plasticizer ranked second in univariate regression, reflecting its significant effect on workability and hydration efficiency, which indirectly enhances compressive strength.

Water content showed moderate importance across both ranking methods, indicating its balanced but controlled influence on strength through the water–cement ratio. Fine and coarse aggregates exhibited lower univariate scores but maintained consistent RReliefF values, suggesting a stabilizing effect rather than direct strength contribution. Blast-furnace slag and fly ash were ranked lowest, reflecting their conditional influence, which depends heavily on proportioning and curing age. Overall, the feature ranking validates the physical relevance of the input variables and supports the robustness of the SVM model.

Table 5 Summary of input variables and their relative importance in predicting concrete compressive strength.

Rank	Feature	Univariate Regression Score	RReliefF Weight
1	Cement	354.661	0.077
2	Super-plasticizer	163.620	0.057
3	Age of testing	132.525	0.097

4	Water	94.884	0.067
5	Fine Aggregate	31.745	0.079
6	Coarse Aggregate	27.784	0.063
7	Blast-furnace Slag	16.482	0.048
8	Fly Ash	4.711	0.033

4 Discussion

The reported SVM model demonstrates reasonably strong predictive capability, but several aspects warrant critical consideration. An R^2 of 0.89 with RMSE ≈5.3 MPa and MAPE ≈11% is consistent with many SVM- or SVR-based concrete strength models, which often achieve R^2 between 0.85–0.97 and MAPE below 10% [66], [67], [68]. However, more recent work shows that optimized or ensemble models (e.g., CatBoost, XGBoost, boosted trees, tuned SVR) can reach $R^2 \geq 0.93$ with RMSE in the 2–4 MPa range and MAPE ≈2–8% [69], [70], [71], suggesting that the present SVM performance is solid but not state-of-the-art. The error analysis highlights mostly stable errors within ±30%, yet some extreme negative errors below −150%. Comparable studies generally report error bands within ±10–30% for well-tuned models [72], [73], [74]. The presence of such large, isolated errors suggests either influential outliers or regions of the input space poorly covered by training data. Other authors explicitly address this with outlier removal, K-fold cross-validation, or additional regularization/hyperparameter tuning to improve robustness [75], [76], [77]. The current single 80:20 split is less statistically reliable than repeated or K-fold validation commonly adopted in recent literature [78], [79]. Feature-importance results (cement and age as dominant, followed by water and superplasticizer) agree well with sensitivity and SHAP analyses in other studies, which consistently rank age, cement, water content, and superplasticizer among the most influential. However, the relatively low ranking of slag and fly ash should be interpreted cautiously: multiple works show their importance grows at longer ages or specific replacement levels. Relying solely on univariate and RReliefF scores may underrepresent interaction effects that more advanced interpretability tools (e.g., SHAP, partial dependence plots) capture. Overall, the SVM model is technically sound and physically consistent, but its validation strategy, handling of extreme outliers, and comparison with more advanced ML approaches could be strengthened to match current best practice in concrete strength prediction[72], [73], [74], [75], [76], [77], [78], [79].

5 Conclusion and recommendation

The present study demonstrates that Support Vector Machine (SVM) regression is an effective and reliable approach for predicting the compressive strength of concrete using mix composition parameters and curing age. Based on a dataset of 1,133 observations, the developed SVM model successfully captured the complex, nonlinear relationships governing concrete strength development. The obtained performance indicators, including an R^2 value of 0.89, RMSE of 5.32 MPa, and MAPE of 11.38%, confirm that the model provides reasonably accurate and stable predictions, consistent with the physical behavior of concrete materials and comparable with many existing SVM-based studies reported in the literature. Descriptive and advanced statistical analyses further verified that the dataset adequately represents a wide spectrum of concrete mix designs and testing ages, supporting the robustness of the modeling process. Feature ranking results reinforced established engineering knowledge by identifying cement content and curing age as the most influential factors, followed by super-plasticizer and water content, thereby validating the interpretability and physical relevance of the model outputs. Nevertheless, the occurrence of isolated large prediction errors highlights the inherent challenges associated with heterogeneous material behavior and extreme mix proportions.

Despite its satisfactory performance, several limitations suggest directions for improvement and future work. The reliance on a single 80:20 train–test split may limit statistical robustness; therefore, future studies should employ K-fold or repeated cross-validation to enhance reliability and reduce sampling bias. Incorporating systematic outlier detection and treatment strategies could further improve prediction stability, particularly for rare or extreme mix designs. Additionally, while the SVM model performed well, recent advances in ensemble and gradient-boosting algorithms have demonstrated superior accuracy and robustness; comparative analysis with models such as XGBoost, CatBoost, or hybrid ensembles is recommended to benchmark performance against state-of-the-art methods. Further exploration of advanced interpretability techniques, including SHAP values or partial dependence analysis, would provide deeper insights into feature interactions and age-dependent effects of supplementary cementitious materials such as slag and fly ash. Expanding the dataset to include additional curing conditions, environmental factors, or material sources may also enhance model generalization. Overall, the findings confirm the practical applicability of SVM for concrete strength prediction while underscoring the importance of advanced validation, interpretability, and model comparison to achieve higher predictive reliability in engineering applications.

6 References

[1] S. Nawaz et al., “Flexural Performance Evaluation Of Glass Fiber Reinforced Polymer Bars ({GFRP}) In Doubly Reinforced Beams: An Experimental Approach,” 2025.

[2] W. Khan et al., “Transforming Expansive Clays {FROM} Waste {TO} Value: Dual Ash Stabilization Strategy {FOR} Augmenting Black Cotton Subgrade Soil Performance,” vol. 11, pp. 5065–5073, 2025, doi: 10.64252/cz07k049.

[3] B. T. Pham, “BGG-REPT and ROF-REPT: ensemble machine learning models for the prediction of compressive strength of concrete,” Innov. Infrastruct. Solut., 2025, doi: 10.1007/s41062-025-01869-3.

[4] T.-G. Nguyen, T. D. Tran, and P. D. Nguyen, “Prediction of Concrete Compressive Strength Using Artificial Neural Network with Bayesian Optimization,” Mater. Emerg. Technol. Sustain., 2025, doi: 10.1142/s3060932125500074.

[5] A. Kashem, R. Karim, S. C. Malo, P. Das, S. D. Datta, and M. Alharthai, “Hybrid data-driven approaches to predicting the compressive strength of ultra-high-performance concrete using SHAP and PDP analyses,” Case Stud. Constr. Mater., 2024, doi: 10.1016/j.cscm.2024.e02991.

[6] Z. Shen, A. Deifalla, P. Kamiński, and A. Dyczko, “Compressive Strength Evaluation of Ultra-High-Strength Concrete by Machine Learning,” Materials (Basel)., vol. 15, 2022, doi: 10.3390/ma15103523.

[7] A. Ahmad et al., “Prediction of Geopolymer Concrete Compressive Strength Using Novel Machine Learning Algorithms,” Polymers (Basel)., vol. 13, 2021, doi: 10.3390/polym13193389.

[8] A. Ahmad, K. Ostrowski, M. Maślak, F. Farooq, I. Mehmood, and A. Nafees, “Comparative Study of Supervised Machine Learning Algorithms for Predicting the Compressive Strength of Concrete at High Temperature,” Materials (Basel)., vol. 14, 2021, doi: 10.3390/ma14154222.

[9] M. N. Amin, B. Iftikhar, K. Khan, M. F. Javed, A. M. AbuArab, and M. F. Rehman, “Prediction model for rice husk ash concrete using AI approach: Boosting and bagging algorithms,” Structures, 2023, doi: 10.1016/j.istruc.2023.02.080.

[10] Y. Huang, Y. Lei, X. Luo, and C. Fu, “Prediction of compressive strength of rice husk ash concrete: A comparison of different metaheuristic algorithms for optimizing support vector regression,” Case Stud. Constr. Mater., 2023, doi: 10.1016/j.cscm.2023.e02201.

[11] Y. Gao, Z. Li, Y. Li, Z. Zhu, and J. Zhu, “Development of chemistry-informed interpretable model for predicting compressive strength of recycled aggregate concrete containing supplementary cementitious materials,” J. Clean. Prod., 2023, doi: 10.1016/j.jclepro.2023.138733.

[12] Y. Gao, J. Lin, J. Zhou, and M. Zhu, “Using Stacking Machine Learning Models to Predict High-Performance Concrete Compressive Strength,” Proc. 2024 4th Int. Conf. Artif. Intell. Big Data Algorithms, 2024, doi: 10.1145/3690407.3690420.

[13] K. Zhang, X. Li, S. Zhang, and S. Zhang, “A Bio-Inspired Adaptive Probability IVYPSO Algorithm with Adaptive Strategy for Backpropagation Neural Network Optimization in Predicting High-Performance Concrete Strength,” Biomimetics, 2025, doi: 10.3390/biomimetics10080515.

[14] D.-L. Nguyen and T.-D. Phan, “Predicting the compressive strength of ultra-high-performance concrete: an ensemble machine learning approach and actual application,” Asian J. Civ. Eng., 2024, doi: 10.1007/s42107-023-00984-9.

[15] G. Abdullah et al., “Boosting-based ensemble machine learning models for predicting unconfined compressive strength of geopolymer stabilized clayey soil,” Sci. Rep., vol. 14, 2024, doi: 10.1038/s41598-024-52825-7.

[16] M. S. Khan et al., “Explainable AutoML models for predicting the strength of high-performance concrete using Optuna, SHAP and ensemble learning,” Front. Mater., 2025, doi: 10.3389/fmats.2025.1542655.

[17] S. Latif, “Developing a boosted decision tree regression prediction model as a sustainable tool for compressive strength of environmentally friendly concrete,” Environ. Sci. Pollut. Res., vol. 28, pp. 65935–65944, 2021, doi: 10.1007/s11356-021-15662-z.

[18] S. Lee, N. H. Nguyen, A. Karamanli, J. Lee, and T. Vo, “Super learner machine-learning algorithms for compressive strength prediction of high performance concrete,” Struct. Concr., vol. 24, pp. 2208–2228, 2022, doi: 10.1002/suco.202200424.

[19] Z. Chao, Z. Li, Y. Dong, D. Shi, and J. Zheng, “Estimating compressive strength of coral sand aggregate concrete in marine environment by combining physical experiments and machine learning-based techniques,” Ocean Eng., 2024, doi: 10.1016/j.oceaneng.2024.118320.

[20] M. S. N. Tak, Y. Feng, and M. Mahgoub, “Advanced Machine Learning Techniques for Predicting Concrete Compressive Strength,” Infrastructures, 2025, doi: 10.3390/infrastructures10020026.

[21] G.-H. Fang, Z.-M. Lin, C.-Z. Xie, Q.-Z. Han, M.-Y. Hong, and X.-Y. Zhao, “Optimized Machine Learning Model for Predicting Compressive Strength of Alkali-Activated Concrete Through Multi-Faceted Comparative Analysis,” Materials (Basel)., vol. 17, 2024, doi: 10.3390/ma17205086.

[22] X. Cui, Q. Wang, R. Zhang, J. Dai, and S. Li, “Machine learning prediction of concrete compressive strength with data enhancement,” J. Intell. Fuzzy Syst., vol. 41, pp. 7219–7228, 2021, doi: 10.3233/jifs-211088.

[23] P. Li, Z. Zhang, and J. Gu, “Prediction of Concrete Compressive Strength Based on ISSA-BPNN-AdaBoost,” Materials (Basel)., vol. 17, 2024, doi: 10.3390/ma17235727.

[24] W. Taffese and L. Espinosa-Leal, “Multitarget regression models for predicting compressive strength and chloride resistance of concrete,” J. Build. Eng., 2023, doi: 10.1016/j.jobe.2023.106523.

[25] P. Ellappan, L. Keshav, K. C. P. Raja, and G. Sanijya, “Improving forecasting of concrete strength using advanced machine learning methods,” Matéria (Rio Janeiro), 2025, doi: 10.1590/1517-7076-rmat-2024-0789.

[26] L. Kalabarige, J. Sridhar, S. Subbaram, P. Prasath, and R. Gobinath, “Machine Learning Modeling Integrating Experimental Analysis for Predicting Compressive Strength of Concrete Containing Different Industrial Byproducts,” Adv. Civ. Eng., 2024, doi: 10.1155/2024/7844854.

[27] M. Katlav and F. Ergen, “Improved forecasting of the compressive strength of ultra-high-performance concrete (UHPC) via the CatBoost model optimized with different algorithms,” Struct. Concr., vol. 26, pp. 212–235, 2024, doi: 10.1002/suco.202400163.

[28] M. Jibril et al., “High strength concrete compressive strength prediction using an evolutionary computational intelligence algorithm,” Asian J. Civ. Eng., vol. 24, pp. 3727–3741, 2023, doi: 10.1007/s42107-023-00746-7.

[29] Y. Xu et al., “Computation of High-Performance Concrete Compressive Strength Using Standalone and Ensembled Machine Learning Techniques,” Materials (Basel)., vol. 14, 2021, doi: 10.3390/ma14227034.

[30] X. Hu et al., “Predicting triaxial compressive strength of high-temperature treated rock using machine learning techniques,” J. Rock Mech. Geotech. Eng., 2022, doi: 10.1016/j.jrmge.2022.10.014.

[31] H. Bentegri et al., “Assessment of compressive strength of eco-concrete reinforced using machine learning tools,” Sci. Rep., vol. 15, 2025, doi: 10.1038/s41598-025-89530-y.

[32] H. Li, H. Chung, Z. Li, and W. Li, “Compressive Strength Prediction of Fly Ash-Based Concrete Using Single and Hybrid Machine Learning Models,” Buildings, 2024, doi: 10.3390/buildings14103299.

[33] A. A. B. Loureiro and R. Stefani, “Comparing the performance of machine learning models for predicting the compressive strength of concrete,” Discov. Civ. Eng., 2024, doi: 10.1007/s44290-024-00022-w.

[34] T. A. Assegie, A. O. Salau, and T. U. Badrudeen, “Estimation of concrete compression using regression models,” Bull. Electr. Eng. Informatics, 2022, doi: 10.11591/eei.v11i5.4210.

[35] D. Dao et al., “A Sensitivity and Robustness Analysis of GPR and ANN for High-Performance Concrete Compressive Strength Prediction Using a Monte Carlo Simulation,” Sustainability, 2020, doi: 10.3390/su12030830.

[36] G. Arora, D. Kumar, and B. Singh, “Tree based Regression Models for Predicting the Compressive Strength of Concrete at High Temperature,” IOP Conf. Ser. Earth Environ. Sci., vol. 1327, 2024, doi: 10.1088/1755-1315/1327/1/012015.

[37] K. Logeswaran et al., “Machine Learning Model for Concrete Strength Prediction,” 2024 15th Int. Conf. Comput. Commun. Netw. Technol., pp. 1–6, 2024, doi: 10.1109/icccnt61001.2024.10724333.

[38] A. Kumar et al., “Compressive Strength Prediction of Lightweight Concrete: Machine Learning Models,” Sustainability, 2022, doi: 10.3390/su14042404.

[39] V. Rathakrishnan, S. B. Beddu, and A. Ahmed, “Predicting compressive strength of high-performance concrete with high volume ground granulated blast-furnace slag replacement using boosting machine learning algorithms,” Sci. Rep., vol. 12, 2022, doi: 10.1038/s41598-022-12890-2.

[40] J.-F. Jia, X.-Z. Chen, Y. Bai, Y.-L. Li, and Z.-H. Wang, “An interpretable ensemble learning method to predict the compressive strength of concrete,” Structures, 2022, doi: 10.1016/j.istruc.2022.10.056.

[41] L. Wang, J. Zhao, and S. Gao, “Strength property estimation of high-performance concrete with novel admixtures using optimal regression-based approaches,” J. Intell. Fuzzy Syst., vol. 45, pp. 6655–6666, 2023, doi: 10.3233/jifs-230132.

[42] L. Chen, Y. Nouri, N. Allahyarsharahi, H. Naderpour, D. R. Eidgahee, and P. Fakharian, “Optimizing compressive strength prediction in eco-friendly recycled concrete via artificial intelligence models,” Multiscale Multidiscip. Model. Exp. Des., 2024, doi: 10.1007/s41939-024-00641-x.

[43] M. Ahmad et al., “Supervised Learning Methods for Modeling Concrete Compressive Strength Prediction at High Temperature,” Materials (Basel)., vol. 14, 2021, doi: 10.3390/ma14081983.

[44] A. H. A. Ahmed, W. Jin, and M. A. H. Ali, “Comparative analysis of intelligent models for predicting compressive strength in recycled aggregate concrete,” Model. Earth Syst. Environ., vol. 10, pp. 5273–5291, 2024, doi: 10.1007/s40808-024-02063-7.

[45] A. H. A. Ahmed, W. Jin, and M. A. H. Ali, “Prediction of compressive strength of recycled concrete using gradient boosting models,” Ain Shams Eng. J., 2024, doi: 10.1016/j.asej.2024.102975.

[46] A. Beskopylny et al., “Prediction of the Compressive Strength of Vibrocentrifuged Concrete Using Machine Learning Methods,” Buildings, 2024, doi: 10.3390/buildings14020377.

[47] T.-D. Phan, “Fast prediction of the compressive strength of high-performance concrete through a k-nearest neighbor approach,” Asian J. Civ. Eng., pp. 1–16, 2023, doi: 10.1007/s42107-023-00756-5.

[48] S. Paudel, A. Pudasaini, R. Shrestha, and E. Kharel, “Compressive strength of concrete material using machine learning techniques,” Clean. Eng. Technol., 2023, doi: 10.1016/j.clet.2023.100661.

[49] Z. Fei, S. Liang, Y. Cai, and Y. Shen, “Ensemble Machine-Learning-Based Prediction Models for the Compressive Strength of Recycled Powder Mortar,” Materials (Basel)., vol. 16, 2023, doi: 10.3390/ma16020583.

[50] H. Nguyen, T. Vu, T. Vo, and H. Thai, “Efficient machine learning models for prediction of concrete strengths,” Constr. Build. Mater., 2021, doi: 10.1016/j.conbuildmat.2020.120950.

[51] V. S. Aswal, B. Singh, and R. Maheshwari, “Machine learning-based model for prediction of concrete strength,” Multiscale Multidiscip. Model. Exp. Des., 2024, doi: 10.1007/s41939-024-00609-x.

[52] A. Shishegaran, H. Varaee, T. Rabczuk, and G. Shishegaran, “High correlated variables creator machine: Prediction of the compressive strength of concrete,” ArXiv, vol. abs/2009.0, 2020, doi: 10.1016/j.compstruc.2021.106479.

[53] I. K. Harith, W. Nadir, M. Salah, and M. Hussien, “Prediction of high-performance concrete strength using machine learning with hierarchical regression,” Multiscale Multidiscip. Model. Exp. Des., vol. 7, pp. 4911–4922, 2024, doi: 10.1007/s41939-024-00467-7.

[54] R. K. Tipu, Suman, and V. Batra, “Enhancing prediction accuracy of workability and compressive strength of high-performance concrete through extended dataset and improved machine learning models,” Asian J. Civ. Eng., pp. 1–22, 2023, doi: 10.1007/s42107-023-

00768-1.

[55] M. Jibril et al., “Implementation of Nonlinear Computing Models and Classical Regression for Predicting Compressive Strength of High-Performance Concrete,” Appl. Eng. Sci., 2023, doi: 10.1016/j.apples.2023.100133.

[56] B. Pan, W. Liu, P. Zhou, and D. Wu, “Predicting the Compressive Strength of Recycled Concrete Using Ensemble Learning Model,” IEEE Access, vol. 13, pp. 2958–2969, 2025, doi: 10.1109/access.2024.3519669.

[57] H. Khandagale, M. Salunkhe, V. Kamble, A. Manjarekar, P. Halkarnikar, and S. Redekar, “Optimizing Concrete Compressive Strength Prediction: A Comparative Study of Machine Learning Algorithms,” 2024 Int. Conf. Electr. Electron. Comput. Technol., vol. 1, pp. 1–6, 2024, doi: 10.1109/iceect61758.2024.10738970.

[58] M. M. Islam et al., “Prediction of compressive strength of high-performance concrete using optimization machine learning approaches with SHAP analysis,” J. Build. Pathol. Rehabil., 2024, doi: 10.1007/s41024-024-00445-z.

[59] M. Alyami et al., “Application of metaheuristic optimization algorithms in predicting the compressive strength of 3D-printed fiber-reinforced concrete,” Dev. Built Environ., 2024, doi: 10.1016/j.dibe.2023.100307.

[60] M. Elshaarawy, M. Alsaadawi, and A. Hamed, “Machine learning and interactive GUI for concrete compressive strength prediction,” Sci. Rep., vol. 14, 2024, doi: 10.1038/s41598-024-66957-3.

[61] A. Punitha, V. Kumar, R. Swetha, B. Jagadeesh, R. Karthikeyan, and K. Charulatha, “Prediction of compressive strength of blended concrete with Alccofine and GGBFS by applying ensemble machine learning algorithms,” J. Struct. Integr. Maint., vol. 10, 2025, doi: 10.1080/24705314.2025.2489858.

[62] A. Ahmad, W. Ahmad, F. Aslam, and P. Joyklad, “Compressive strength prediction of fly ash-based geopolymer concrete via advanced machine learning techniques,” Case Stud. Constr. Mater., 2021, doi: 10.1016/j.cscm.2021.e00840.

[63] N. Amruthamol, P. Shahbaz, K. Kapoor, and R. Kumar, “Bayesian learning approach to foresee compressive strength of concrete,” Mater. Today Proc., 2023, doi: 10.1016/j.matpr.2023.07.263.

[64] F. Aslam and M. Shahab, “Supplementary Cementitious Materials in Blended Cement Concrete: Advancements in Predicting Compressive Strength through Machine Learning,” Mater. Today Commun., 2023, doi: 10.1016/j.mtcomm.2023.107725.

[65] L. Liu, B. Du, K. Yu, and W. Wei, “Prediction of compressive strength of high-performance concrete via coupled meta-heuristic random forest regression techniques,” Multiscale Multidiscip. Model. Exp. Des., pp. 1–15, 2023, doi: 10.1007/s41939-023-00256-8.

[66] Y. Wu and Y. Zhou, “Hybrid machine learning model and Shapley additive explanations for compressive strength of sustainable concrete,” Constr. Build. Mater., 2022, doi: 10.1016/j.conbuildmat.2022.127298.

[67] M. Wang, J. Kang, W. Liu, J. Su, and M. Li, “Research on prediction of compressive strength of fly ash and slag mixed concrete based on machine learning,” PLoS One, vol. 17, 2022, doi: 10.1371/journal.pone.0279293.

[68] W. Zhang, D. Liu, and K. Cao, “Prediction of concrete compressive strength using support vector machine regression and non-destructive testing,” Case Stud. Constr. Mater., 2024, doi: 10.1016/j.cscm.2024.e03416.

[69] S. Pal, L. H. Trang, V. Hieu, D. Nguyen, D. Q. Vu, and I. Prakash, “Investigation of Support Vector Machines with Different Kernel Functions for Prediction of Compressive Strength of Concrete,” J. Sci. Transp. Technol., 2024, doi: 10.58845/jstt.utt.2024.en.4.2.55-68.

[70] M. Alyami et al., “Predictive Modeling for Compressive Strength of 3D Printed Fiber-Reinforced Concrete Using Machine Learning Algorithms,” Case Stud. Constr. Mater., 2023, doi: 10.1016/j.cscm.2023.e02728.

[71] Y.-M. Hong, “Performance Comparison of Machine Learning Models for Concrete Compressive Strength Prediction,” Materials (Basel)., vol. 17, 2024, doi: 10.3390/ma17092075.

[72] D. Feng et al., “Machine learning-based compressive strength prediction for concrete: An adaptive boosting approach,” Constr. Build. Mater., 2020, doi: 10.1016/j.conbuildmat.2019.117000.

[73] A. Kumar, S. Sen, and S. Sinha, “Support vector machine-based prediction model for the compressive strength for concrete reinforced with waste plastic and fly ash,” Asian J. Civ. Eng., vol. 26, pp. 1429–1447, 2025, doi: 10.1007/s42107-024-01256-w.

[74] T. Cihan, “Prediction of Concrete Compressive Strength and Slump by Machine Learning Methods,” Adv. Civ. Eng., 2019, doi: 10.1155/2019/3069046.

[75] A. Abd and S. Abd, “Modelling the strength of lightweight foamed concrete using support vector machine (SVM),” Case Stud. Constr. Mater., vol. 6, pp. 8–15, 2017, doi: 10.1016/j.cscm.2016.11.002.

[76] A. Beskopylny et al., “Concrete Strength Prediction Using Machine Learning Methods CatBoost, k-Nearest Neighbors, Support Vector Regression,” Appl. Sci., 2022, doi: 10.3390/app122110864.

[77] M. M. Woldeamanuel and H.-K. Kim, “Parametric study on global estimation models for compressive strength adopting various machine learning algorithms in concrete,” Eng. Appl. Artif. Intell., vol. 136, p. 108888, 2024, doi: 10.1016/j.engappai.2024.108888.

[78] V. Vakharia and R. Gujar, “Prediction of compressive strength and portland cement composition using cross-validation and feature ranking techniques,” Constr. Build. Mater., 2019, doi: 10.1016/j.conbuildmat.2019.07.224.

[79] Q.-H. Le et al., “Machine learning based models for predicting compressive strength of geopolymer concrete,” Front. Struct. Civ. Eng., vol. 18, pp. 1028–1049, 2024, doi: 10.1007/s11709-024-1039-5.