



## Comparison of Predictive Modeling Concrete Compressive Strength with Machine Learning Approaches

G. Airlangga

Information System Study Program, Faculty of Engineering, Atma Jaya Catholic University of Indonesia, South Jakarta, Indonesia

Email: [gregorius.airlangga@atmajaya.ac.id](mailto:gregorius.airlangga@atmajaya.ac.id)

### ARTICLE INFO

#### Article History :

Article entry : 24 – 03 – 2024  
Article revised : 27 – 03 – 2024  
Article received : 29 – 04 – 2024

#### Keywords :

Concrete Compressive Strength, Machine Learning, Prediction, SHAP Values.

**IEEE Style in citing this article :**  
G. Airlangga, "Comparison of Predictive Modeling Concrete Compressive Strength with Machine Learning Approaches," *U Karst*, vol. 8, no. 1, pp 28 – 41, 2023, doi: 10.30737/ukarst.v8i1.5532.

### ABSTRACT

Accurately predicting concrete compressive strength is fundamental for optimizing mix designs, ensuring structural integrity, and advancing sustainable construction practices. Increased demands for safer, more durable infrastructure necessitate effective predictive concrete compressive strength models. This research aims to compare the effectiveness of six machine learning models such as Linear Regression, Random Forest, Support Vector Regression (SVR), K-Nearest Neighbors (KNN), Gradient Boosting, and XGBoost to predict concrete compressive strength. Used a dataset of 1030 instances with varying mixture compositions, conducted extensive exploratory data analysis, and applied feature engineering and data scaling to enhance model performance. Assessments were performed with a 5-fold cross-validation approach with the R-squared ( $R^2$ ) metric. In addition, the SHAP value is used to understand the influence of each feature on the compressive strength results. The results revealed that XGBoost significantly outperformed other models, achieving an average  $R^2$  value of 0.9178 with a standard deviation of 0.0296. Notably, Random Forest and Gradient Boosting also demonstrated robust capabilities. Based on our experiment, these models effectively predicted compressive strengths close to actual measured values, confirming their practical applicability in civil engineering. SHAP values provided insights into the significant impact of age and cement quantity on model outputs. These results highlight the advanced ensemble methods' effectiveness in concrete compressive strength prediction and underscore the importance of feature engineering in enhancing model accuracy.

## 1. Introduction

Concrete is a quintessential material in civil engineering, holds the civilization's infrastructure in its robust, unyielding grip [1]–[3]. The spectrum of its applications, spanning

bridges, skyscrapers, tunnels, and highways, underscores its pivotal role in development [4]–[6]. However, its compressive strength is the core attribute underpinning its universal application [7]–[9]. The concrete compressive strength, a paramount indicator of the structural integrity and longevity of constructions, is a non-linear function of its age and composition [10]–[12]. This intricate dependency on various factors, including cement, blast furnace slag, fly ash, water, superplasticizer, coarse aggregate, and fine aggregate, alongside the curing period, delineates the complexity of predicting concrete's compressive strength [13]. The urgency for accurate predictions stems from the escalating demand for sustainable construction practices [14]. Modern engineering projects necessitate materials that exhibit superior performance, are resource-efficient and are environmentally friendly. Optimizing concrete mixes to achieve desired strength outcomes without excess material use becomes imperative [15]. Thus, predicting concrete strength accurately is not merely an academic exercise but a practical necessity that echoes the broader call for sustainable development in the construction industry.

Historically, the quest to predict concrete compressive strength has traversed through empirical models based on laboratory experiments to sophisticated statistical analyses [16]. Initial efforts heavily relied on trial and error, with mixed designs adjusted based on iterative testing. However, this foundational approach proved time-consuming and resource-intensive [17]–[19]. The advent of computational models marked a paradigm shift, allowing for the exploration of complex relationships between mix ingredients and compressive strength [20]–[22]. The literature abounds with studies employing various statistical and machine-learning methods to model concrete strength [23]. The body of research reflects a vibrant exploration of computational approaches, from regression analyses to more complex algorithms like artificial neural networks, support vector machines, and decision trees [24]. Yet, despite these advancements, challenges persist. The inherent variability in raw materials, environmental conditions, and mixing processes introduces significant unpredictability, making model accuracy and generalizability ongoing concerns [25].

The sophistication in concrete compressive strength prediction increasingly gravitates towards ensemble methods and advanced machine learning techniques [26]. By leveraging the strengths of multiple predictive models, various approaches aim to enhance accuracy and robustness [27]. Furthermore, the integration of domain knowledge into computational models represents a promising frontier, offering the potential to refine predictions by incorporating insights into the chemical and physical interactions within the concrete mix [26], [28], [29].

This research seeks to contribute to the evolving landscape of concrete strength prediction by addressing identified gaps in the literature. Specifically, while previous studies have demonstrated the efficacy of various predictive models, there remains a need for comprehensive comparative analyses that elucidate the relative performance of these models across diverse datasets [30], [31]. Moreover, the potential of feature engineering and advanced scaling techniques to improve model performance has not been fully explored. Recognizing these gaps, our study addresses critical gaps in predicting concrete compressive strength by focusing on the comparative effectiveness of various machine learning models and their optimization through feature engineering and data scaling.

The research aims to evaluate the predictive accuracy of models such as linear regression, random forest, support vector regression, K-nearest neighbors, gradient boosting, and XGBoost. So, how well these models predict concrete strength and determine the impact of sophisticated feature engineering and scaling techniques on their performance can be ascertained. This is crucial for developing more accurate and adaptable models in civil engineering, ensuring the structural integrity and sustainability of concrete constructions. The comparative results of six predictive models will explain their relative strengths and weaknesses in estimating concrete compressive strength, thereby assisting practitioners and researchers in selecting the most suitable models for their specific needs. Additionally, it will show how feature engineering and data scaling can significantly improve model accuracy, providing insights that guide the enhancement of predictive methods. This dual approach advances our understanding of model capabilities in a practical engineering context and promotes the refinement of methodologies for greater generalizability and precision in material science predictions.

## 2. Research Method

### 2.1 Research Design

This study employs a quantitative research design focusing on predicting concrete compressive strength using various statistical and machine-learning models. The quantitative approach is chosen due to the nature of the problem, which involves analyzing numerical data related to concrete mixtures and their compressive strength. The research design encompasses data collection from existing datasets, preprocessing of data, feature engineering, application of multiple predictive models, and evaluation of these models based on their performance metrics. The models included in the study are Linear Regression, Random Forest, Support

Vector Regression (SVR), K-nearest neighbors (KNN), Gradient Boosting, and XGBoost. This comparative approach allows for a robust analysis of each model's efficacy in predicting the target variable.

## 2.2 Population and Samples

The population of interest in this study consists of various concrete mixtures, each defined by a specific combination of ingredients. The dataset can be downloaded from <https://www.kaggle.com/datasets/vinayakshanawad/cement-manufacturing-concrete-dataset> [32], and the dataset comprises 1030 instances of concrete mixtures, including eight quantitative input variables: Cement, Blast Furnace Slag, Fly Ash, Water, Superplasticizer, Coarse Aggregate, Fine Aggregate, and Age. These variables have been meticulously chosen due to their pivotal roles in influencing the concrete's compressive strength, a crucial determinant of the material's applicability in diverse construction scenarios. Cement is selected as it acts as the primary binder in concrete, directly affecting the mixture's strength and longevity. The inclusion of Blast Furnace Slag and Fly Ash as supplementary cementitious materials is justified by their ability to enhance the mechanical properties of concrete and its sustainability credentials by reducing cement necessity. Water, indispensable for the hydration reaction with cement, is measured precisely, as its excess can dilute the concrete mix, adversely impacting structural integrity. The superplasticizer is considered for its capacity to enhance mix workability without increasing water content, thus allowing for higher compressive strength achievements. The choice of coarse and fine aggregate stems from their role as structural fillers, which contribute to the mix's volume, affecting its density and strength. Finally, age is accounted for due to the significant maturation of concrete strength over time, particularly in the initial curing phases, making it a critical factor in the mix's performance evaluation. The output variable, concrete compressive strength measured in MPa, serves as a quantifiable metric for assessing the efficacy of the mix design. This dataset, therefore, stands as a comprehensive representation of the broader population of concrete mixtures used in civil engineering.

## 2.3 Instruments

Several software tools and libraries are utilized for data analysis, including Python (for data preprocessing and analysis), Scikit-learn and XGBoost for implementing the machine learning models. The SHAP library is used for model interpretability, specifically for understanding the impact of each feature on the predictions of the Random Forest model.

## 2.4 Data Analysis

The analysis begins with an exploratory data analysis (EDA) to understand the dataset's characteristics, including the distribution of variables, presence of outliers, and potential correlations between features. Following EDA, feature engineering is applied to create new features that might improve model performance. Data scaling is performed using `StandardScaler` and `MinMaxScaler` to normalize the feature values, facilitating more efficient learning by the models. For model evaluation, the study employs a cross-validation approach, specifically 5-fold cross-validation, to ensure the models' performance is not dependent on a particular data partitioning. This approach provides a more generalized performance metric across different subsets of the data. The main performance metric used for model evaluation is the R-squared ( $R^2$ ) value, which measures the proportion of variance in the compressive strength that is predictable from the input features. The methodology culminates in the application of `RandomizedSearchCV` for hyperparameter tuning of the Random Forest model, aiming to optimize its performance. Finally, SHAP values are calculated for the Random Forest model to interpret the model's predictions and understand the influence of each feature on the compressive strength outcome.

## 3. Results and Discussions

The R-Square results indicate a notable variance in model performance. XGBoost emerged as the front-runner, demonstrating an exemplary average R-squared value of 0.9178, which suggests a high predictive capability. This model's performance is particularly noteworthy compared to more traditional models, such as Linear Regression, which exhibited an average R-squared value of 0.5886, indicating a moderate fit. Gradient Boosting also performed robustly, with an R-squared value that signals strong predictive power. The performance spectrum of these models underscores the advanced capabilities of ensemble learning techniques in capturing complex, non-linear relationships within data.

A deeper exploration into model interpretability is afforded by SHAP values, which reveal the impact of each feature on the model output. For instance, age and cement content emerge as significant predictors, reflecting their crucial role in determining concrete strength. The distribution of SHAP values for these features indicates their varied influence across different observations, providing insights that could guide more nuanced concrete mix designs. Expanding upon these findings, the discussion integrates them into the broader landscape of civil engineering by exploring their practical implications. The significance of age and cement

content in predicting concrete strength underscores the necessity for precise control over these variables in mix design and quality assurance processes. Such insights can help engineers optimize concrete formulations for specific performance requirements, including durability, sustainability, and cost-efficiency.

Furthermore, understanding the variability in feature influences aids in developing more robust construction standards and can contribute to the ongoing advancement of predictive modeling applications in civil engineering projects. This integration of machine learning insights into practical engineering practices not only enhances the reliability of construction materials but also supports the innovation of building techniques that respond to environmental and economic pressures. It reflects on how integrating machine learning models, particularly those exhibiting high predictive accuracy, can revolutionize material optimization in concrete manufacturing. By doing so, it extends beyond the numbers to probe the potential real-world impact of these findings, such as reducing material waste and enhancing the durability of structures.

**Table 1.** Comparison Results of Machine Learning Methods.

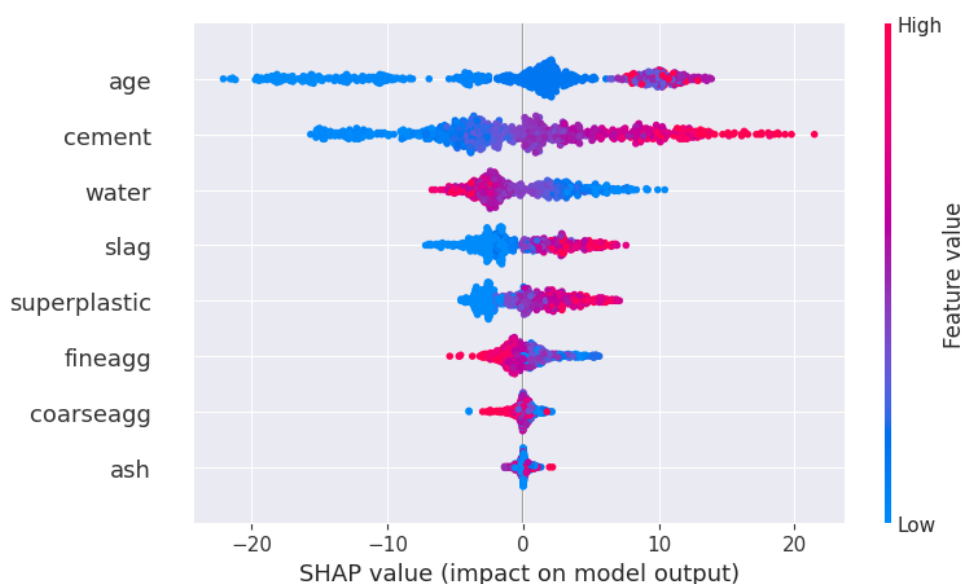
Methods	Average R <sup>2</sup>
Linear Regression	0.5885913293726979 +/- 0.048832066348542244
Random Forest	0.8985148680592616 +/- 0.02471727390076229
SVR	0.6108332127739228 +/- 0.031633314002622946
KNN	0.672174941010522 +/- 0.046894492260017094
Gradient Boosting	0.8928078766409893 +/- 0.018404399492875638
XGBoost	0.9177996618927435 +/- 0.029649696040465096

Source: Author's Analysis Results (2024).

As presented in the **Table 1**, Linear Regression presented an average R<sup>2</sup> value of approximately 0.59 with a standard deviation of 0.049, indicating a moderate fit to the data. However, the spread of residuals in the residual plot demonstrates that the model has limitations in capturing the complexity of the data, as evidenced by the heteroscedastic pattern – the variance of residuals is not constant across the range of fitted values. Furthermore, Random Forest and Gradient Boosting models performed significantly better, with average R<sup>2</sup> values of approximately 0.90 and 0.89, respectively. The low standard deviation for these models

suggests a consistent performance across different folds in cross-validation. By integrating multiple decision trees, these ensemble learning models effectively captured the non-linear relationships in the data.

The Support Vector Regression (SVR) and K-Nearest Neighbors (KNN) models yielded average  $R^2$  values of 0.61 and 0.67, respectively. While outperforming the Linear Regression model, these methods did not achieve the high predictive accuracies of the ensemble models. The SVR model's performance indicates a limited capacity to manage the multi-dimensional feature space, and KNN's performance suggests sensitivity to the feature space's dimensionality and the scale of the data. The standout performer was the XGBoost model, with an impressive average  $R^2$  value of approximately 0.92, which signifies a superior fit to the dataset compared to the other models. The model's ability to leverage gradient boosting and tree pruning mechanisms, along with sophisticated handling of missing values and regularization to avoid overfitting, contributed to its high performance.



Source: Author's Analysis Results (2024).

**Figure 1.** SHAP Results.

As presented in **Figure 1**, the SHAP (SHapley Additive exPlanations) summary plot provides insights into the feature importance and impact on the model output. In our study, the SHAP summary plot for the Random Forest model revealed that age and cement are the most influential factors in predicting concrete compressive strength. Higher SHAP values for age suggest that longer curing times have a strong positive effect on strength, aligning with the understanding that concrete continues to gain strength over time. Similarly, higher cement content is positively correlated with strength, which is consistent with the known properties of concrete mixtures. Water, slag, and superplasticizer features also show varied impacts on the

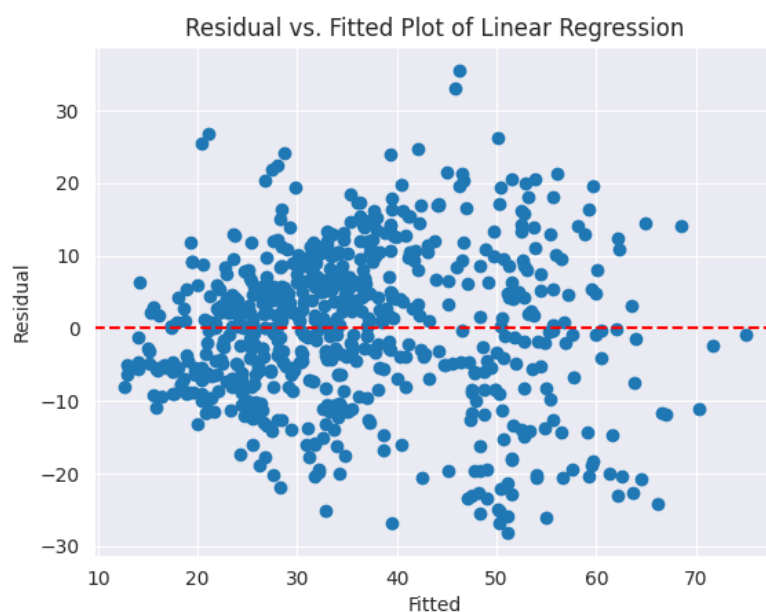
model output, indicating their roles in the concrete's strength development. Notably, when present in lower amounts (indicated by low feature values), the superplasticizer seems to have a more negative impact on strength, likely due to its role in reducing water content and thus affecting the hydration process.

This divergence in model efficacy invites a deeper inquiry into the underlying dynamics of concrete strength prediction. For instance, SHAP values have illuminated the paramount importance of age and cement content as predictors, reinforcing the critical influence of these variables on concrete strength. Such insights are consistent with fundamental concrete science principles and align with findings from recent studies. For example, research by Dixit et al. demonstrated the pronounced impact of curing age on strength development, a phenomenon our study corroborates through quantitative analysis [33]. Similarly, the emphasis on cement content aligns with Arifuzzaman et al. work, highlighting the material's central role in strength attainment [34]. The discussion of these results within the broader landscape of civil engineering and material science not only underscores the potential of advanced machine learning models in material optimization but also delineates a practical pathway for applying these models. The insights gained from this study can directly inform the development and refinement of concrete mix designs to enhance the sustainability and durability of concrete infrastructures. By integrating predictive modeling techniques into the mix design process, engineers can achieve more precise control over material properties, leading to cost-effective and environmentally friendly constructions. This approach enables the proactive adjustment of mixture components based on predictive outcomes, ensuring that each formulation meets specific performance standards while minimizing waste and maximizing resource efficiency. It echoes the sentiments of pioneering research in the field, as the work of previous research has advocated adopting machine learning techniques in optimizing concrete mix designs [35]–[37].

Moreover, the comparative analysis of our results with those of previous studies elucidates a significant advancement in predictive model development. While Linear Regression has been a staple in earlier research due to its simplicity and interpretability, our findings suggest a shift towards more sophisticated models like XGBoost and Random Forest for higher accuracy in complex scenarios. This transition mirrors the evolving landscape of construction material research, where the demand for precision and efficiency necessitates a departure from traditional methodologies. Our investigation's reliance on ensemble models and the nuanced understanding facilitated by SHAP analysis not only enriches the existing body of knowledge but also offers a practical guide for future concrete mix optimizations. The clear



depiction of variable impacts through SHAP values provides a roadmap for adjusting mix components to achieve desired strength levels is a strategy for the efficacy of machine learning in construction material innovation [38].



Source: Author's Analysis Results (2024).

**Figure 2.** Residual Versus Fitted Plot.

**Figure 2** shows that the results emphasize the capability of ensemble models and advanced tree-based algorithms in predicting complex, non-linear relationships inherent in construction materials data. The superior performance of Random Forest and XGBoost over traditional Linear Regression and simpler machine learning models underscores the importance of model selection in predictive analytics for material science. The higher standard deviation in the  $R^2$  values of the Linear Regression model points to potential overfitting or underfitting issues. It could also indicate that the assumptions of Linear Regression, such as homoscedasticity and linearity, might not hold for this dataset.

#### 4. Conclusion

The study's findings emphasize the complexity inherent in predicting concrete compressive strength and the potential of machine learning models to capture this complexity. The superior performance of ensemble methods, particularly XGBoost with an average  $R^2$  value of 0.9178, suggests that these models are more adept at handling the nonlinearity and high dimensionality of the problem space. Additionally, the feature impact analysis via SHAP values underscores the importance of both quantitative mix components and curing time, providing valuable insights for practitioners seeking to optimize concrete formulations for desired

strength outcomes. The research contributes to the field by demonstrating the applicability and effectiveness of various predictive models, and it also highlights the critical role of feature engineering and model interpretability in understanding and improving predictions. Future work could explore further optimization of model parameters, integration of additional features that may influence concrete strength, and the application of these findings in practical mix design and quality control processes.

## 5. Acknowledgements

We express our deepest gratitude to Atma Jaya Catholic University of Indonesia for the unwavering support and resources throughout this research project. Our sincere thanks also extend to the Information System Study Program, whose rigorous academic environment and collaborative ethos have facilitated our work.

## References

- [1] B. Calder, *Architecture: from prehistory to climate emergency*. Penguin UK, 2021.
- [2] V. Nguyen, "Urban Interrupted: an Ethnography of Urbanization and Its Others in Contemporary Beijing," The University of Chicago, 2020.
- [3] S. Timcke, *Algorithms and the end of politics: How technology shapes 21st-century American life*. Bristol University Press, 2021.
- [4] H. Kim, N. Lamichhane, C. Kim, and R. Shrestha, "Innovations in Building Diagnostics and Condition Monitoring: A Comprehensive Review of Infrared Thermography Applications," *Buildings*, vol. 13, no. 11, p. 2829, 2023.
- [5] Q. Zaheer, S. Qiu, S. M. A. Hassan Shah, C. Ai, and J. Wang, "A Novel Framework for Boundary-Preserving Semantic Segmentation of Cracks Using Multisource Visual Fusion (Mvf) and Attention-Based Hierarchical Objective Refinement Method," *Available SSRN 4680261*.
- [6] S. A. Moayyedi, H. Rezaei, A. Kalantari, and R. Jankowski, "Effects of Deck-Abutment Pounding on the Seismic Fragility Curves of Box-Girder Highway Bridges," *J. Earthq. Eng.*, pp. 1–30, 2023.
- [7] S. Nunna, A. R. Ravindran, J. Mroszczok, C. Creighton, and R. J. Varley, "A review of the structural factors which control compression in carbon fibres and their composites," *Compos. Struct.*, vol. 303, p. 116293, 2023.
- [8] T. Tsuru *et al.*, "Intrinsic factors responsible for brittle versus ductile nature of refractory high-entropy alloys," *Nat. Commun.*, vol. 15, no. 1, p. 1706, 2024.
- [9] K. Dwivedi, S. Joshi, R. Nair, M. S. Sapre, and V. Jatti, "Optimizing 3D printed diamond lattice structure and investigating the influence of process parameters on their mechanical integrity using nature-inspired machine learning algorithms," *Mater. Today Commun.*, vol. 38, p. 108233, 2024.
- [10] R. Wang, J. Zhang, Y. Lu, and J. Huang, "Towards designing durable sculptural elements: Ensemble learning in predicting compressive strength of fiber-reinforced nano-silica modified concrete," *Buildings*, vol. 14, no. 2, p. 396, 2024.
- [11] M. Shariati *et al.*, "A novel hybrid extreme learning machine--grey wolf optimizer (ELM-GWO) model to predict compressive strength of concrete with partial replacements for cement," *Eng. Comput.*, pp. 1–23, 2022.

- [12] P. Kumar and B. Pratap, "Feature engineering for predicting compressive strength of high-strength concrete with machine learning models," *Asian J. Civ. Eng.*, vol. 25, no. 1, pp. 723–736, 2024.
- [13] M. J. Moradi, M. Khaleghi, J. Salimi, V. Farhangi, and A. M. Ramezani-pour, "Predicting the compressive strength of concrete containing metakaolin with different properties using ANN," *Measurement*, vol. 183, p. 109790, 2021.
- [14] A. Waqar, I. Othman, N. Saad, M. Azab, and A. M. Khan, "BIM in green building: Enhancing sustainability in the small construction project," *Clean. Environ. Syst.*, p. 100149, 2023.
- [15] H. F. Campos, N. S. Klein, and J. Marques Filho, "Proposed mix design method for sustainable high-strength concrete using particle packing optimization," *J. Clean. Prod.*, vol. 265, p. 121907, 2020.
- [16] Y. Wu and Y. Zhou, "Hybrid machine learning model and Shapley additive explanations for compressive strength of sustainable concrete," *Constr. Build. Mater.*, vol. 330, p. 127298, 2022.
- [17] M. Hu *et al.*, "Recent applications of machine learning in alloy design: A review," *Mater. Sci. Eng. R Reports*, vol. 155, p. 100746, 2023.
- [18] N. J. Szymanski, Y. Zeng, H. Huo, C. J. Bartel, H. Kim, and G. Ceder, "Toward autonomous design and synthesis of novel inorganic materials," *Mater. Horizons*, vol. 8, no. 8, pp. 2169–2198, 2021.
- [19] T. Nasrin, F. Pourkamali-Anaraki, and A. M. Peterson, "Application of machine learning in polymer additive manufacturing: A review," *J. Polym. Sci.*, 2023.
- [20] N. Ninduwezuor-Ehiobu *et al.*, "Tracing the evolution of ai and machine learning applications in advancing materials discovery and production processes," *Eng. Sci. & Technol. J.*, vol. 4, no. 3, pp. 66–83, 2023.
- [21] K. Sarkar, A. Shiuly, and K. G. Dhal, "Revolutionizing concrete analysis: An in-depth survey of AI-powered insights with image-centric approaches on comprehensive quality control, advanced crack detection and concrete property exploration," *Constr. Build. Mater.*, vol. 411, p. 134212, 2024.
- [22] H. Jia, G. Qiao, and P. Han, "Machine learning algorithms in the environmental corrosion evaluation of reinforced concrete structures-A review," *Cem. Concr. Compos.*, vol. 133, p. 104725, 2022.

- [23] C. Frenkel, D. Bol, and G. Indiveri, “Bottom-up and top-down approaches for the design of neuromorphic processing systems: tradeoffs and synergies between natural and artificial intelligence,” *Proc. IEEE*, 2023.
- [24] T. D. Akinosho *et al.*, “Deep learning in the construction industry: A review of present status and future innovations,” *J. Build. Eng.*, vol. 32, p. 101827, 2020.
- [25] W. Zhang, H. Li, Y. Li, H. Liu, Y. Chen, and X. Ding, “Application of deep learning algorithms in geotechnical engineering: a short critical review,” *Artif. Intell. Rev.*, pp. 1–41, 2021.
- [26] M. H. Mobarak *et al.*, “Scope of machine learning in materials research—A review,” *Appl. Surf. Sci. Adv.*, vol. 18, p. 100523, 2023.
- [27] A. I. Torre-Bastida, J. D\`iaz-de-Arcaya, E. Osaba, K. Muhammad, D. Camacho, and J. Del Ser, “Bio-inspired computation for big data fusion, storage, processing, learning and visualization: state of the art and future directions,” *Neural Comput. Appl.*, pp. 1–31, 2021.
- [28] Y. Cheng *et al.*, “Weakly guided adaptation for robust time series forecasting,” *Proc. VLDB Endow.*, vol. 17, no. 4, pp. 766–779, 2023.
- [29] A. H. Khalaf *et al.*, “Emerging AI technologies for corrosion monitoring in oil and gas industry: A comprehensive review,” *Eng. Fail. Anal.*, p. 107735, 2023.
- [30] A. D. A. Bin Abu Sofian *et al.*, “Advances, Synergy, and Perspectives of Machine Learning and Biobased Polymers for Energy, Fuels, and Biochemicals for a Sustainable Future,” *Energy & Fuels*, vol. 38, no. 3, pp. 1593–1617, 2024.
- [31] J. Guo *et al.*, “A deep feature learning method for remaining useful life prediction of drilling pumps,” *Energy*, vol. 282, p. 128442, 2023.
- [32] V. Shanawad, “Cement Manufacturing Concrete Dataset.” 2020.
- [33] A. Dixit, G. Geng, H. Du, and S. Dai Pang, “The role of age on carbon sequestration and strength development in blended cement mixes,” *Cem. Concr. Compos.*, vol. 133, p. 104644, 2022.
- [34] M. Arifuzzaman, A. F. Alfuhaid, A. B. M. S. Islam, M. T. Bhuiyan, M. H. Tito, and A. Gul, “From Mix Design to Strength Prediction: Ensemble Learning Application on the Performance of High-Performance Concrete,” in *2024 ASU International Conference in Emerging Technologies for Sustainability and Intelligent Systems (ICETSYS)*, 2024, pp. 1584–1588.

- [35] O. R. Abuodeh, J. A. Abdalla, and R. A. Hawileh, "Assessment of compressive strength of Ultra-high Performance Concrete using deep machine learning techniques," *Appl. Soft Comput.*, vol. 95, p. 106552, 2020.
- [36] 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.*, vol. 16, p. e00840, 2022.
- [37] V. M. John, M. Quattrone, P. C. R. A. Abrao, and F. A. Cardoso, "Rethinking cement standards: Opportunities for a better future," *Cem. Concr. Res.*, vol. 124, p. 105832, 2019.
- [38] K. Stergiou, C. Ntakolia, P. Varytis, E. Koumoulos, P. Karlsson, and S. Moustakidis, "Enhancing property prediction and process optimization in building materials through machine learning: A review," *Comput. Mater. Sci.*, vol. 220, p. 112031, 2023.