Application of Advance Machine Learning Models on Predicting

Compressive Strength of UHPC

K Akhilendra Sharma1(0009-0002-7100-7392), and Rahul Biswas1(0000-0001-8697-7565)

1Applied Mechanics Department,

Visvesvaraya National Institute Of Technology, Nagpur, India

[akhilendrasharma.er@gmail.com](mailto:akhilendrasharma.er@gmail.com)

[rahulbiswas@apm.vnit.ac.in](mailto:rahulbiswas@apm.vnit.ac.in)

**Abstract.** Ultra-High-Performance Concrete (UHPC) is an advanced material popular for its remarkable strength and durability, which makes it one of the most preferred materials for advanced high strength structural applications. However, its exceptional compressive strength relies heavily on expensive ingredients, such as cement, silica fume, superplasticizers and other carbon intensive ingredients. The objective of this study is to develop robust machine learning models to reduce experimentation time, enhance cost efficiency and minimize waste in construction. By comparing various Machine Learning models, we strive to achieve the most accurate predictions and thus establish a non-linear relationship between mix design and Compressive strength of UHPC. The models evaluated include Gradient Boosting, Random Forest (RF), and Decision Trees (DT). Each model was trained and tested on a comprehensive dataset of 810 test observations with 15 input variables that influence UHPC's compressive strength. Hyperparameters of the Gradient Boosting model were tuned in such a way that it gives the best performance at a minimal computational effort. Models were compared based on performance metrics such as root mean square error (RMSE), and the coefficient of determination (R²). The results demonstrate that ensemble methods, particularly Gradient Boosting and Random Forest, excel in prediction accuracy and robustness. This statistical analysis was substantiated by Taylor diagram for visual understanding. Taylor Diagram provides a quick visual summary, thereby helping in the identification of models with robust predictive performance. Further SHapley Additive exPlanations (SHAP), a model-agnostic interpretability tool based on game theory, was used. SHAP analysis was conducted to establish the feature significance and relationship between input features and compressive strength. Consequently, this understanding of feature significance enables us to optimize the mix to achieve a superior yet sustainable material.

**Keywords:** Random Forest, Gradient Boosting, Decision Tree, Ultra-High Performance Concrete.

# Introduction

Ultra-High Performance Concrete (UHPC) represents a substantial breakthrough in the construction materials sector, with exceptional mechanical behavior such as very high strength under compression and durability against environmental degradation [1]. Consequently, UHPC is opted for a myriad of structural applications that span from high-rise buildings to long-span bridges [1]. Due to the dense packing of particles and low porosity, it has higher mechanical properties of 150Mpa-810Mpa [1 - 3]. However, producing a material with such high compressive strength necessitates a high dosage of cement, low water-binder ratio, supplementary cementitious materials (silica fume, fly ash, ground granulated blast-furnace slag) and steel fibers, which ultimately lead to environmental concerns [3]. Overcoming these challenges is necessary to promote the use of UHPC, and achieving this requires formulating an optimal mix which reduces environmental impact while maintaining the material’s strength. Machine learning algorithms have been proven to predict outcomes precisely, thereby reducing the expenses and time for experimentation. Moreover, ensemble machine learning models such as Random Forest regressor and Gradient Boosting have shown remarkable predictive performance over conventional regression models, especially in handling imbalanced and noisy datasets [4, 5].

In this study we performed hyperparameter tuning of Gradient Boosting model to enhance the performance of the model. Also, we performed feature selection by removing weak features. This approach enables us to enhance the predictability of the models and understand the concrete behavior in a better way. In the field of civil engineering and concrete technology, many machine learning studies have overlooked how different features affect compressive strength. While these studies focus on predicting compressive strength, the specific impact of each feature on compressive strength has been insufficiently explored. This gap highlights the need for practical explanation of the underlying logic behind the predictions and their relationship with the features, such that it can be related to real world scenarios and fundamentals of concrete technology. We have conducted SHapley Additive exPlanations (SHAP) analysis to explain the feature importance and their individual contribution to the compressive strength of UHPC.

# Methodology

## Data Collection

A total of 810 experimental datasets were gathered from published literature [6] to predict and compare the compressive strength of Ultra-High Performance Concrete (UHPC) using three machine learning algorithms: Gradient Boosting, Random Forest, and Decision Trees.

The dataset included in the literature have 15 input features, such as Cement (C), Silica Fume (SF), Slag (S), Fly Ash (FA), Quartz Powder (QP), Nano Silica (NS), Limestone Powder (LP), Water (W), Coarse Aggregate (Gravel), Fine Aggregate (Sand), Fiber (Fi), Superplasticizer (SP), Temperature (T), Relative Humidity (RH), and Age.The statistical properties of the input and output features of the dataset are mentioned in Table 1, which gives us a clear understanding of the range of the parameters.

## Feature Selection and Data Pre-processing

Pearson correlation analysis, as shown in Fig. 2a, was conducted to explore the degree of correlation (DOC) between these features and the compressive strength (f), which led to following observations:

1. Compared to other parameters, cement (C), silica fume (SF), fiber (Fi), and age showed a considerably higher correlation with (f).

2. Features like fly ash, nano silica, sand, gravel, and relative humidity displayed a negative correlation with (f).

To improve model performance, features with correlations below 0.2 were excluded, as shown in Fig. 2b. As a result, the 8 remaining features used for model training included: Cement, Silica Fume, Quartz Powder, Nano Silica, Sand, Fiber, Superplasticizer, and Age. The data was also checked for missing values and outliers, finding no missing values and only a few outliers that required no significant adjustments. The final dataset was split into training and testing sets, with 80% used for training and 20% for testing.

**Table 1.** Properties of Dataset features

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Mean | SD | Min | Max |
| C(Kg/m3) | 737.914642 | 173.457225 | 270 | 1251.2 |
| S(Kg/m3) | 25.194568 | 74.365464 | 0 | 375 |
| SF(Kg/m3) | 136.98721 | 104.144596 | 0 | 433.7 |
| LP(Kg/m3) | 41.929506 | 133.131451 | 0 | 1058.2 |
| QP(Kg/m3) | 33.270988 | 79.673876 | 0 | 397 |
| FA(Kg/m3) | 26.264938 | 67.461703 | 0 | 356 |
| NS(Kg/m3) | 3.638642 | 7.775957 | 0 | 47.5 |
| W(Kg/m3) | 179.891136 | 25.568235 | 90 | 272.6 |
| Sand (Kg/m3) | 995.328519 | 283.268562 | 0 | 1502.8 |
| Gravel (Kg/m3) | 154.781481 | 357.569121 | 0 | 1195 |
| Fi (Kg/m3) | 56.044444 | 75.230584 | 0 | 234 |
| SP(Kg/m3) | 30.030926 | 13.993525 | 1.1 | 57 |
| RH (%) | 97.985185 | 7.808415 | 50 | 100 |
| T(oC) | 23.871605 | 16.253806 | 0 | 210 |
| Age(days) | 37.093827 | 53.115855 | 1 | 365 |
| f (MPa) | 123.131496 | 40.238665 | 28.51 | 220.5 |

A screenshot of a grid

Description automatically generated

**(a)**

A screenshot of a grid

Description automatically generated

**(b)**

**Fig. 1.** Pearson Correlation Matrix between input and target variables. **(a)** Input Features from Literature. **(b)** After excluding weak features

## Model Training

Decision Tree regressor uses a flowchart like structure to recursively split the data into subsets, known as branches, to predict an output. This simple machine learning model is prone to overfitting on complex data [7 - 9]. To address this limitation, we used a bagging ensemble method, Random Forest. In this model combination of multiple decision trees is to enhance prediction accuracy and reduce overfitting. Each tree in the Random Forest is trained on a bootstrapped sample of the data, with random feature selections at each node. The final prediction is an average of the predictions from all trees, providing a more robust and generalizable model [7 - 10]. Even though Random Forests (RF) are simple and have low computational costs, they often perform surprisingly well in most regression problems [6].

Gradient Boosting is a boosting ensemble technique. This model combines numerous small decision trees but in an intelligent sequential process. Each new model corrects the errors of the previous ones by focusing on the residuals or errors left by the prior models. This iterative approach improves the model's performance by refining its predictions over multiple stages [10].

Hyperparameter tuning was carried out by trial-and-error method, and on gradual increase in n\_estimators and learning\_rate, an improvement in performance was observed in Gradient Boosting model. Following a certain point, the improvement in performance appeared to plateau, hence finally n\_estimators was set to a value of 10000 and learning\_rate as 0.3.

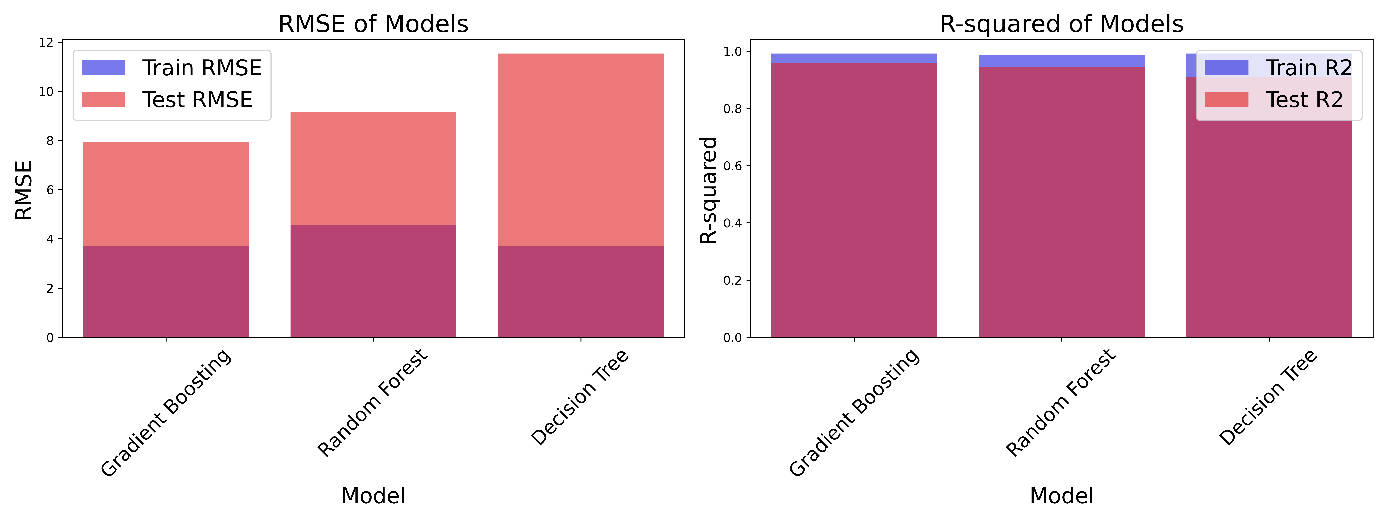
# Results and Discussions

## Hyperparameter settings and Performance Metrics

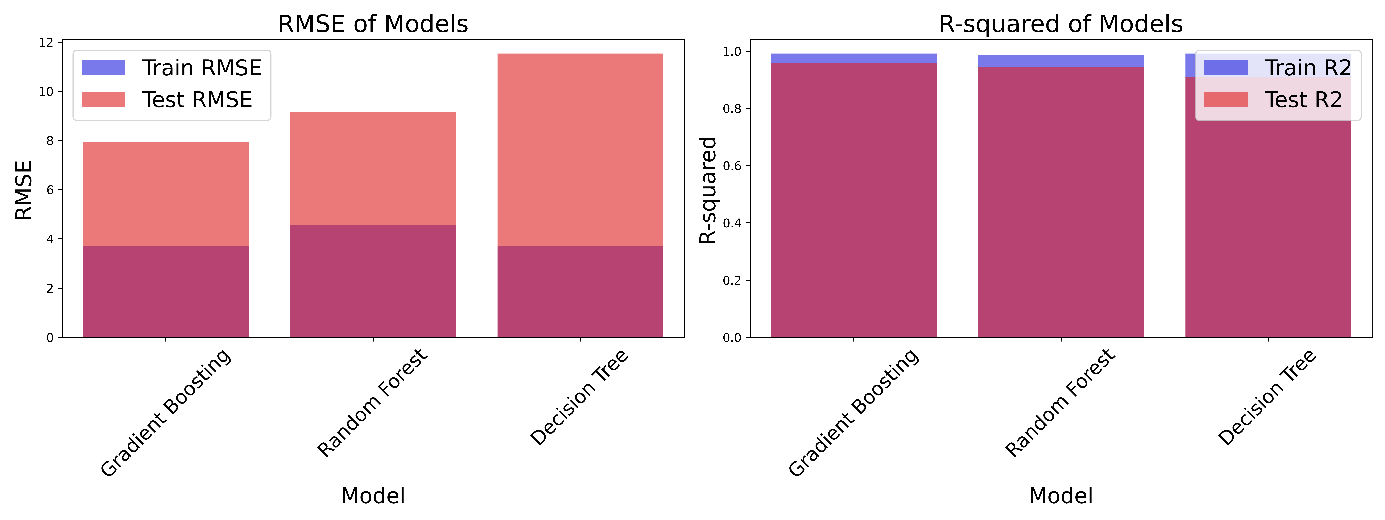
This section presents how each of the models have performed based on their specific hyperparameter configurations, with results presented in terms of the performance metrics - RMSE and R2 values [7 – 9]. With default settings of hyperparameters Decision Tree Regressor was trained and considered as the baseline for comparing the performance of more complex models such as ensemble models. The RMSE and R2 values of DT model for training stage are 3.6983 and 0.9916 respectively. But for the testing stage, RMSE = 11.5353 and R2 = 0.9108, which indicates that the model exhibits overfitting issue. To counter this, we have employed ensemble machine learning methods. The bagging ensemble method, Random Forest, outperformed Decision Tree in testing stage, with obtained RMSE = 4.5798 and R2 = 0.9872 in the training stage and for testing stage, RMSE = 9.1467 and R2 = 0.9439. However, Gradient Boosting emerged as the top-performing model when tuned to hyperparameter settings: n\_estimators = 10000 and random\_state = 42 and learning\_rate=0.3. And the improvement can be attributed to the sequential learning method that is adopted in the model. During the training stage the Gradient Boosting model achieved an RMSE of 3.698 and R2 of 0.9916 and for testing stage RMSE of 7.93 and R2 of 0.96. As shown in Fig. 3a, b, c, the scatter plot for regression shows how the predicted values deviate from the actual values for their respective models. The ranking of the models based on their performance in testing data is presented in Table 2 and depicted visually in Fig 2.

**Table 2.** Ranking of models based on their performance

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Rank | Model | RMSE(Train) | RMSE(Test) | R2(Train) | R2(Test) |
| 1 | Gradient Boosting | 3.698 | 7.93 | 0.9916 | 0.96 |
| 2 | Random Forest | 4.5798 | 9.146 | 0.9872 | 0.9439 |
| 3 | Decision Tree | 3.6983 | 11.5353 | 0.99168 | 0.9108 |

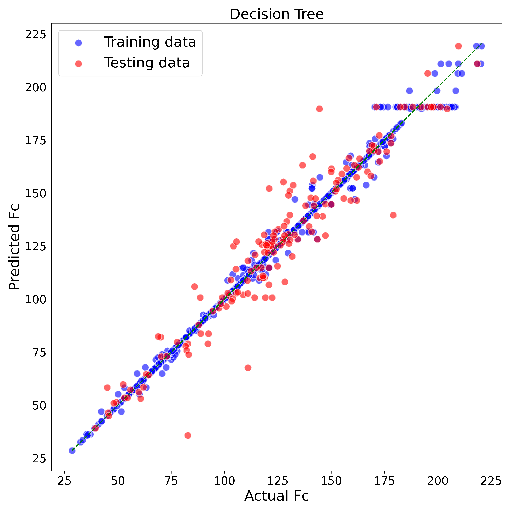


**(a)**

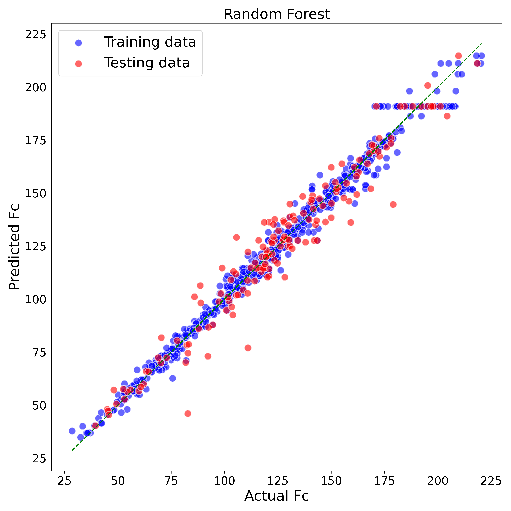


(**b)**

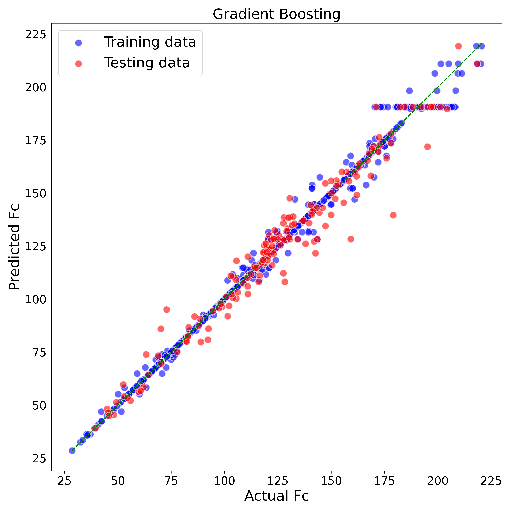
**Fig. 2.** Comparison of models’ performance based on **(a)** RMSE values, **(b)** R2 values



**(a)**



**(b)**



**(c)**

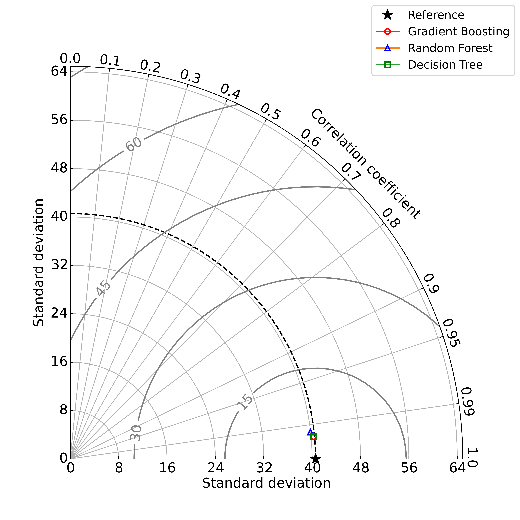
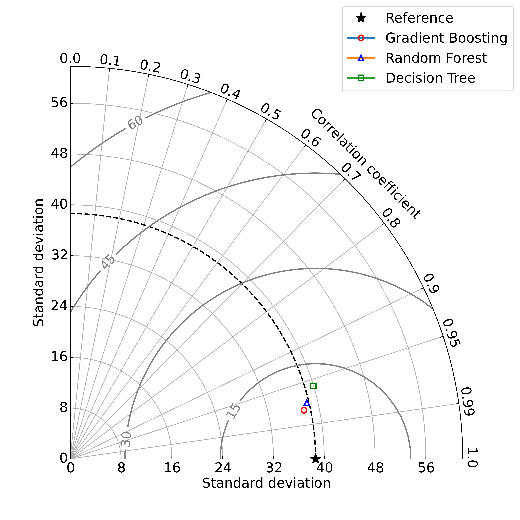
**Fig. 3.** Scatter plot of models **(a)** Decision Tree, **(b)** Random Forest, **(c)** Gradient Boosting

## Taylor Diagram

A Taylor diagram uses three statistical metrics, RMSE, standard deviation and correlation coefficient. Radial distance from origin shows standard deviation, angle between axes represents correlation coefficient and distance from reference point denotes the RMSE. By plotting multiple models on a single Taylor diagram, one can easily identify which model most closely resembles the reference dataset. Models that lie closer to the reference point have higher correlation, lower RMSE, and a standard deviation similar to the reference, indicating better performance. This makes the Taylor diagram a powerful tool for comparing the accuracy of different predictive models in a visually intuitive way [11]. As observed in Fig. 4a, in the training phase, Decision Tree and Gradient Boosting are coinciding with each other and are closer to the reference point as compared to Random Forest. However, in the testing phase Gradient boosting is closest to the reference point followed by Random Forest and Decision Tree. Hence, the graphical summary by Taylor diagram reinforces the statistical results which favored the robustness of Gradient Boosting model.

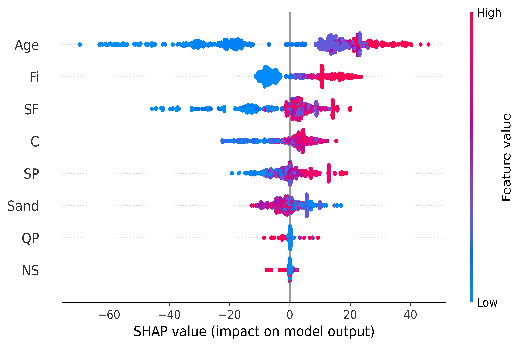
## SHAP Analysis

SHapley Additive exPlanations(SHAP) assigns importance value to each feature for a particular prediction [12, 13]. SHAP uses game theory principles to determine feature importance through a two-step process: first, it trains a classification model with all the features during the initial interaction. Then, it calculates SHAP values for each feature and ranks them to identify which features are most influential in predicting the target outcome. As shown in Fig. 4b, mean absolute SHAP values of features which show higher influence on the compressive strength of UHPC are: Age: 20.555417, Fi: 9.731076, SF: 7.413811, C: 5.298315,SP: 4.404854, Sand: 3.549891.The Age feature has exhibited the highest Mean Absolute SHAP value, whereas, QP (SHAP value = 0.691695) and NS (SHAP value = 0.364115) have the lowest Mean Absolute SHAP values among the primary features analyzed. However, from Fig. 4a, it can be observed that as compared to other features, Sand, when decreased, results in an increase in compressive strength. This can be attributed to the improved particle packing and better chances for supplementary cementitious materials like Silica fume and Nano silica to participate in the pozzolanic reactions. But on further reduction of sand after a certain limit, there could be a potential reduction of compressive strength due to issues like high brittleness. Recognizing the significance of every characteristic enables us to improve the model and make better and practical changes to the UHPC mix design and thus achieve a superior yet sustainable material.

1. **(b)**

**Fig. 4.** Taylor Diagrams **(a)** For Training stage(**b)** For Testing stage

 A graph of a bar chart

Description automatically generated with medium confidence

1. **(b)**

**Fig. 5.** SHAP Analysis **(a)** SHAP global explanation for XGBoost model

(**b)** Mean Absolute SHAP values for XGBoost Model

# Conclusions

This study has evaluated the performance of three machine learning models in predicting the compressive strength of Ultra-High-Performance Concrete (UHPC) through performance assessment based on the metrics RMSE and R² values. The models analyzed include Decision Tree, Random Forest and Gradient Boosting. The Decision Tree model established a baseline performance with an RMSE of 11.5353 and an R² of 0.9108 in the testing stage, providing a comparative reference for the more robust models. The Random Forest model demonstrated superior performance over the Decision Tree, achieving an RMSE of 9.1467 and an R² of 0.9439 in testing,

owing to its ensemble nature that mitigates overfitting. The Gradient Boosting model emerged as the best performing model, with an RMSE of 10.4221 and an R² of 0.9272 in the testing stage, attributed to its sequential learning approach. Following this predictive modelling an explanatory feature significance analysis was performed using SHAP, which indicated the impact of each characteristic on compressive strength of UHPC. These observations can be used to optimize the mix design to reduce the negative environmental impact of the material.

In summary, the findings highlight the Gradient Boosting model as the most proficient in predicting UHPC compressive strength, followed by Random Forest. The results underscore the potential of hyperparameter optimization of these machine learning models in enhancing the accuracy of UHPC property predictions. Future work should focus on further optimization of Gradient Boosting and Random Forest models and exploration of hybrid models to combine the advantages of different approaches. Additionally, expanding the dataset may provide further improvements in predictive performance.

# References

1. Vernet, C. P. (2004). Ultra-Durable concretes: structure at the micro- and nanoscale. MRS Bulletin, 29(5), 324–327. <https://doi.org/10.1557/mrs2004.98>
2. Richard, P., & Cheyrezy, M. (1995). Composition of reactive powder concretes. Cement and Concrete Research, 25(7), 1501–1511. <https://doi.org/10.1016/0008-8846(95)00144-2>
3. Bajaber, M., & Hakeem, I. (2021). UHPC evolution, development, and utilization in construction: a review. Journal of Materials Research and Technology, 10, 1058–1074. <https://doi.org/10.1016/j.jmrt.2020.12.051>
4. Mienye, I. D., & Sun, Y. (2022). A survey of ensemble learning: Concepts, algorithms, applications, and prospects. IEEE Access, 10, 99129–99149. <https://doi.org/10.1109/access.2022.3207287>
5. Dong, X., Yu, Z., Cao, W., Shi, Y., & Ma, Q. (2019). A survey on ensemble learning. Frontiers of Computer Science, 14(2), 241–258. <https://doi.org/10.1007/s11704-019-8208-z>
6. Marani, A., Jamali, A., & Nehdi, M. L. (2020b). Predicting Ultra-High-Performance concrete compressive strength using tabular generative adversarial networks. Materials, 13(21), 4757. <https://doi.org/10.3390/ma13214757>
7. Zhou, Z. (2021c). Machine learning. In Springer eBooks. <https://doi.org/10.1007/978-981-15-1967-3>
8. Rebala, G., Ravi, A., & Churiwala, S. (2019b). An introduction to machine learning. In Springer eBooks. <https://doi.org/10.1007/978-3-030-15729-6>
9. Kubat, M. (2017b). An introduction to machine learning. In Springer eBooks. <https://doi.org/10.1007/978-3-319-63913-0>
10. Korstanje, J. (2021c). Advanced Forecasting with Python. In Apress eBooks. <https://doi.org/10.1007/978-1-4842-7150-6>
11. Taylor, K. E. (2001). Summarizing multiple aspects of model performance in a single diagram. Journal of Geophysical Research Atmospheres, 106(D7), 7183–7192. <https://doi.org/10.1029/2000jd900719>
12. Lundberg, S. M., & Lee, S. (2017). A unified approach to interpreting model predictions. arXiv (Cornell University). <https://doi.org/10.48550/arxiv.1705.07874>.
13. Wang, H., Liang, Q., Hancock, J. T., & Khoshgoftaar, T. M. (2024). Feature selection strategies: a comparative analysis of SHAP-value and importance-based methods. Journal of Big Data, 11(1). <https://doi.org/10.1186/s40537-024-00905-w>