E-ISSN : 2541-5794
P-ISSN : 2503-216X

Journal of Geoscience,
Engineering, Environment, and Technology
Vol 10 No 2 2025

RESEARCH ARTICLE

Optimization of Machine Learning Algorithms Through Outlier Data Separation for Predicting Concrete Compressive Strength

Faisal Ananda [1,*], Hendra Saputra[1] , Nurul Fahmi[2] ,Eko Prayitno[2] , Sinatu Sadiah Shapie[3], Mohamad Azwan Bin Ikhwat[4], Mohd Nur Azmi Nordin[5], Andicha Zain[6], Fadhillah Binti Mohd. Nasir[7]

[1] Politeknik Negeri Bengkalis, Department of Civil Engineering, Jl. Bathin Alam-Sei Alam Bengkalis, Indonesia.
[2] Politeknik Negeri Bengkalis, Department of Informatics Engineering, Jl. Bathin Alam-Sei Alam Bengkalis, Indonesia.
[3] Polytechnic Port Dickson, Department of Civil Engineering, Jl. Jalan Pantai 71050 Si Rusa Negeri Sembilan, Malaysia.
[4] Polytechnic Malacca, Department of Civil Engineering, Plaza Pandan Malim, 75250 Melaka, Malaysia
[5] Universiti Teknikal Malaysia Melaka, Department of Mechanical Engineering, Hang Tuah Jaya, 76100 Durian Tunggal, Melaka, Malaysia
[6] Former Doctoral Candidate in Mechanical Engineering, University of Miskolc, Egyetem út 1, 3515 Miskolc, Hungary
[7] Polytechnic Merlimau, Department of Civil Engineering, Jalan Merlimau-Jasin, 77300 Merlimau, Melaka, Malaysia

* Corresponding author : faisalananda@polbeng.ac.id
Tel.:+62-8127635964 fax: -

Received: Mar 28, 2025; Accepted: Jun 3, 2025.
DOI: 10.25299/jgeet.2025.10.02.21896

Abstract

This study investigates the comparative performance of ten machine learning models—Linear Regression, SVM, Neural Network, Decision Tree, Random Forest, Gradient Boosting, AdaBoost, XGBoost, LightGBM, and CatBoost—in predicting concrete compressive strength. The research emphasizes practical applications in construction, where accurate predictions can improve material design and structural reliability. Through detailed evaluation using MAE, RMSE, and R^2 metrics, CatBoost and Linear Regression emerged as top-performing models. A rigorous hyperparameter tuning process, employing grid search, significantly enhanced models like SVM and Neural Network, increasing their R^2 by over 80%. However, tuning occasionally led to reduced performance due to overfitting or unsuitable parameter selection. Outlier analysis using the Z-score method revealed nuanced effects across models: while SVM and Decision Tree benefited from outlier removal, models like Neural Network and CatBoost experienced performance degradation, indicating their reliance on diverse data patterns. These findings underscore the importance of tailored tuning and outlier handling strategies. Future work will incorporate advanced optimization techniques (e.g., Bayesian optimization) and robust cross-validation to further improve model generalization and stability.

Keywords: Predictive Models, Outlier, Evaluations, hyperparameter tuning

1. Introduction

The optimization of machine learning algorithms for predicting concrete compressive strength has emerged as a pivotal area of research within civil engineering. Concrete, being one of the most widely used construction materials, necessitates accurate predictions of its compressive strength to ensure structural integrity and safety. Traditional methods for assessing concrete strength often involve time-consuming and destructive testing procedures, which can be both costly and impractical. In contrast, machine learning (ML) offers a promising alternative by enabling non-destructive and efficient predictions based on historical data (Hassan et al., 2024; Tak et al., 2025).

The accuracy of compressive strength forecasts could be greatly improved by recent developments in machine learning techniques. Numerous studies have looked into modeling the intricate interactions between concrete mix design parameters and compressive strength results using sophisticated algorithms like ensemble learning and deep learning (Alyami et al., 2025; Luo et al., 2024; Vargas et al., 2024). By using big datasets to find trends and forecast outcomes, these models lessen the need for empirical formulas that might not take into consideration all relevant variables (Anwar, 2025; Dong, 2025). For example, studies have demonstrated that deep learning models can predict concrete strength more accurately than conventional regression techniques, underscoring the significance of algorithm selection in attaining the best outcomes (Guzmán-Torres et al., 2024; R. Kumar et al., 2024; Mahmood et al., 2023) .

However, the existence of outliers in the dataset presents a significant obstacle to using machine learning for this purpose. Measurement errors, changes in material characteristics, or irregularities in the mixing process can all result in outliers. These outliers have the potential to distort the results and produce erroneous forecasts if they are not addressed (Liu et al., 2024). To increase the resilience and dependability of machine learning models, outliers must be removed. To lessen the impact of these anomalies and guarantee that the models are trained on high-quality data, methods like data normalization and recursive outlier elimination have been proposed (Parmo and Wardhana, 2024; Sathiparan, 2025). Furthermore, it has been demonstrated that employing strong statistical techniques to identify and manage outliers greatly improves model performance (Jiang et al., 2021; Li et al., 2024).

Furthermore, it has been demonstrated that adding domain knowledge to machine learning frameworks improves their predictive power. Researchers can improve their models to more accurately represent the underlying physical mechanisms determining concrete strength by integrating knowledge from material science and engineering (Liu et al., 2025; Miyan et al., 2024). In addition to increasing forecast accuracy, this multidisciplinary approach promotes a better comprehension of the variables affecting tangible performance (Fan et al., 2025; Khan, 2023). For instance, research has shown that adding environmental and chemical composition variables to prediction models can result in more precise concrete strength estimates (Lv et al., 2025; Moutassem and Chidiac, 2016).

In addition, the use of hybrid models which blend machine learning with conventional engineering principles has become more popular. By combining the advantages of both strategies, these models enable more thorough forecasts that take into account both data-driven insights and well-established engineering knowledge (Asteris et al., 2021; Dash, 2023). This pattern is part of a larger trend in the construction sector to adopt data-driven approaches, which are becoming more and more important for enhancing quality control and maximizing material use (Liang et al., 2024; Marchiori et al., 2025).

In conclusion, a major breakthrough in the field of concrete compressive strength prediction has been made possible by the optimization of machine learning algorithms in conjunction with efficient outlier elimination techniques. The potential for machine learning to transform concrete testing and quality assurance is becoming more and more clear as the construction sector adopts data-driven approaches. The purpose of this study is to investigate these developments by offering a thorough evaluation of the state of knowledge at the moment and suggesting areas for further research.

While the adoption of machine learning (ML) techniques for predicting concrete compressive strength continues to grow, the majority of existing studies tend to concentrate narrowly on identifying the highest-performing algorithms, often neglecting the combined impact of parameter tuning strategies and outlier management on overall model performance. This presents a notable research gap, particularly given that models such as Support Vector Machines (SVM) and Neural Networks are highly sensitive to data distribution and the presence of noise. Additionally, there has been limited investigation into how these models respond to outlier removal and hyperparameter tuning within the context of realistic and heterogeneous construction datasets. To address this gap, the present study offers a comprehensive and comparative analysis of the effects of parameter tuning and outlier handling across several ML algorithms, including CatBoost, SVM, Neural Networks, Linear Regression, and Decision Trees. The novelty of this research lies in its contextual examination of algorithm sensitivity to extreme values, as well as the implementation of adaptive tuning and validation techniques. Consequently, this study not only enhances methodological insight but also delivers practical contributions toward improving prediction accuracy and quality control in civil engineering applications, reinforcing the relevance of ML in real-world construction environments.

2. Methodology

The methodology for this research focuses on optimizing machine learning algorithms to predict concrete compressive strength while effectively eliminating outliers from the dataset. The study utilizes the data set provided https://www.kaggle.com/code/abirchodha/cement-slump-svr/input (Chodha, 2024), which contains various parameters influencing concrete strength, including material composition, mixing ratios, and curing conditions.

2.1. Data Collection and Preprocessing

The initial step involves collecting the dataset, which includes comprehensive records of concrete mixtures and their corresponding compressive strength values. The dataset is then used without undergoing preprocessing steps, such as handling missing values, normalizing the data, or converting categorical variables into numerical formats

2.2. Machine Learning Model Selection

Following the outlier elimination process, ten different machine learning algorithms are employed to predict concrete compressive strength. These algorithms include:

Linear Regression
Support Vector Machine (SVM)
Decision Tree
XGBoost
Neural Network
Random Forest
Gradient Boosting
AdaBoost
LightGBM
CatBoost

Each of these algorithms is selected for its unique strengths in handling various types of data and its ability to model complex relationships between input features and the target variable (Huang et al., 2023).

2.3. Model Training and Evaluation

The selected machine learning models are trained using the preprocessed dataset. A portion of the data is reserved for testing to evaluate the performance of each model. The models are assessed based on several metrics, including Mean Absolute Error (MAE), Mean Squared Error (MSE), and R-squared values, to determine their predictive accuracy (Talpur et al., 2025).

Mean Absolute Error (MAE), R-squared (R^2), and Root Mean Squared Error (RMSE) are three primary evaluation metrics commonly used in machine learning, particularly for regression tasks. Below is a detailed elaboration of these metrics:

2.3.1 Mean Absolute Error (MAE):

Description: MAE calculates the average absolute difference between the model's predictions and the actual values.

Advantage: This metric is intuitive and provides a direct representation of the average error without assigning additional weight to larger errors.

Objective: The higher the MAE value, the better the model's ability to explain the data.

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|x_i - x| \quad (1)$$

MAE = *Mean Absolute Error*
n = *number of data points*
xi = *predicted value*

2.3.2 R-squared (R^2):

Description: R^2 measures how well the model explains the variation in the target data compared to a baseline model (mean of the target).

Advantage: This metric helps evaluate the level of fit between the predictions and the actual data. R^2 values range from 0 to 1, where a value closer to 1 indicates a high level of fit.

Objective: The higher the R^2 value, the better the model's ability to explain the data.

$$R^2 = 1 - \frac{SSR}{SST} \quad (2)$$

$$SSR = \sum x_i - \bar{x} \quad (3)$$

$$SST = \sum x - \bar{x} \quad (4)$$

SSR = Sum Squared Regression
xi = Prediction Average
x = Average
R^2 = R Squared

2.3.2 Root Mean Squared Error (RMSE):

Description: RMSE is the square root of the mean of the squared errors. This metric is more sensitive to larger errors as it assigns additional weight through squaring.

Advantage: It is well-suited for identifying outliers or significant errors that impact the model's performance.

Objective: A lower RMSE value indicates overall better predictions.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - x)^2} \quad (5)$$

2.4. Optimization and Hyperparameter Tuning

To further enhance the performance of the models, hyperparameter tuning is conducted using techniques Grid Search. This process involves systematically testing different combinations of hyperparameters to identify the optimal settings for each algorithm (Choi et al., 2023).

2.5. Outlier Detection and Elimination

To improve the accuracy of predictive models, it is essential to identify and remove outliers from the dataset. The method I used is the Z-score method. The Z-score method calculates the standard score for each data point, indicating how far an element deviates from the mean in terms of standard deviations. Data points with a Z-score greater than 3 or less than -3 are considered outliers and are removed from the dataset (Aggarwal et al., 2019).

2.6. Results Analysis

The results from each model are analyzed to identify the most effective algorithm for predicting concrete compressive strength. These findings are compared with existing literature to validate the effectiveness of the proposed methodology and to highlight the improvements achieved through outlier removal and model optimization(Ghafoorian Heidari et al., 2024).

3. Results and Discussion

3.1. Machine Learning Performance Review

In the ever evolving domain of machine learning, the evaluation and comparison of model performance remain critical to ensuring the reliability and robustness of predictive systems. Performance reviews in this context go beyond mere accuracy, delving into metrics that comprehensively measure how well a model captures patterns, generalizes to unseen data, and minimizes errors.

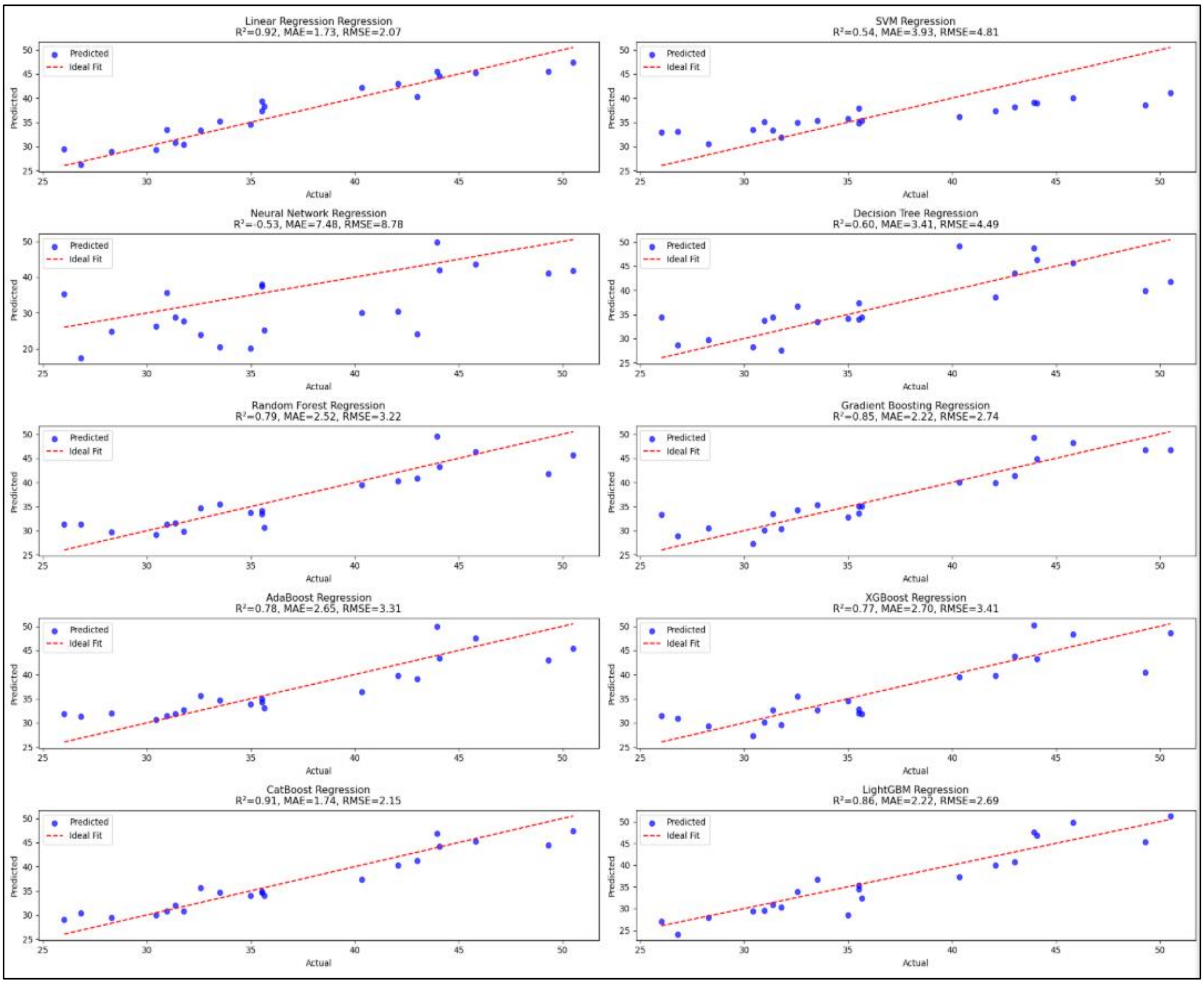

Fig. 1. Regression Results from Machine Learning

In the ever evolving domain of machine learning, the evaluation and comparison of model performance remain critical to ensuring the reliability and robustness of predictive systems. Performance reviews in this context go beyond mere accuracy, delving into metrics that comprehensively measure how well a model captures patterns, generalizes to unseen data, and minimizes errors.

One of the key aspects of a machine learning performance review involves selecting appropriate evaluation metrics. Metrics such as Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and R-squared (R^2) are pivotal in assessing regression models. While MAE provides insights into the average magnitude of errors, RMSE emphasizes larger errors due to its sensitivity to squared differences, offering a nuanced perspective on model reliability. Similarly, R^2 quantifies how effectively the model explains variance in the target variable, acting as a holistic measure of performance.

Based on the results of the visual analysis presented in Figure 2, conducted using models such as Gradient Boosting, Support Vector Machine (SVM), Decision Tree, XGBoost, Neural Network, Random Forest, K-Nearest Neighbors (KNN), AdaBoost, LightGBM, and CatBoost, the accuracy levels of these models and their error patterns were identified. To evaluate accuracy and error patterns, three evaluation metrics were used: Mean Absolute Error (MAE), R-squared (R^2), and Root Mean Squared Error (RMSE). The results of these metrics, obtained from running the models, are presented in Table 1.

The data utilized in this study comprises concrete data collected from a dataset. During the model analysis process, the data was split into two parts: 80% for training and 20% for testing. This splitting process was performed directly within the modeling framework.

Table 1. Machine Learning Model Performance Based on MAE, R^2, and RMSE.

Model ML	MAE	R^2	RMSE
Linear Regression	1.730261	0.915435	2.06851
SVM	3.9314	0.541977	4.814011
Neural Network	7.475249	-0.525024	8.784189
Decision Tree	3.407143	0.602373	4.485402
Random Forest	2.5185	0.794447	3.224964
Gradient Boosting	2.215504	0.852045	2.736073
AdaBoost	2.65459	0.783474	3.309929
XGBoost	2.703838	0.769589	3.414402
CatBoost	1.736123	0.9083	2.154007
LightGBM	2.219146	0.856856	2.691226

In the evaluation of machine learning model performance on the concrete dataset, three primary metrics were used to analyze accuracy and error patterns: MAE, R^2, and RMSE. CatBoost emerged as the top-performing model with an MAE of 1.736123, an R^2 of 0.9083, and an RMSE of 2.154007. This demonstrates its ability to deliver highly accurate predictions by effectively capturing complex relationships in the data. Similarly, Linear Regression also exhibited strong results, achieving an MAE of 1.730261, an R^2 of 0.915435, and an RMSE of 2.06851. This indicates that even a simpler model can perform competitively under the right conditions. On the other hand, SVM and Neural Network struggled to achieve satisfactory results. SVM recorded an MAE of 3.9314, an R^2 of 0.541977, and an RMSE of 4.814011, reflecting moderate accuracy. Neural Network, however, performed poorly, with an MAE of 7.475249, an R^2 of -0.525024, and an RMSE of 8.784189, indicating a failure to model the relationships in the data effectively.

Models such as Gradient Boosting, LightGBM, and Random Forest also delivered notable results, with R^2 values exceeding 0.85 and MAE and RMSE remaining within acceptable ranges. For instance, Gradient Boosting achieved an MAE of 2.215504 and an R^2 of 0.852045, while LightGBM attained an R^2 of 0.856856 and an RMSE of 2.691226. These models demonstrated strong predictive capabilities and stand as competitive alternatives to CatBoost and Linear Regression. Decision Tree and AdaBoost also exhibited reasonable performance, although their accuracy was slightly lower compared to the boosting and ensemble-based methods.

Visual analysis through regression scatter plots illustrated the differences in performance. CatBoost and LightGBM showed predicted values closely aligned with the diagonal ideal-fit line, indicating high accuracy, whereas SVM and Neural Network displayed notable deviations. Comparative bar charts of the metrics further highlighted CatBoost and Linear Regression as the most reliable models, while Neural Network and SVM lagged significantly.

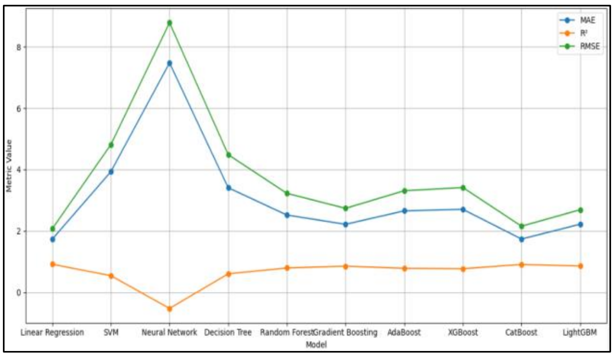

Fig. 2. Machine Learning Model Performance

3.2. Optimization and Hyperparameter Tuning

Optimization and hyperparameter tuning are crucial steps in improving the performance of machine learning models. Optimization focuses on adjusting model parameters and configurations to minimize error and maximize predictive accuracy.

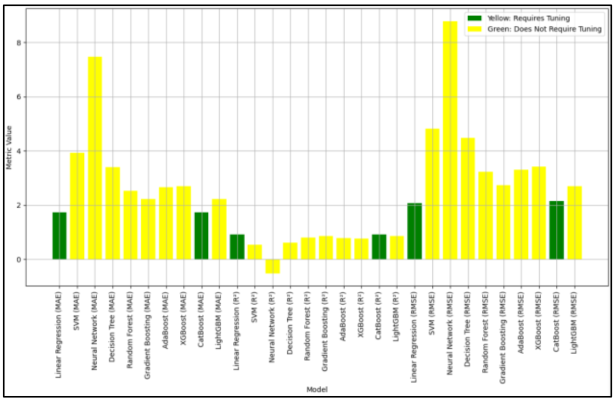

Fig. 3. Tuned Machine Learning Model

Hyperparameter tuning, on the other hand, involves finding the best combination of hyperparameters settings that control the model's behavior such as learning rate, number of estimators, or regularization strength. Proper tuning ensures that the model generalizes well to unseen data while avoiding overfitting or underfitting.

The evaluation process involved assessing the performance of various machine learning models using key metrics such as R^2, MAE, and RMSE. Thresholds were established to determine model adequacy, with minimum values set at $R^2 \geq 0.90$, MAE ≤ 2.5, and RMSE ≤ 3.5. Models meeting these thresholds were marked in green, signifying no need for further tuning, while those falling short were highlighted in yellow, indicating the necessity for hyperparameter tuning. A detailed results table was generated with these color-coded markings for clarity.

Based on Figure. 3 the analysis of the graph and the performance thresholds, only Linear Regression and CatBoost do not require tuning as they have met the criteria with low errors and high R^2 values. The other models, including SVM, Neural Network, Decision Tree, Random Forest, Gradient Boosting, LightGBM, and AdaBoost, need tuning. These models fail to meet the thresholds due to either higher MAE, RMSE, or lower R^2 values. For instance, SVM and Neural Network exhibit significant prediction errors and low R^2, indicating weak performance. Similarly, AdaBoost and Decision Tree, while performing moderately, still produce errors that can be reduced through optimization. The algorithm was tuned using the grid search technique, which offers several advantages, such as being more systematic, having a high likelihood of finding the best combination, being consistent, and being well-suited for small to medium-sized datasets.

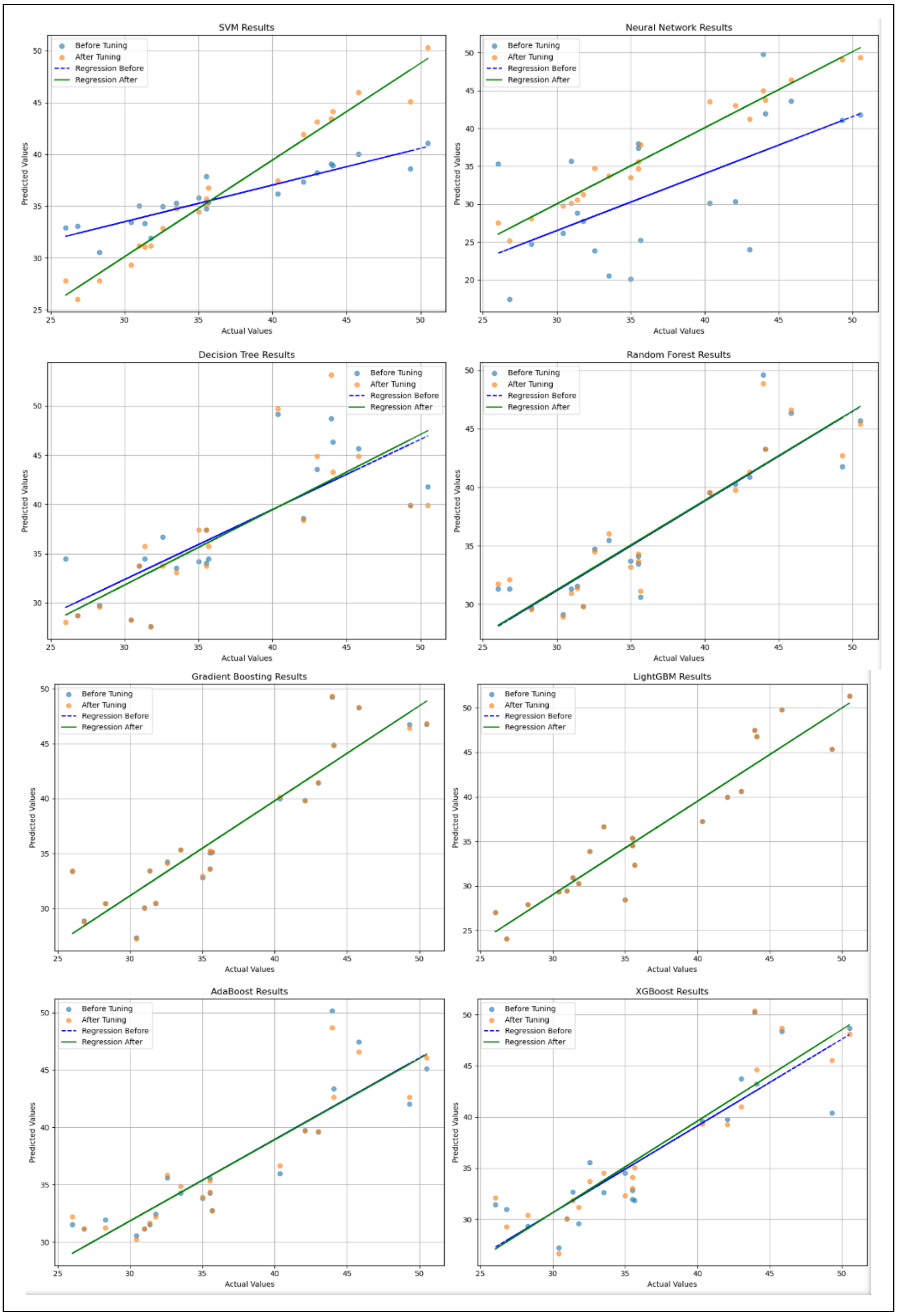

Fig. 4. Performance Comparison: Before and After Tuning

Table 2. Performance Metrics Before and After Tuning

Model ML	MAE Before	MAE After	R^2 Before	R^2 After	RMSE Before	RMSE After
SVM	3.93	0.83	0.54	0.97	4.81	1.31
Neural Network	7.48	1.06	-0.53	0.97	8.78	1.32
Decision Tree	3.41	3.43	0.60	0.57	4.49	4.69
Random Forest	2.52	2.50	0.79	0.80	3.22	3.16
Gradient Boosting	2.22	2.21	0.85	0.85	2.74	2.75
LightGBM	2.22	2.22	0.86	0.86	2.69	2.69
AdaBoost	2.62	2.48	0.77	0.80	3.40	3.16
XGBoost	2.70	2.26	0.77	0.85	3.41	2.79

Table 3. Performance Improvements Percentage (%)

Model ML	MAE	R^2	RMSE
SVM	85.79	285.52	72.83
Neural Network	-0.80	-6.16	84.99
Decision Tree	0.86	1.00	-4.56
Random Forest	0.40	-0.05	2.26
Gradient Boosting	0.00	0.00	-0.66
LightGBM	5.34	4.05	0.00
AdaBoost	16.29	9.94	7.07
XGBoost	85.79	285.52	18.26

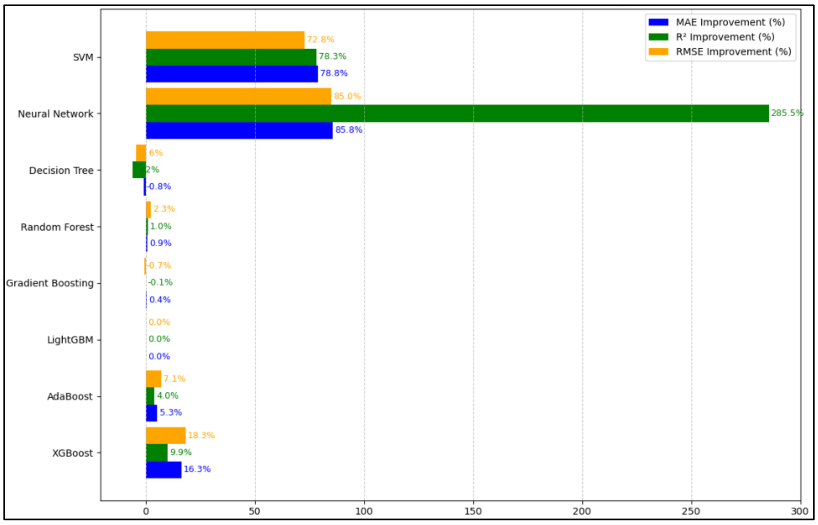

Fig 5. Performance Improvements Percentage (%)

According to Figure 4 and further detailed in Figure 5 illustrates the significant effects of the hyperparameter tuning process on various machine learning models in the context of concrete data. Hyperparameter tuning aims to optimize the model parameters, such as learning rate, max depth, the number of estimators, or kernel functions, enabling the models to learn the patterns in the concrete data more effectively. As a result, performance improvements are observed in metrics such as MAE, R^2, and RMSE, where the average prediction error decreases (MAE), the model's ability to capture data variability improves (R^2), and the error deviation becomes smaller (RMSE).

From the conditions before tuning, several issues were identified, including:

- Default Hyperparameters: The model was operating with default values that might not be suitable for the analyzed dataset (A Ilemobayo et al., 2024).
- Scaling or Data Normalization Issues: The data may not have been processed correctly, making it difficult for the model to interpret the values (Wang et al., 2025).
- Model Complexity: For SVM, the kernel used might not have been suitable, while for the Neural Network, the number of neurons or layers might have been either too few or excessive (N. Kumar et al., 2024).

The substantial benefit of this tuning lies in the accuracy of predicting concrete compressive strength, which is a critical parameter in designing concrete structures. With optimized models, such as Neural Network and XGBoost, which demonstrate significant improvements, the prediction results become more accurate and reliable for real-world applications, such as identifying the optimal concrete mix design or detecting potential material failures prior to construction (Tak et al., 2025). Overall, tuning helps enhance efficiency, reduce risks, and support data-driven decision-making in the construction industry. The visualization in this diagram clearly highlights the positive impacts of tuning while facilitating the evaluation of which models are most responsive to the optimization process.

After hyperparameter tuning, it is possible for the performance of machine learning models to decline instead of improve, as highlighted in the results (Rusman et al., 2023). This phenomenon occurs due to several factors. One common reason is overfitting, where the chosen parameters are overly tailored to the training data, leading to reduced generalization and poorer performance on test data. Additionally, tuning may fail to find the optimal parameters if the parameter grid is too narrow or does not encompass values that best suit the dataset. For models like Decision Tree or Gradient Boosting, their sensitivity to data quality (such as noise or outliers) can amplify adverse effects during tuning, causing deviations in performance. Furthermore, if the dataset is limited or imbalanced, tuning could inadvertently make the model learn irrelevant patterns, degrading its predictive capability (Anugerah Simanjuntak et al., 2024). For certain models, such as LightGBM, default parameters may already perform well, meaning tuning offers no significant benefits and occasionally worsens results. This underscores the importance of evaluating tuning outcomes carefully using test data to confirm its effectiveness.

3.3. Outlier Detection and Elimination

Outliers are data points that are significantly different or distant from most other data within a dataset. In the context of statistics or data analysis, outliers are often considered "unusual" data that can influence the results of analysis or predictive models. In the process of analyzing data to identify outliers, the Z-Score method is used.

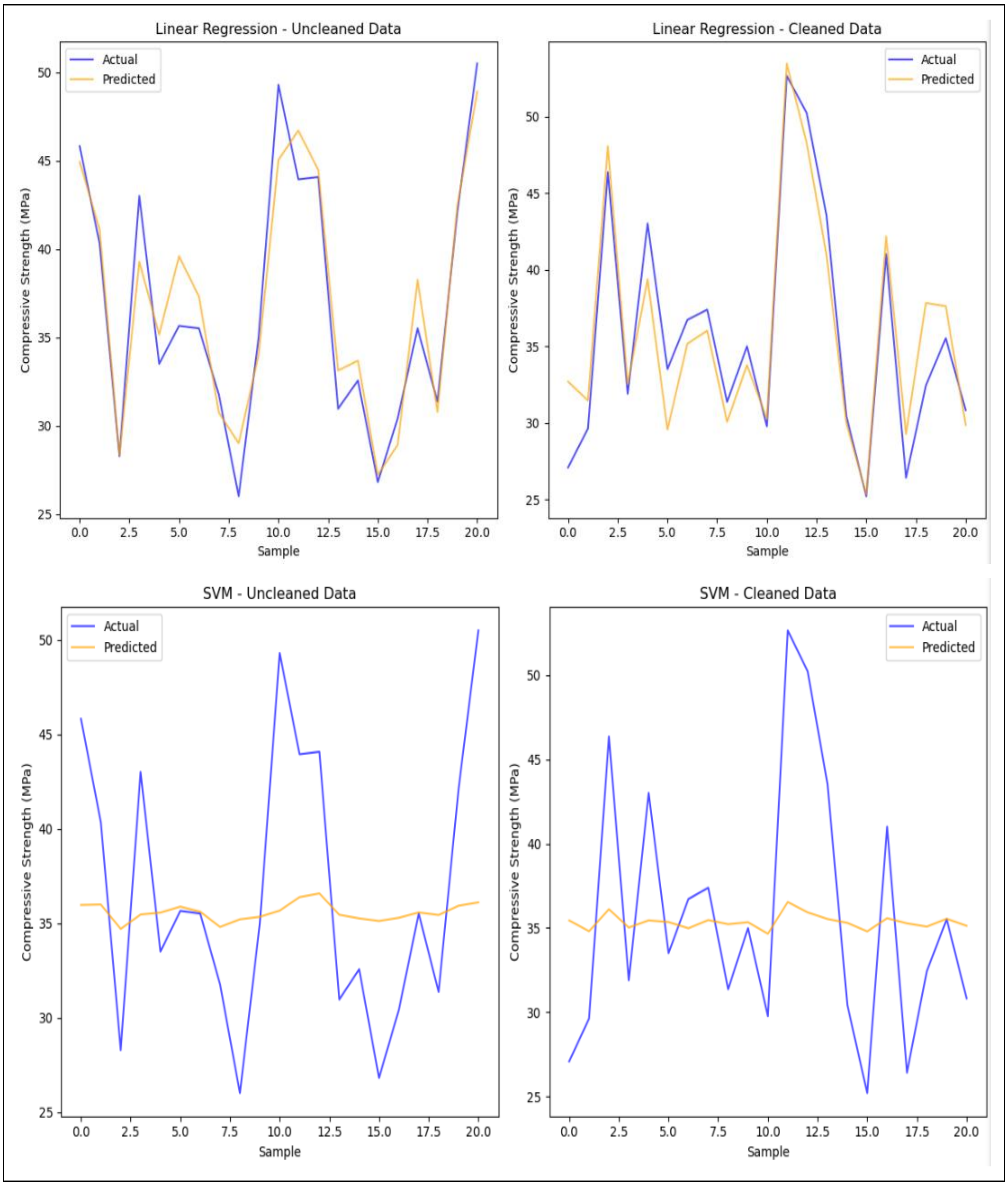

Fig 6. Comparison of actual and predicted concrete compressive strength before and after outliers are removed using the Linear Regression and SVM methods.

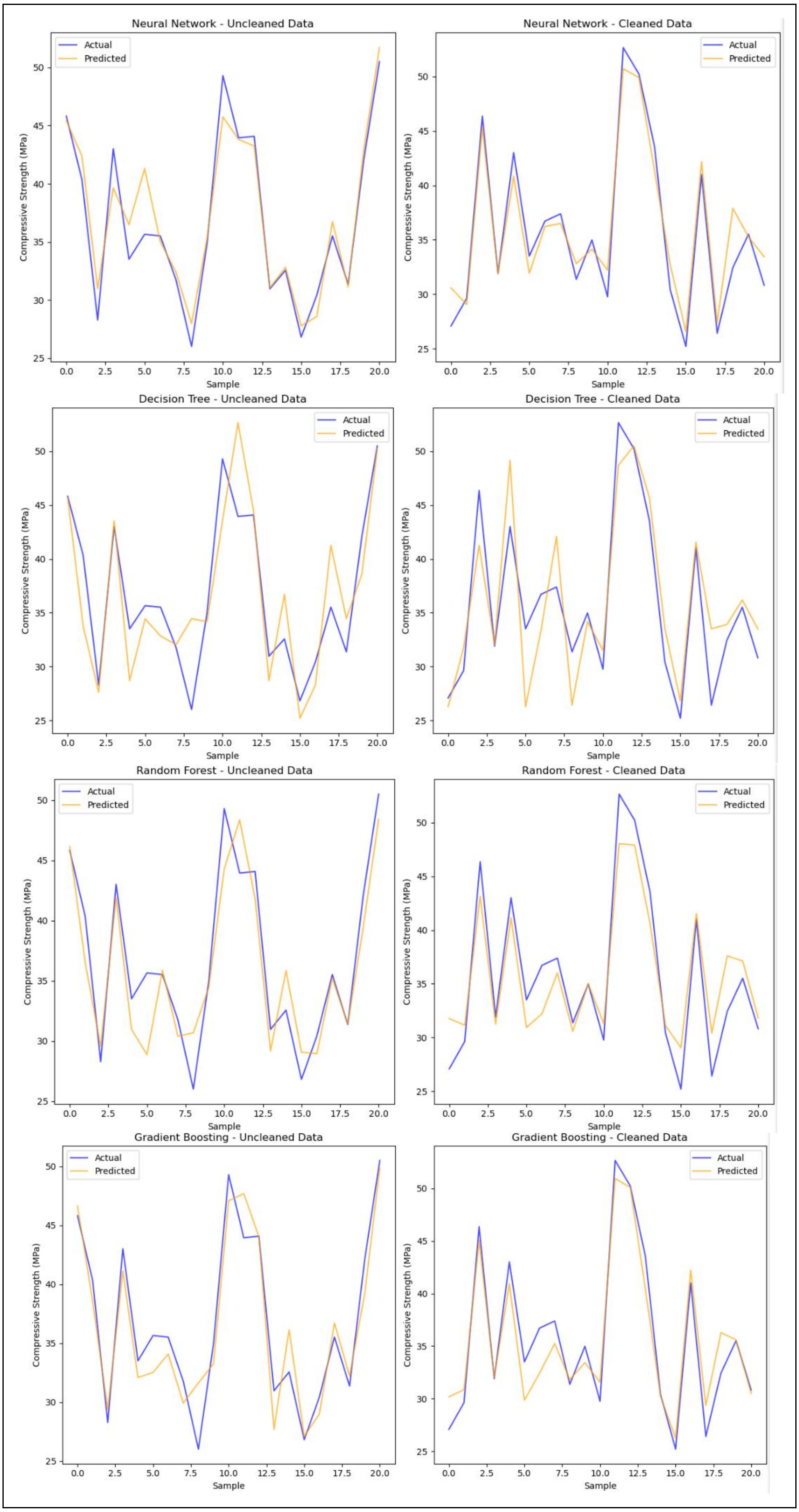

Fig 7. Comparison of actual and predicted concrete compressive strength before and after outliers are removed using the Neural Networks, Decision Tree, Random Forest and Gradient Boosting methods.

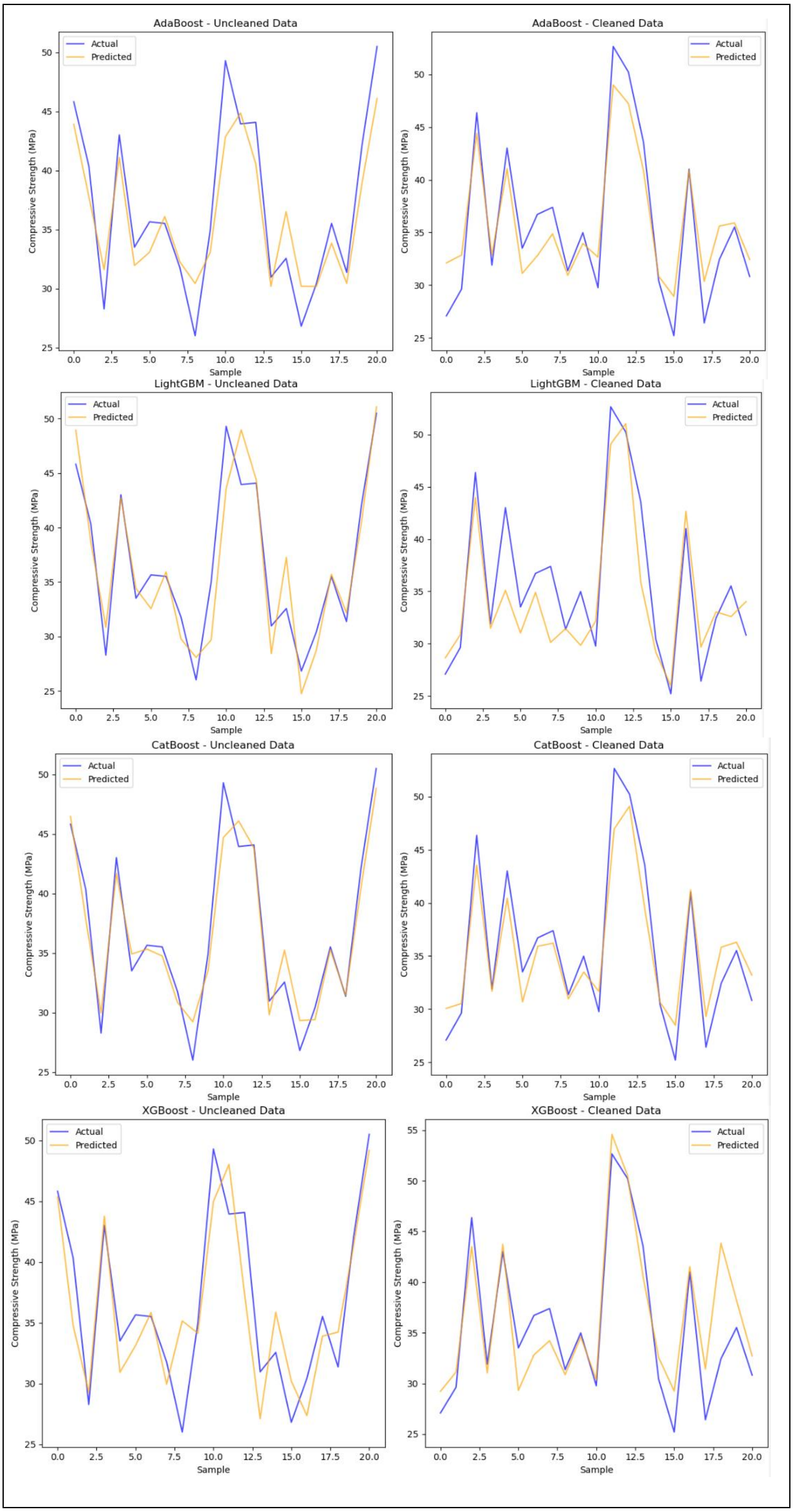

Fig 8. Comparison of actual and predicted concrete compressive strength before and after outliers are removed using the AdaBoost, XGBoost, LightGBM and Cat Boost methods.

From Table 4, it can be seen that the Z-Score method retains almost all of the original data, leaving 101 out of 103 data points. This is because the Z-Score method identifies outliers using the mean and standard deviation, making it more adaptive in preserving data that has a nearly normal distribution.

Table 4. Comparison of Z-score and IQR methods

Method	Uncleaned Data	Cleaned Data	R2	MSE
Z score	103	101	0.87	7.61

The data presented in Table 5 based on Figures 6, 7 and 8 illustrates the performance comparison of various machine learning models before and after outlier removal, based on the R^2 value. Outlier removal has a significant impact on prediction accuracy, with some models showing positive changes in accuracy while others experience negative changes. For example, the SVM model shows a substantial increase in R^2, increasing from 0.065774 to 0.089339, which corresponds to a change in accuracy of 35.83%, although the condition is far from perfect in terms of prediction accuracy (Touati et al., 2025). This increase indicates that the SVM benefits from outlier removal, likely because it is sensitive to extreme values that can distort its predictions. Similarly, the Decision Tree shows a marked positive change, with R^2 increasing by 28.32%, indicating its ability to generalize better without the influence of outliers (Gokcesu et al., 2019). Conversely, models like Neural Network and CatBoost exhibit negative accuracy changes, with R^2 decreasing by 17.52% and 8.70%, respectively (El Hachimi et al., 2025; Yu et al., 2024). This decline may be attributed to the loss of valuable information contained within the outliers, which could have contributed to the predictive power of these models. Neural Network, for example, relies on complex patterns in the data, and the removal of outliers might disrupt these patterns, leading to reduced performance.

Table 5. Machine Learning Model Performance Based R^2 on before and after outliers are removed

Model ML	R^2 (Uncleaned Data)	R^2 (Cleaned Data)	Accuracy Change (%)
Linear Regression	0.91304	0.892171	-2.29
SVM	0.065774	0.089339	35.83
Neural Network	0.892516	0.736192	-17.52
Decision Tree	0.533923	0.685147	28.32
Random Forest	0.842795	0.878786	4.27
Gradient Boosting	0.859871	0.901632	4.86
AdaBoost	0.831834	0.849613	2.14
XGBoost	0.742747	0.784034	5.56
CatBoost	0.846171	0.772529	-8.70
LightGBM	0.930109	0.893892	-3.89

The mixed results highlight the importance of understanding the nature of the data and the characteristics of the machine learning models used. Models that are highly sensitive to outliers, such as SVM, tend to benefit from their removal, while models that leverage diverse data points, including outliers, may suffer from reduced accuracy. According to these results it can be seen that to handle outliers effectively can vary, depending on the model type and application context. What should be remembered is that although removing outliers often improves models that are prone to overfitting (Ayiah-Mensah et al., 2025), it can also have adverse effects on models that rely on data variability for robust learning. According to the analysis's findings, a number of models, including Neural Network and CatBoost, saw a drop in performance following tuning. This is most likely because of overfitting or the loss of crucial data brought on by outlier removal. Several tactics will be used to avoid this in subsequent analyses, such as using more stringent cross-validation methods like k-fold cross-validation to make sure the model doesn't overfit to particular data subsets. Additionally, models that are prone to overfitting will be subjected to early halting and regularization approaches (such L1/L2), which enable training to end before performance deteriorates. Additionally, the hyperparameter tuning strategy will be enhanced through the use of more effective and flexible techniques like Bayesian optimization and randomized search. More consideration will be given to outlier removal, particularly for models that depend on data diversity for efficient learning. Rather than eliminating outliers entirely, data modification or more resilient scaling techniques may be used.

4. Conclusion

The results of the study show that outlier management and tuning techniques have a significant impact on how well machine learning algorithms estimate the compressive strength of concrete. Models like SVM and Neural Network exhibited significant gains after tweaking, underscoring the significance of choosing the correct parameters, whilst CatBoost and Linear Regression models showed optimal performance without the need for additional tuning. However, models that depend on data variability, like Neural Network and CatBoost, suffered while models that are sensitive to extreme data, like SVM and Decision Tree, improved as a result of outlier removal. To prevent overfitting and performance deterioration, a contextual approach is therefore required, including the use of cross-validation, regularization, and adaptive tuning techniques. Practically speaking, these findings greatly increase the effectiveness and precision of concrete planning and quality control procedures in the building sector, solidifying machine learning as a pertinent and trustworthy instrument for data-driven civil engineering decision making.

References

A Ilemobayo, J., Durodola, O., Alade, O., J Awotunde, O., T Olanrewaju, A., Falana, O., Ogungbire, A., Osinuga, A., Ogunbiyi, D., Ifeanyi, A., E Odezuligbo, I., E Edu, O., 2024. Hyperparameter Tuning in Machine Learning: A Comprehensive Review. J. Eng. Res. Rep. 26, 388–395. https://doi.org/10.9734/jerr/2024/v26i61188

Alyami, M., Ullah, I., AlAteah, A.H., Alsubeai, A., Alahmari, T.S., Farooq, F., Alabduljabbar, H., 2025. Machine learning models for predicting the compressive strength of cement-based mortar materials: Hyper tuning and optimization. Structures 71, 107931. https://doi.org/10.1016/j.istruc.2024.107931

Anugerah Simanjuntak, Rosni Lumbantoruan, Kartika Sianipar, Rut Gultom, Mario Simaremare, Samuel Situmeang, Erwin Panggabean, 2024. Research and Analysis of IndoBERT Hyperparameter Tuning in Fake News Detection. J. Nas. Tek. Elektro Dan Teknol. Inf. 13, 60–67. https://doi.org/10.22146/jnteti.v13i1.8532

Anwar, M.K., 2025. A comparative performance analysis of machine learning models for compressive strength prediction in fly ash-based geopolymers concrete using reference data. Case Stud. Constr. Mater.

Asteris, P.G., Skentou, A.D., Bardhan, A., Samui, P., Pilakoutas, K., 2021. Predicting concrete compressive strength using hybrid ensembling of surrogate machine learning models. Cem. Concr. Res. 145, 106449. https://doi.org/10.1016/j.cemconres.2021.106449

Ayiah-Mensah, F., Bosson-Amedenu, S., Baah, E.M., Addor, J.A., 2025. Advancements in seasonal rainfall forecasting: A seasonal auto-regressive integrated moving average model with outlier adjustments for Ghana’s Western Region. Sci. Afr. 28, e02632. https://doi.org/10.1016/j.sciaf.2025.e02632

Chodha, A., 2024. Cement Slump SVR.

Choi, J.-H., Kim, D., Ko, M.-S., Lee, D.-E., Wi, K., Lee, H.-S., 2023. Compressive strength prediction of ternary-blended concrete using deep neural network with tuned hyperparameters. J. Build. Eng. 75, 107004. https://doi.org/10.1016/j.jobe.2023.107004

Dash, P.K., 2023. Influence of chemical constituents of binder and activator in predicting compressive strength of fly ash-based geopolymer concrete using firefly-optimized hybrid ensemble machine learning model. Mater. Today Commun.

Dong, Y., 2025. A new method to evaluate features importance in machine-learning based prediction of concrete compressive strength. J. Build. Eng.

El Hachimi, C., Belaqziz, S., Khabba, S., Daccache, A., Ait Hssaine, B., Karjoun, H., Ouassanouan, Y., Sebbar, B., Kharrou, M.H., Er-Raki, S., Chehbouni, A., 2025. Physics-informed neural networks for enhanced reference evapotranspiration estimation in Morocco: Balancing semi-physical models and deep learning. Chemosphere 374, 144238. https://doi.org/10.1016/j.chemosphere.2025.144238

Fan, C., Ding, Y., Zheng, Y., 2025. Bond strength and failure mode prediction model for recycled aggregate concrete based on intelligent algorithm optimized support vector machine. Structures 71, 107999. https://doi.org/10.1016/j.istruc.2024.107999

Ghafoorian Heidari, S.I., Safehian, M., Moodi, F., Shadroo, S., 2024. Predictive modeling of the long-term effects of combined chemical admixtures on concrete compressive strength using machine learning algorithms. Case Stud. Chem. Environ. Eng. 10, 101008. https://doi.org/10.1016/j.cscee.2024.101008

Gokcesu, K., Neyshabouri, M.M., Gökcesu, H., Kozat, S.S., 2019. Sequential Outlier Detection Based on Incremental Decision Trees. IEEE Trans. SIGNAL Process. 67.

Guzmán-Torres, J.A., Domínguez-Mota, F.J., Alonso-Guzmán, E.M., Tinoco-Guerrero, G., Martínez-Molina, W., 2024. ConcreteXAI: A multivariate dataset for concrete strength prediction via deep-learning-based methods. Data Brief 53, 110218. https://doi.org/10.1016/j.dib.2024.110218

Hassan, B.R., Hamid, F.L., Karim, H., Hussein, A.B., Abdulrahman, P.H., Mohammed, L.M., Hatim, L.A., Mahmood, S.S., 2024. Predicting shear strength of fiber-reinforced concrete beams reinforced with longitudinal FRP bars with machine learning techniques. Aust. J. Struct. Eng. 1–16. https://doi.org/10.1080/13287982.2024.2396215

Huang, P., Dai, K., Yu, X., 2023. Machine learning approach for investigating compressive strength of self-compacting concrete containing supplementary cementitious materials and recycled aggregate. J. Build. Eng. 79, 107904. https://doi.org/10.1016/j.jobe.2023.107904

Jiang, Y., Wang, Y., Zhang, J., Xie, B., Liao, J., Liao, W., 2021. Outlier detection and robust variable selection via the penalized weighted LAD-LASSO method. J. Appl. Stat. 48, 234–246. https://doi.org/10.1080/02664763.2020.1722079

Khan, M.I., 2023. Intelligent data-driven compressive strength prediction and optimization of reactive powder concrete using multiple ensemble-based machine learning approach. Constr. Build. Mater.

Kumar, N., Martin, H., Raubal, M., 2024. Enhancing Deep Learning-Based City-Wide Traffic Prediction Pipelines Through Complexity Analysis. Data Sci. Transp. 6, 24. https://doi.org/10.1007/s42421-024-00109-x

Kumar, R., Althaqafi, E., Patro, S.G.K., Simic, V., Babbar, A., Pamucar, D., Singh, S.K., Verma, A., 2024. Machine and deep learning methods for concrete strength Prediction: A bibliometric and content analysis review of research trends and future directions. Appl. Soft Comput. 164, 111956. https://doi.org/10.1016/j.asoc.2024.111956

Li, Z., Pei, T., Ying, W., Srubar, W.V., Zhang, R., Yoon, J., Ye, H., Dabo, I., Radlińska, A., 2024. Can domain knowledge benefit machine learning for concrete property prediction? J. Am. Ceram. Soc. 107, 1582–1602. https://doi.org/10.1111/jace.19549

Liang, K., Zhao, J., Zhang, Z., Guan, W., Pan, M., Li, M., 2024. Data-driven AI algorithms for construction machinery. Autom. Constr. 167, 105648. https://doi.org/10.1016/j.autcon.2024.105648

Liu, L., Xie, C., Hu, W., Li, Y., 2024. Recursive Elimination of “Outliers” to Get Benchmark Dataset. IEEE Access 12, 98319–98325. https://doi.org/10.1109/ACCESS.2024.3427783

Liu, Y., Yu, H., Guan, T., Chen, P., Ren, B., Guo, Z., 2025. Intelligent prediction of compressive strength of concrete based on CNN-BiLSTM-MA. Case Stud. Constr. Mater. 22, e04486. https://doi.org/10.1016/j.cscm.2025.e04486

Luo, X., Li, Y., Wang, Q., Mu, J., Liu, Y., 2024. Machine learning based modeling for predicting the compressive strength of solid waste material-incorporated Magnesium Phosphate Cement. J. Clean. Prod. 442, 141172. https://doi.org/10.1016/j.jclepro.2024.141172

Lv, Z., Chen, C., Tan, Z., Hu, B., Jin, J., 2025. Establishment of strength evolution model and life prediction for diatomite-modified concrete under freeze-thaw conditions. Constr. Build. Mater. 463, 140049. https://doi.org/10.1016/j.conbuildmat.2025.140049

Mahmood, M.S., Elahi, A., Zaid, O., Alashker, Y., Șerbănoiu, A.A., Grădinaru, C.M., Ullah, K., Ali, T., 2023. Enhancing compressive strength prediction in self-compacting concrete using machine learning and deep learning techniques with incorporation of rice husk ash and marble powder. Case Stud. Constr. Mater. 19, e02557. https://doi.org/10.1016/j.cscm.2023.e02557

Marchiori, R., Song, S., Moon, J., 2025. Developing heat stress training assessments: A training-driven methodology approach to enhance safety in the construction industry. J. Safety Res. 92, 262–271. https://doi.org/10.1016/j.jsr.2024.11.027

Miyan, N., Krishnan, N.M.A., Das, S., 2024. Integrating data imputation and augmentation with interpretable machine learning for efficient strength prediction of fly ash-based alkali-activated concretes. J. Build. Eng. 98, 111248. https://doi.org/10.1016/j.jobe.2024.111248

Moutassem, F., Chidiac, S.E., 2016. Assessment of concrete compressive strength prediction models. KSCE J. Civ. Eng. 20, 343–358. https://doi.org/10.1007/s12205-015-0722-4

Parmo, P., Wardhana, I., 2024. Investigating the Influence of Cement and Additive Normalization on Concrete Compressive Strength: A Statistical Analysis. Comput. Eng. Phys. Model. 7. https://doi.org/10.22115/cepm.2024.484143.1335

Rusman, J., Haryati, B.Z., Michael, A., 2023. Optimisasi Hiperparameter Tuning pada Metode Support Vector Machine untuk Klasifikasi Tingkat Kematangan Buah Kopi. J. Komput. Dan Inform. 11, 195–202. https://doi.org/10.35508/jicon.v11i2.12571

Sathiparan, N., 2025. Predicting compressive strength in cement mortar: The impact of fly ash composition through machine learning. Sustain. Chem. Pharm. 43, 101915. https://doi.org/10.1016/j.scp.2025.101915

Tak, M.S.N., Feng, Y., Mahgoub, M., 2025. Advanced Machine Learning Techniques for Predicting Concrete Compressive Strength.

Talpur, S.A., Thansirichaisree, P., Anotaipaiboon, W., Mohamad, H., Zhou, M., Ejaz, A., Hussain, Q., Saingam, P., Chaimahawan, P., 2025. Data-driven prediction of failure loads in low-cost FRP-confined reinforced concrete beams. Compos. Part C Open Access 17, 100579. https://doi.org/10.1016/j.jcomc.2025.100579

Touati, S., Boumediri, H., Karmi, Y., Chitour, M., Boumediri, K., Zemmouri, A., Moussa, A., Fernandes, F., 2025. Performance analysis of steel W18CR4V grinding using RSM, DNN-GA, KNN, LM, DT, SVM models, and optimization via desirability function and MOGWO. Heliyon 11, e42640. https://doi.org/10.1016/j.heliyon.2025.e42640

Vargas, J.F., Oviedo, A.I., Ortega, N.A., Orozco, E., Gómez, A., Londoño, J.M., 2024. Machine-Learning-Based Predictive Models for Compressive Strength, Flexural Strength, and Slump of Concrete. Appl. Sci. 14, 4426. https://doi.org/10.3390/app14114426

Wang, C., Li, C., Feng, Y., Wang, S., 2025. Predicting hydropower generation: A comparative analysis of Machine learning models and optimization algorithms for enhanced forecasting accuracy and operational efficiency. Ain Shams Eng. J. 16, 103299. https://doi.org/10.1016/j.asej.2025.103299

Yu, J., Chang, X., Hu, S., Yin, H., Wu, J., 2024. Combining travel behavior in metro passenger flow prediction: A smart explainable Stacking-Catboost algorithm. Inf. Process. Manag. 61, 103733. https://doi.org/10.1016/j.ipm.2024.103733