Engineered Science
DOI: https://dx.doi.org/10.30919/es1182

Hybrid Deep Learning Approach for Accurate Prediction of Flowability in Ultra-High-Performance Concrete

Ayat Mahmoud Al-Hinawi[1 # *] Radwan A. Alelaimat,[1 #] Esraa Alhenawi[2 #] and Mohammad I. AlBiajawi[3 #]

Abstract

By implementing several machine learning (ML), deep learning (DL), and hybrid deep learning models, the research methodology included a systematic approach, which included data separation, exploratory data analysis (EDA), artificial neural networks (ANN), K-Nearest neighbors (kNN), convolutional neural networks (CNN), long short-term memory (LSTM), Gated recurrent units (GRU), and convolutional neural network long short-term memory/gated recurrent units hybrid models. Also, the mean absolute error (MAE), R-squared (R^2), and Root Mean Square Error (RMSE) were utilized to evaluate these models. Our results demonstrate that hybrid deep learning models, specifically the CNN-GRU configuration, achieve better performance in predicting ultra-high-performance concrete (UHPC) flowability compared to individual Deep Learning models and traditional Machine Learning approaches. The CNN-GRU model exhibited the best predictive accuracy with a RMSE of 1.360066 and MAE of 1.036573. Additionally, feature selection techniques enhanced the performance of certain models, with the feature-selected random forest model showing notable improvements in accuracy, achieving an RMSE of 1.032841 and MAE of 0.767066. Infrastructure durability and building processes can be improved with higher Ultra-High-Performance Concrete flowability prediction, which improves the effectiveness of various operations of the UHPC mixture design and benefits the application.

Keywords: Accurate prediction; Hybrid deep learning; Flowability; Ultra-high-performance concrete.
Received: 28 March 2024; Revised: 30 May 2024; Accepted: 12 June 2024.
Article type: Research article.

1. Introduction

Flowing Ultra-High-Performance Concrete (UHPC) has good mechanical properties.[1] Successful mixture design and construction require accurate UHPC flowability calculations, as illustrated by the comparison of predicted and experimental values shown in Fig. 1 which elucidates the discrepancies between predicted and actual flowability values, emphasizing the challenges in modeling such a sophisticated material property and underscoring the importance of refining predictive methodologies for better precision in construction practices.

[1] *Department of Allied Engineering Sciences, Hashemite University, Zarqa, 13133 Jordan.*
[2] *Faculty of Information Technology, Zarqa University, Zarqa, 13110, Jordan.*
[3] *Faculty of Civil Engineering Technology, Universiti Malaysia Pahang AL-Sultan Abdullah, Persiaran Tun Khalil Yaakob, Gambang, Pahang, 26300, Malaysia.*
[#] *All authors contributed to this work equally.*
**Email:* ayat05@hu.edu.jo (A. M. Al-HInawi)

The x-axis represents the dataset index, which sequentially identifies each data point or observation in the dataset. The left y-axis measures the flowability of UHPC in centimeters (cm), while the right y-axis quantifies the error in these measurements, also in centimeters (cm).

UHPC fluidity refers to the ability of UHPC mix to flow, spread, and fill the formwork under its own weight, with minimal mechanical assistance. This critical property influences the ease of placement and quality of the concrete in structures, particularly where complex shapes or dense reinforcement configurations are involved.

Fluidity in UHPC is indicative of the material's rheological properties—specifically its viscosity and yield stress. High fluidity suggests low viscosity and yield stress, allowing the concrete to flow easily, which is crucial for achieving a homogeneous distribution without mechanical vibration. This property is vital for ensuring that UHPC fills the formwork and encapsulates the reinforcement without creating honeycombs or voids, thus ensuring structural integrity and durability. Fluidity affects the workability of UHPC, determining how well it can be handled, pumped, and finished in construction applications.

Recent ML and DL advances have excelled in prediction problems. Flowability makes UHPC excellent for high-strength structural parts, precast components, and bridge engineering.[2,3] Hydration, microstructure, characteristics, mix design, and additive integration have been intensively researched in UHPC development and application. Fiber reinforcing, nano-silica (NS), steel slag, and other cementitious materials have been studied. Arora *et al.* (2018)[4] and Alsalman (2017)[5] highlighted UHPC binder selection, rheology, and mechanical properties. Wang *et al.* (2015)[1] and Yoo and Banthia (2016)[6] study UHPC and fiber reinforcement hydration, microstructure, mechanical characteristics, and mix design.

This work develops and tests UHPC flowability prediction algorithms using (ML), (DL), and hybrid DL. For each intelligence technique (ML, DL, hybrid DL), the study selects features. This study examined 135 UHPC mix combinations with 21 flowability-affecting inputs. This study aims to precisely forecast UHPC's flowability, which affects its workability and location during construction. To accomplish this, the investigation begins with (EDA). This step seeks to understand the dataset and link input parameters to UHPC flowability.

High strength, durability, and workability make UHPC a sophisticated building material. UHPC has fine aggregates, high cementitious components, and low water-to-cement ratio. A better combination design and materials give UHPC higher mechanical properties, environmental resistance, and lower permeability.

UHPC is liquid and solid complex.[7] More pieces, combinations, relative proportioning, and features make this concrete touzher to predict. This project will develop four Artificial Neural Networks (ANN)-based analytical models to estimate compressive strengths (CS) and slump flow over one, seven, and 28 days. Portland cement and silica fume were substituted by variable-particle limestone, recycled glass, and fluid catalytic cracking (FCC). Initial, seventh, and eighth-day CSs and slump flow yielded prediction error values of 2.400 MPa, 2.638 MPa, 2.064 MPa, and 7.245 mm. ANN models with limestone, silica fume, FCC, and recycled glass powder predicted UHPC slump flow and CSs.

Underwater concrete (UWC) mix rheological and mechanical qualities and constituent material sensitivity are predicted by (ANN).[8] ANN can approximate and accomplish tasks independently by mimicking organic neurons. From ideation to training to validation, their research describes how they created the neural network (NN) model. The ANN model was trained and tested using 175 distinct UWC permutations from 9 experiments. Data is organized by pattern. Each pattern inputs that alter UWC mixture behavior and outputs the rheological or mechanical property to be represented. The ANN model can accurately predict the attributes above for new underwater concrete mixtures generated within the realistic range of the training phase input parameters with absolute errors of 4.6, 10.6, and 4.4%. NNs may predict flowable concrete's CS and new concrete properties, according to Jayaseelan *et al.* (2019).[9] Standardized lab investigations yielded complete data. One nanoparticle-two microparticle combination made flowable concrete. A NN model using BFGS Quasi-Newton, Fletcher-Powell, Polak-Ribiere, Gradient Descent (GD) with Adaptive Linear Back Propagation, and Levenberg-Marquardt back propagations predicts better than 90%. The model properly predicted flowable concrete's new characteristics and CS.

However, the size of the dataset and atypical data utilized for training prediction models limit both their accuracy and their capacity for generalization.[11] ML techniques can predict concrete attributes and speed up advanced concrete design, which is vital considering the variety of concrete types used in construction. Fig. 2 provides an overview of different types of concrete. Their research suggests an ML framework for predicting (UHPC) compressive, workability, porosity, and flexural. The framework has three parts: Data (1) is conflicting. Using isolation forests, integrated mutual information, and univariate (LR), (2) an unsupervised anomaly identification

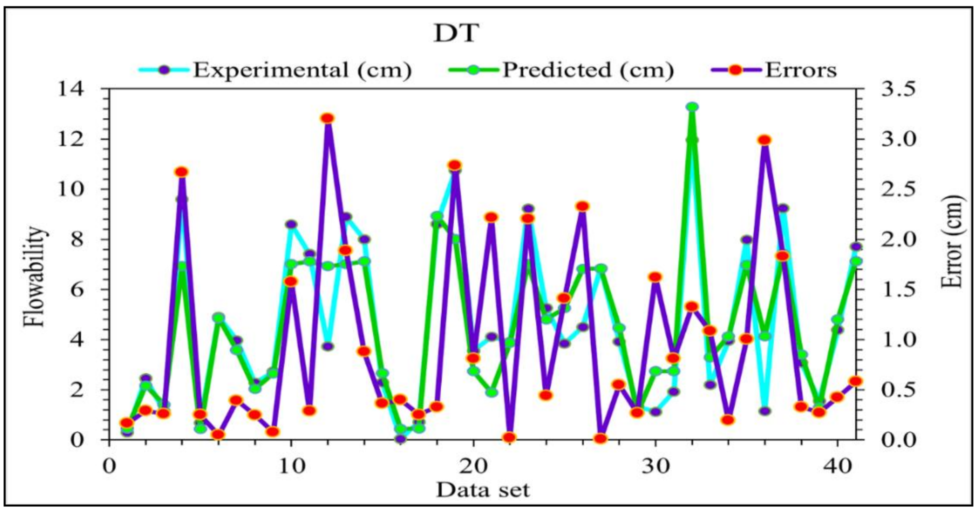

Fig. 1 DT predicted and experimental error values—UHPC flowability, reproduced with permission from [10] Copyright @Qian *et al.*, 2022.

method is used to find and eliminate outliers and irrelevant variables in a dataset ML hyperparameters can be tuned using a tree-structured Parzen estimator and k-fold cross-validation, as demonstrated in reference (3) The learning framework and Light Gradient Boosting Machine build auto-tuned prediction models. The developed approach demonstrates strong predictive accuracy. Auto-tune models study mixture design variances on characteristics. Reduce experimentation to improve material development.

Material composition determines UHPC CS. A prediction model that works with an experimental dataset frequently requires complex algorithms like the ANN to experimentally capture this link.[12] Its opaqueness makes mathematically characterizing its interior workings difficult for scientists. In their research, they apply Sequential Feature Selection (SFS) and Neural Interpretation Diagram (NID) to identify key ANN elements. ANN was trained using 110 (UHPC) CS tests. The (ANN) outperformed the model employing all eight characteristics in prediction accuracy, with a correlation coefficient of 1% and NMSE of 0.012 compared to 0.5% and 0.035. The technique produced a parametric analysis and non-linear regression (LR) model based on four variables. Combining ANN with SFS and NID enhanced model accuracy and showed how ANN could estimate CS for different UHPC permutations.

Choudhary *et al.* estimate UHPC CS using back-propagation neural network (BPNN) and (SFS).[13] ML models were constructed using a literature-based 110-point, eight-material constituent database. BPNN and SFS were interchangeable for response variable characteristics. Thus, the BPNN with selected characteristics outperformed the model with all features (0.816) ($R^2 = 0.991$). Civil engineering ML case studies benefit from ANN with SFS prediction model accuracy. AI advancements have fostered intelligent UHPC creation. Fan provide a reliable Modified Andreasen and Andersen (MAA) model and Genetic Algorithm based Artificial Neural Network GA-ANN UHPC design and features prediction.[14] A GA-ANN model for UHPC features prediction outperforms established approaches in fitting goodness and prediction accuracy using 80 mixtures as a training dataset. GUIs are simple for GA-ANN prediction applications. Finally, MAA and GA-ANN models provide a new UHPC mix-design method. GA-ANN and property criteria for final optimization, MAA model for preliminary mixture design. They found that AI can build a dense particle-packing skeleton UHPC.

Nurlan found few studies predicting fresh or hardened Self-Compacting Concrete (SCC).[16] Effective Radial Basis Function NN (RBFNN) models predict fresh and toughened self-compacting concrete qualities. RBFNN parameters are optimized using ant-lion optimization (ALO) and biogeography optimization (BBO). The results show strong learning and testing. The association between observed and predicted SCC characteristics using hybrid models shows great training and approximation accuracy. ALO-RBFNN beat literature and BBO in D flow, L-box, V-funnel, and CS. ALO's RBFNN model outperforms others in determining optimal method parameters.

New UHPC improves mechanical, rheological, and durability.[17] Manufacturers must adjust UHPC constituent quantities for strength, flowability, and cost. Traditional concrete mixture design requires expensive and time-consuming testing. Many goals have been achieved by Statistical Mixture Design (SMD), design of experiments (DOE) design, and mathematical optimization. Conventional methods involve multiple (LR) and other objective functions. After modeling a problem, mathematical programming and simplex algorithms can find optimal solutions. Data that does not fit basic regression models like multiple (LR) requires a more adaptable technique that allows high-accuracy nonlinear models and varied multi-objective mixture formation circumstances. The authors propose a steel-fiber-enhanced

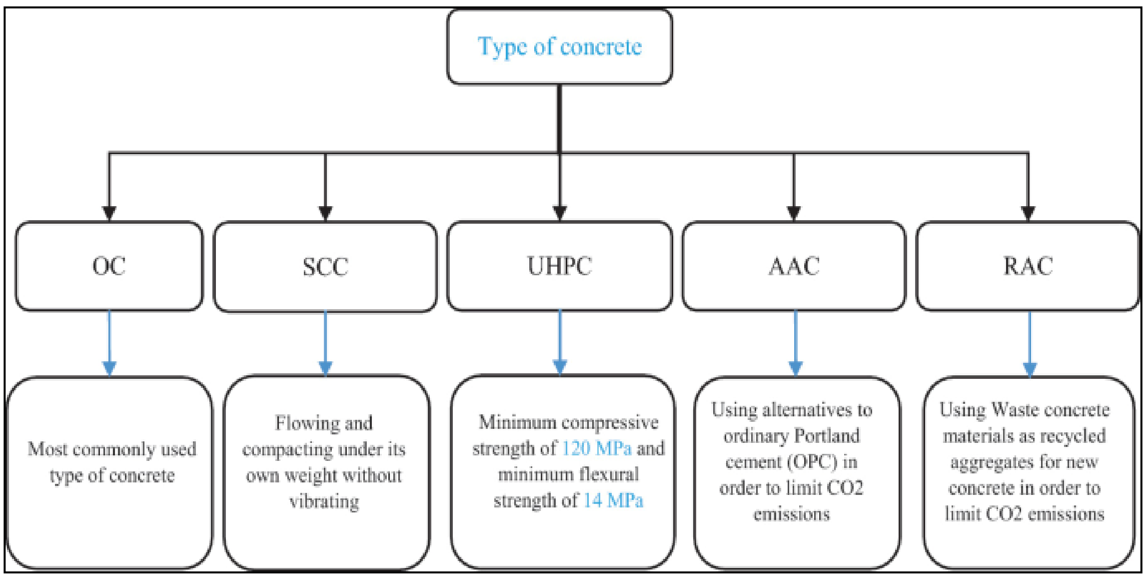

Fig. 2 Types of Concrete, reproduced with permission from [15] Copyright Asghari *et al.*, 2023.

UHPC efficiency approach. ANNs and Gaussian process regression are used in multi-objective learning. Theory and outcomes are based on solid experiments. Experimental results validate the multi-objective mixed design and optimization method for steel fiber reinforced UHPC. The method streamlines (UHPC) trial design and optimizes strength, flowability, and cost.

UHSC is a good civil engineering building material.[18] Construction time and cost are reduced by soft computing concrete quality estimates. Their study calculated UHSC CS utilizing sophisticated soft computing. They investigated XGBoost, AdaBoost, and Bagging. RMSE, R^2, and MAE were used to evaluate algorithms. The model was statistically assessed. With lower R^2 (0.90) and errors, XGBoost soft computing was more correct. XGBoost estimates UHSC CS. SHapley Additive exPlanations noted that curing time increased UHSC CS the most. Their findings will let specialists quickly and precisely evaluate UHSC's CS.

Khan and Suthar were stated UHPC's strength and durability make it popular.[19] Their 28-day CS prediction comparison of M5P and random forest (RF) models in UHPC. Their investigation includes 236 readings. 70% of 236 readings-157-are used for training, while the other 30 are tested. Compare (RF) and M5P to determine which predicts UHPC CS better. Performance measurements like Corelation Coefficient (CC), RMSE, and MAE determine model soundness. Random Forest (RF) outperformed M5P in testing with CC, RMSE, and MAE values of 0.8568, 16.005, and 12.03.

Many UHPC experiments have been conducted, according to Marani *et al.* (2020).[20] UHPC and its heterogeneous composition have nonlinear engineering properties that standard statistical approaches cannot define. To build accurate and insightful nonlinear materials science property prediction tools, effective and creative approaches are needed to concentrate key experimental data. ML may reveal complex data patterns. UHPC CS is predicted using modern ML algorithms and a large experimental dataset of 810 test results and 15 academic input variables. Tabular generative adversarial networks generated 6513 synthetic data points for (RF), additional trees, and gradient boosting regression models. Models were tested on 810 fresh experimental data sets after training on simulated data. Predictively, models performed well. Parametric tests revealed UHPC strength development methods and parameters using the models.

UHPC has high CS, strain hardening under stress, and self-healing.[21] UHPC self-healing prediction models are ignored despite study. Data-driven AI and ML models are increasingly predicting attributes. Multi-physics modeling predicts cement-based materials' chemical, physical, and mechanical properties. Whale Optimization Algorithm (WOA), Grey Wolf Optimizer (GWO), and Flower Pollination Algorithm (FPA) with Xgboost meta-heuristic algorithms were employed to create a ML model to predict UHPC self-healing. The model was based on main experimental research on UHPC fracture-sealing over six months under sustained crack tensile stress in severe conditions. Four mathematical measures evaluated the model's anticipated accuracy. The REC and Taylor diagrams revealed that optimal models worked effectively with various optimization strategies. SHAP found that exposure length and fracture diameter best predicted self-healing. The study found that ML could predict UHPC self-repairing and identify key factors.

The paper addresses the environmental issues caused by the production of cement, particularly its significant CO2 emissions that contribute to global warming and other health problems.[25] It explores the use of fly ash (FA), a by-product from thermal power stations, as a sustainable replacement for ordinary Portland cement in the construction industry. The study focuses on the development of net-zero mortars using FA to enhance its reactivity and effectiveness. The paper evaluates the mechanical properties of these mortars through CS tests, varying the water-to-cement ratio and the use of a super plasticizer. The findings demonstrate that the best-performing mortar includes fine FA with specific mix proportions, confirming the potential of FA as a viable and environmentally friendly cement substitute.

The paper investigates the use of a hybrid cement mixture to address environmental concerns related to marine and industrial waste disposal.[26] This study aligns with the United Nations Sustainable Development Goals (UNSDG) and COP27 climate actions by exploring low carbon technologies in the cement industry. The research evaluates the bonding behavior and strength of this 3R hybrid cement in various masonry and mortar tests, assessing chemical characteristics, workability, CS, and other properties. The findings highlight the environmental and economic benefits of using hybrid cement and demonstrate superior performance in terms of the CS of brick masonry prisms.

The paper addresses the imperative of transitioning to net-zero construction materials in alignment with UNSDG' objective to mitigate carbon emissions by 2050.[27] Focusing on the replacement of cement with pozzolanic materials to reduce its carbon footprint, the study employs gene expression programming (GEP) and artificial neural network (ANN) machine learning (ML) techniques to predict the CS of FA concrete. By utilizing a dataset compiled from various sources, encompassing input parameters, the study demonstrates the applicability of ML in forecasting the short- and long-term CS of FA concrete toward carbon neutrality in infrastructure. Comparing GEP and ANN models based on statistical parameters, the study reveals GEP's superiority in estimating CS at both time intervals. Furthermore, the GEP model offers a simplified equation for predicting CS at different ages of net-zero FA concrete, providing valuable insights for sustainable construction practices.

In their research paper, Onyelowe and Ebid (2023b)[28] explore the prediction of CS in high-performance concrete (HPC) containing FA and slag using ML techniques. They employ genetic programming (GP), ANN, and evolutionary polynomial regression (EPR) on a dataset comprising 1030 entries with variables. The study divides the data into training and validation sets, with results indicating that ANN outperforms GP and EPR models, exhibiting the highest consistency between predicted and measured values. ANN's superior performance is reflected in its accuracy metrics, including a high coefficient of determination (R^2), low mean absolute error (MAE), mean squared error (MSE), and root mean squared error (RMSE). The study suggests that ANN could serve as a decisive model in HPC design, particularly regarding CS, contributing to the development of sustainable structures in the built environment.

The paper focuses on advancing concrete technology towards carbon neutrality, particularly by optimizing net-zero concrete mixes incorporating industrial waste materials.[29] The study aims to develop predictive models for CS at different ages using novel metaheuristic techniques. Utilizing a database of 1133 net-zero concrete mix records, the study evaluates performance metrics including MAE, MSE, and R2 score. AutoML tools, particularly AutoSklearn and AutoGluon, were employed to swiftly generate models with high accuracy. AutoGluon notably outperformed AutoSklearn and traditional methods like Support Vector Regression (SVR) and RF, achieving a superior R^2 of 92.6% with lower MAE and MSE. Notably, BFS emerges as a significant contributor to ecofriendly concrete strength, showcasing its potential as a sustainable alternative to cement compared to FA. The study underscores the efficacy of AutoML techniques in surpassing previous methods, highlighting their role in advancing ecofriendly construction practices.

In their study, Onyelowe *et al.* (2023d)[30] investigate the utilization of NS as a precursor to enhance the CS of mortar, crucial for sustainable building construction. They produced multiple mortar mixes using NS precursor (NSP) and employed Advanced ML (AML) techniques to predict CS, training and validating models with various NSP ratios. The NS precursor significantly influenced mortar strength, particularly due to its contribution to pozzolanic reactions forming C-S-H gel. Model accuracies were assessed using performance indices and Taylor charts, revealing the superiority of the ANN model. Despite lacking a closed-form expression, the ANN model outperformed others, demonstrating potential for NSP application in sustainable construction as a reliable pozzolanic material.

A summary of recent studies related to various concrete types and their predicted properties can be found in Table 1.

Table 1. Related work summarization.

Study	Objective	Methodology	Key Findings
Nurlan (2022)[16]	Develop efficient RBFNN models for SCC fresh and hardened properties	RBFNN, ALO, BBO	ALO-RBFNN outperformed BBO and achieved high accuracy in predicting SCC properties. ALO demonstrated superior optimization capabilities.
Qian *et al.* (2022)[10]	Predict UHPC flowability and CS	DT, BA, GB	GB provided accurate predictions for UHPC flowability and CS, outperforming other algorithms. Limestone powder content and curing time were identified as influential parameters.
Qian *et al.* (2023)[22]	Predict UHPC flexural strength	SVM, MLP, GB	GB achieved better prediction accuracy for UHPC flexural strength than SVM and MLP. Steel fiber composition was identified as the most influential parameter.
Sadrossadat *et al.* (2022)[17]	Develop ML-based models and optimize UHPC mixture design	ANN, GPR, PSO	ANN and GPR. PSO was employed for multi-objective mixture design and optimization, considering strength, flowability, and cost.
Shen *et al.* (2022)[18]	Estimate UHSC CS using soft computing methods	XGBoost, AdaBoost, Bagging	XGBoost demonstrated higher accuracy in estimating UHSC CS compared to other methods. Curing time was identified as the most influential parameter.
Abuodeh *et al.* (2020)[12]	Develop a predictive model for UHPC CS	SFS, NID, ANN	SFS and NID helped determine important material constituents for ANN, resulting in more accurate predictions.
Al Sarfin *et al.* (2023)[23]	Create a locally sourced, non-proprietary UHPC mix design based on the required qualities.	RF, SVM, GB	ML models trained on mix design parameters can reverse-engineer UHPC mix proportions for desired qualities.
Choudhary *et al.* (2021)[13]	Forecast UHPC CS	BPNN, SFS	BPNN with selected features achieved higher accuracy compared to models with all features.
Fan *et al.* (2021)[14]	Accurate design and characteristics prediction of UHPC	MAA model, GA-ANN	The GA-ANN model outperformed traditional methods regarding fitting goodness and prediction accuracy for UHPC characteristics.
Farouk *et al.* (2022)[24]	Predict UHPC and steel reinforcing bar bond	AI, IEPANN)	The IEPANN model accurately predicted bond strength between UHPC and reinforcing bars. Monte Carlo simulation

	strength		quantified uncertainties.
Kumar *et al.* 2023)[25]	Develop sustainable cement alternatives to reduce CO2 emissions and achieve carbon neutrality.	Utilization of FA in place of cement; experiments with different granularities and mix proportions, including the use of super plasticizers.	Fine FA mortar with a specific water-to-cement ratio and super plasticizer showed the best performance, validating FA as an effective cement substitute.
Ravi *et al.* (2023)[26]	Recycled waste materials in cement.	Evaluation of 3R hybrid cement (oyster shell, slag, tyre waste) in brick masonry prisms; tests on bonding, CS, and microstructure.	Hybrid cement significantly enhances the CS of brick masonry prisms, with positive environmental impacts.
Onyelowe *et al.* (2023a)[27]	To predict the 56 days and 180 days CS of net-zero FA concrete using ML techniques	Adopted GEP and ANN	Both GEP and ANN were used to determine CS. GEP outperformed ANN in estimating CS at both 56 days and 180 days. GECSP produced a simplified equation for predicting CS at different ages of net-zero FA concrete.
Onyelowe and Ebid (2023b)[28]	To predict CS in HPC with FA and slag using ML techniques	Employed GP, ANN, and evolutionary EPR on a dataset comprising various concrete mix parameters.	ANN demonstrated superior performance, exhibiting high consistency between predicted and measured values. It showcased higher accuracy metrics (R2, MAE, MSE, RMSE) compared to GP and EPR, suggesting its utility in HPC design for sustainable structures.
Onyelowe *et al.* (2023c)[29]	Develop predictive models for CS in net-zero concrete mixes with industrial waste materials.	Employ superspeed metaheuristic predictive techniques and AutoML tools on a database of 1133 records.	AutoGluon outperforms, achieving R2 of 92.6%; BFS significantly impacts strength, suggesting it as a sustainable cement alternative.
Onyelowe *et al.* (2023d)[30]	Investigate the use of NS precursor in enhancing mortar CS for sustainable building construction.	Utilize AML techniques to predict CS of mortar using various NS precursor ratios, and compare model accuracies using performance indices and Taylor charts.	ANN model outperforms, exhibiting highest accuracy with MAE of 1.47 MPa, MSE of 3.84 MPa, RMSE of 1.96 MPa, and R2 of 0.980. NSP significantly improves mortar strength, confirming its potential as a sustainable pozzolanic material in construction.

2. Experimental

2.1 Research approach

The investigation commences with the collection and upload of the UHPC dataset, followed by configuring initial input settings for the analysis. The primary step involves Exploratory Data Analysis (EDA) to delve into the dataset, aiming to unearth patterns and ascertain the connections between the various input and target feature, UHPC flowability.

Upon concluding the EDA, the dataset is systematically split into an 80:20 ratio for training and testing purposes. This segregation ensures that a substantial portion of the data is utilized for model training, while a distinct set is reserved for performance evaluation, thereby maintaining the robustness of the testing process.

Multiple predictive models are then developed to estimate UHPC flowability, encompassing traditional ML techniques such as RF, Logistic Regression (LR), and k-Nearest Neighbors (kNN). Additionally, several deep learning (DL) approaches are explored, including ANNs, Long Short-Term Memory networks (LSTM), Gated Recurrent Unit networks (GRU), and Convolutional Neural Networks (CNN). To exploit the combined benefits of convolutional and recurrent networks, hybrid models like CNN-LSTM and CNN-GRU are also constructed.

All models are rigorously evaluated using the metrics Root Mean Square Error (RMSE), MAE, and the R^2, which collectively demonstrate the accuracy and precision of the predictions concerning UHPC flowability. To enhance model performance and interpretability, FS using Recursive Feature Elimination (RFE) is applied to each type of model— ML, DL, and hybrid DL. RFE assists in pinpointing ML critical input features that significantly impact UHPC flowability.

- ML Models: Parameters for RF, LR, and kNN are optimized through grid search to determine the most effective settings. For instance, the number of trees in RF and the number of neighbors in kNN are tailored based on the dataset characteristics.
- DL Models: For ANN, LSTM, GRU, and CNN, models are configured with specific layers, neuron counts, and

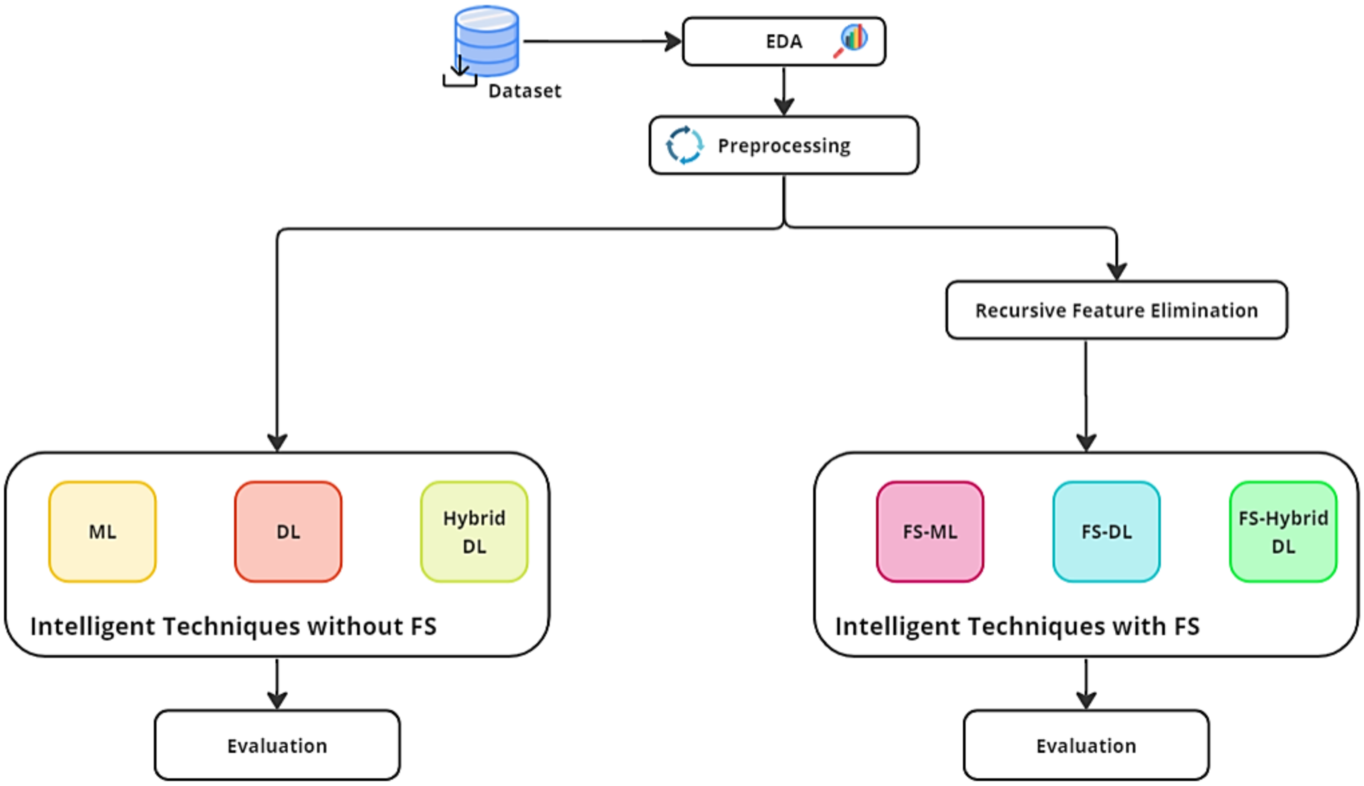

Fig. 3 Research methodology.

activation functions (AF) suited to the data's nature. For example, the number of LSTM blocks is chosen based on the temporal depth required to capture the dependencies in the data effectively.

- Hybrid Models: CNN-LSTM and CNN-GRU models are meticulously designed to first utilize CNN layers for feature extraction, followed by LSTM or GRU layers to analyze temporal sequences, ensuring a comprehensive understanding of both spatial and temporal dynamics.

Finally, models that incorporate FS-labeled FS-ML, FS-DL, and FS-Hybrid DL are constructed to verify the efficacy of the selected features in predicting UHPC flowability accurately. This methodical approach not only aims to develop robust predictive models but also enhances the models' interpretability by focusing on influential features.

The entire methodology, from data handling through model evaluation, is succinctly illustrated in Fig. 3. This figure encapsulates the sequential steps undertaken, offering a clear visualization of the process flow and aiding in the understanding of the methodological rigor applied throughout the study.

2.2 Dataset description

UHPC fluidity estimation uses 135 mix combinations and 21 input parameters. This dataset assessed multiple AI UHPC flowability forecasting algorithms. Parameter details for the dataset are outlined in Table 2, which provides information on the various input features used for UHPC flowability estimation:

Flowability is vital to construction workability and positioning (cm).

UHPC flowability is affected by many elements, including dataset input parameter proportions and attributes. These input features affect UHPC flowability. The research's ML models are trained and tested using the dataset from relevant literature.

2.3 Exploratory Data Analysis (EDA)

To forecast UHPC flowability, we use EDA approaches to analyze the dataset and comprehend the target variable. We use a histogram to visualize the target variable distribution during EDA. The code snippet plots flowability as a histogram. Our visualization uses Python's Seaborn package to select plot style, color, and font scale. To organize data and display occurrence counts for each bin, the histogram has 20 bins. The alpha parameter determines histogram bar transparency. The plot in Fig. 4 shows the flowability variable's distribution, showing its range and frequency. Flowability distribution can be seen in the histogram. It shows the concentration of flowability values around specified ranges or peaks, revealing the target variable's variability and spread. This image helps identify flowability distribution outliers and skewness, which can be important in modeling and analysis. The desired variable, UHPC flowability, can be studied with EDA. Selecting good hybrid deep-learning UHPC flowability

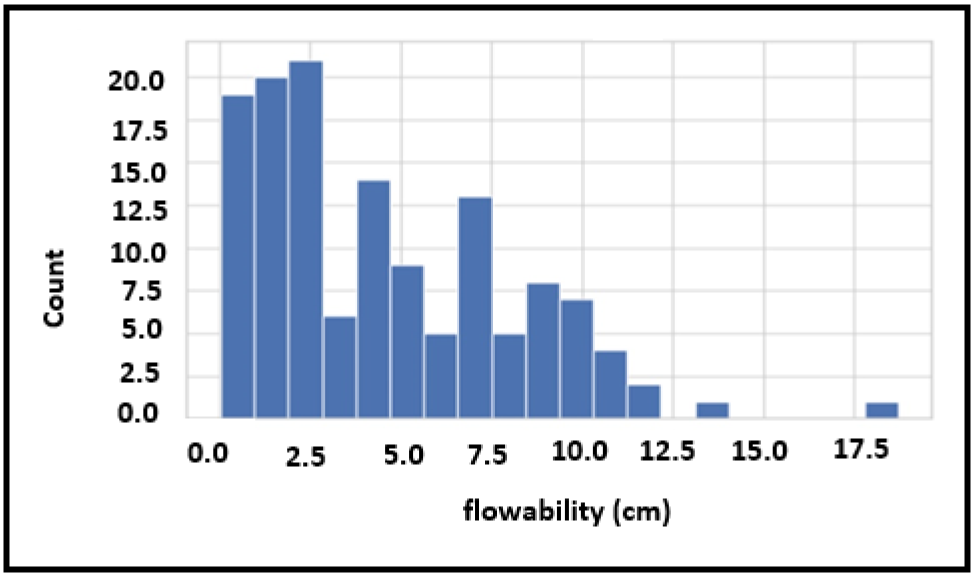

Fig. 4 Data visualization.

Table 2. Input features.

number	features	number	features
1	Cement Content: The amount of cement used in the UHPC mix (Kg/m^3).	12	Polystyrene Fiber Content: The quantity of polystyrene fibers in the mix (%).
2	FA Content: The quantity of FA in the mix (Kg/m^3).	13	Water Content: The amount of water used in the UHPC mix (%).
3	Silica Fume Content: The amount of silica fume added to the mix (Kg/m^3).	14	Type of Cement: The type of cement used.
4	Slag Content: The quantity of slag included in the mix (Kg/m^3).	15	Strength Class of Cement: The strength class of the cement.
5	Nano-Silica Content: The NS content in the mix (Kg/m^3).	16	CS of Cement: The CS of the cement (MPa).
6	Quartz Powder Content: The amount of quartz powder used in the mix (Kg/m^3).	17	Maximum Aggregate Size: The maximum size of the aggregates used (mm).
7	Limestone Powder Content: The quantity of limestone powder in the mix (Kg/m^3).	18	Length of Polystyrene Fiber: The length of the polystyrene fibers (mm).
8	Sand Content: The amount of sand in the UHPC mix (Kg/m^3).	19	Diameter of Polystyrene Fiber: The diameter of the polystyrene fibers (µm & mm).
9	Coarse Aggregates Content: The quantity of coarse aggregates in the mix (Kg/m^3).	20	Length of Steel Fiber: The length of the steel fibers (mm).
10	Super-plasticizers Content: The amount of super-plasticizers used in the mix (Kg/m^3).	21	Diameter of Steel Fiber: The diameter of the steel fibers (µm & mm).
11	Steel Fiber Content: Steel fibre content in the UHPC mix (Kg/m^3).		

forecast modeling and preprocessing methods requires this expertise.

2.4 Preprocessing

ML and analysis require data preparation. Creating and testing UHPC flowability models requires data preprocessing. Data preprocessing requires splitting. The dataset includes training and testing. Data is used to train ML models to uncover 80% patterns and correlations. The remaining 20% examines model performance and fresh data generalization. The data is segregated to prevent overfitting, which happens when a model accomplishes well on training data but not new data. Our testing set lets us evaluate models' predictions on unknown data, enhancing reliability.

Model generalization and performance measures can be tested by data splitting. The models' UHPC flowability prediction accuracy, precision, and performance are measured. Data splitting during preprocessing supports reliable ML model creation and evaluation. The training set educates models using data patterns and correlations, while the testing set evaluates model performance independently. This thorough preprocessing enhances prediction models for real-world UHPC flowability projections.

2.5 Machine learning methods

This work predicts UHPC flowability using ML after preprocessing and partitioning data into training and testing sets. We use common ML techniques to create prediction models. The ensemble learning method RF builds numerous decision trees and combines their predictions for more accurate and resilient outcomes. It excels at complicated datasets and non-linear input parameter-UHPC flowability correlations. Non-parametric KNN classifies by proximity. The K nearest data instances in the training set and the majority label or value from these neighbors predict UHPC flowability. KNN is simple and flexible, making it appropriate for many applications. Traditional regression algorithms like LR use linear functions to model input parameters and flowability. It estimates the coefficients that best suit the data, revealing the direction and size of each input parameter's effect on UHPC flowability. Researchers may discover patterns and relationships from the training data and construct models that predict UHPC flowability depending on input parameters using these ML methods. We may compare the flowability prediction accuracy of these methods due to their strengths and weaknesses.

2.5.1 Deep learning methods

ANN

In the following level of our research, we use DL approaches, starting with ANN architecture. Interconnected layers of artificial neurons mirror the brain's structure and function in the ANN DL model. It seems capable of capturing complex data patterns and linkages. This is our ANN model's architecture. The Keras library builds and trains NNs at a high level. The model's Sequential class arranges layers sequentially. The model starts with a 64-neuron Dense layer. The Rectified Linear Unit (ReLU) AF for this layer introduces non-linearity and lets the model learn complicated representations. The amount of input parameters in our dataset determines the geometry of this layer.

The second Dense layer uses ReLU activation and 64 neurons. The additional hidden layer captures more complex data patterns and linkages. We conclude with a single-neuron output layer to forecast UHPC flowability, a continuous target variable. Since we want the model to output the projected flowability value directly, this output layer has no AF. Compiling the model with the optimizer and loss function (LF) prepares it for training. For this task, we employ the Adam optimizer, an efficient stochastic GD technique, and the MSE to reduce the difference between estimated and real flowability values. The model is trained on training data using a batch size and epochs. Internal model parameters are modified to minimize the LF during training using the optimization procedure. To assess model performance on unobserved data, the validation split parameter is 0.2, using 20% of training data for validation.

LSTM

We forecast UHPC flowability using LSTM, a customized RNN, in our next phase. LSTM excels in sequential data modeling and long-term dependencies. The following describes our LSTM model's architecture. The NN is built and trained using Keras. Model layers are ordered sequentially by the Sequential class. The model starts with a 64-memory LSTM layer. The ReLU AF for this layer introduces non-linearity and lets the model capture complex sequential data patterns. In our dataset, the LSTM layer's input shape matches the number of input parameters.

After the LSTM layer, we add a 64-neuron Dense layer with ReLU activation. Building on the LSTM layer's representations, this hidden layer seeks to capture more complex data patterns and relationships. The model concludes with a single-neuron output layer. This output layer has no AF since we want to anticipate UHPC flowability. It immediately produces projected flowability.

Compiling the model requires defining the optimizer and LF. For this task, we employ the Adam optimizer, and the MSE to decrease the difference between predicted and actual flowability values. The model is then trained on training data using batch size and epochs. Internal model parameters are optimized iteratively to minimize the LF during training. To assess classifier performance on unseen data, the validation split parameter is 0.2, using 20% of training data for validation.

GRU

Researchers used a variation of the RNN architecture to estimate UHPC flowability in our research. GRU is built to handle sequential data, making it ideal for capturing our dataset's temporal dependencies. Researchers implemented the GRU model using the popular Keras package, which simplifies NN construction. We defined the model using Keras' Sequential class to organize layers sequentially.

Our proposed architecture revolves around the GRU layer. The model uses 64 memory units to remember past observations and produce predictions. The GRU layer uses ReLU activation. It creates non-linearity and lets the model capture complex sequential data patterns. The GRU layer input form matches the amount of input parameters in our UHPC flowability dataset. Researchers added a 64-neuron Dense layer after the GRU layer to improve model representation. This layer uses ReLU activation to capture and express increasingly complex data patterns and relationships. Adding the GRU layer and Dense layer helps explain UHPC flowability.

The final prediction is obtained from a single-neuron output layer. This output layer has no AF since we want to anticipate UHPC flowability. The model's estimated flowability value for a given input instance is output directly. To maximize model performance, we provide the Adam optimizer and MSE. To reduce the difference between projected and real flowability values, the Adam optimizer efficiently modifies the model's internal parameters during training. A batch size and number of epochs are used to train the model with the training data. The optimization approach iteratively updates model parameters to minimize the LF during training. With a validation split setting of 0.2, 20% of the training data is saved for validation. This lets us evaluate the model's generalization and performance on unseen data. We want to use the GRU model to capture UHPC flowability's temporal dynamics in our research. This DL approach may help us forecast (UHPC) flowability.

CNN

DL methods in our research end with the CNN architecture. CNNs excel at computer vision and time series data analysis, making them ideal for UHPC flowability prediction. The Sequential class in Keras defines our CNN model's sequential layer design. Conv1D, which convolutions input data in one dimension, is the initial layer. The model captures local patterns and characteristics in input sequences using 64 filters with a kernel size of 3. The ReLU AF adds non-linearity and improves the model's capacity to capture complicated data correlations. Our UHPC flowability dataset has the same amount of input parameters as the Conv1D layer's input shape. A MaxPooling1D layer follows the Conv1D layer to reduce output dimensionality and extract the most important features from convolved information. The pooling process captures crucial data while simplifying model computation. Flatten layers prepare the data for fully linked layers, which convert multidimensional output into a one-dimensional vector. This integrates with subsequent layers seamlessly.

Next, the model gets a 64-neuron Dense layer. ReLU activation on this layer adds non-linearity and expressive power to the model. This layer teaches the model abstract characteristics and high-level representations from flattened input. To anticipate the continuous target variable of UHPC flowability, we add a single-neuron output layer. Since this output layer has no AF, it outputs the projected flowability value directly. The model is constructed with the Adam optimizer and MSE, similar to previous models. To reduce the discrepancy between projected and actual flowability values, the Adam optimizer efficiently modifies model internal parameters during training.

2.5.2 Hybrid-deep learning methods

CNN-LSTM

Researchers use hybrid DL to improve our models' prediction skills. CNN-LSTM architecture combines CNNs with LSTM networks. The CNN-LSTM model starts with a Conv1D layer that convolutions input data in one dimension. This layer captures local input sequence patterns and features with 64 filters and a kernel size 3. Non-linearity from the ReLU AF improves the model's capacity to grasp complicated interactions. UHPC flowability dataset input parameters match Conv1D layer input shape.

A MaxPooling1D layer down samples convolved features and extracts the most important input features after the Conv1D layer. This pooling operation reduces data dimensionality and highlights key information. To capture temporal dependencies and long-term trends, an LSTM layer is added. Since they may use earlier time steps, LSTM networks are ideal for sequential data analysis. ReLU activation is used in the 64-memory LSTM layer. A 64-neuron dense layer is added to the model to capture complex data interactions. ReLU activation on this layer adds non-linearity and expressive power to the model. To predict the continuous target variable of UHPC flowability, we include a single-neuron output layer. Since this output layer has no AF, it outputs the projected flowability value directly. The Adam optimizer and MSE compile the model, as with previous versions. The Adam optimizer adjusts model parameters during training to minimize the gap in flowability between expected and observed values.

CNN-LSTM hybrid models can increase UHPC flowability prediction accuracy and understanding by merging CNNs' local characteristics and patterns with LSTM networks' temporal dependencies. The strengths of both architectures are used to capture spatial and temporal information for a more complete UHPC flowability data analysis.

CNN-GRU

We investigate the CNN-GRU architecture for hybrid DL UHPC flowability prediction. This CNN-GRU network architecture provides unique spatial and temporal insights. Conv1D layers apply one-dimensional convolutions to input data in the CNN-GRU hybrid model. This layer captures local input sequence patterns and features with 64 filters and a kernel size 3. The ReLU AF adds non-linearity and better represents complicated interactions. Conv1D layer input shape matches UHPC flowability dataset parameters. After the Conv1D layer, MaxPooling1D down samples convolved features and keeps the most important input. Pooling decreases data dimensionality and highlights critical features. A GRU layer captures temporal dependencies and long-term trends. LSTM-like GRU networks use prior time steps. In this architecture, the 64-memory GRU layer uses ReLU activation. To mimic complicated data interactions, a 64-neuron dense layer is introduced. To increase representation and non-linearity, this layer undergoes ReLU activation. Finally, a single-neuron output layer forecasts continuous target variable UHPC flowability. This output layer outputs estimated flowability without activation.

As with earlier architectures, the Adam optimizer and MSE optimize the model during compilation. CNN-GRU hybrid models combine spatial feature capture and GRU networks' temporal dependencies to increase UHPC flowability prediction accuracy and understanding. This hybrid method evaluates UHPC flowability using spatial and temporal data.

2.5.3 Feature selection

Our ML, DL, and Hybrid-DL FS step is identical. Predicting UHPC flowability requires choosing the best parameters. We use RFE to eliminate unimportant features and identify the most significant ones depending on model performance. Using RFE, we may reduce dataset dimensionality and focus on key prediction factors. Features are selected using the RF Regressor, which handles high-dimensional datasets and complex variable relationships. Once fitted to training data, RFE objects can rank features by importance.

We acquire the most important UHPC flowability prediction characteristics after (FS). Select features are applied to training and testing datasets to build new datasets with only those features. By consistently selecting features across ML, DL, and Hybrid-DL techniques, we train and test models on fewer important attributes, improving efficiency, interpretability, and prediction performance. Reducing unneeded or duplicated information improves model generalization and UHPC flowability predictions.

2.5.4 Evaluation metrics

Researchers evaluate the prediction models utilizing metrics to establish their performance and accuracy in predicting UHPC flowability in the final step of our process. Evaluation metrics include MAE, RMSE, and R^2.

Final step: Researchers analyze prediction models using metrics to determine their performance and accuracy in predicting UHPC flowability. However, RMSE estimates the square root of the average squared discrepancies between anticipated and actual values.

RMSE equation:

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

where, n is the number of samples, y_i is the real value, and $\widehat{y}_i$ is the expected value.

It penalizes big errors more than MAE since it considers magnitude and direction. As with MAE, lower RMSE indicates better model accuracy and performance.

MAE is another evaluation metric used to measure the accuracy of prediction models. It calculates the average absolute differences between the expected and real values. The equation for MAE is:

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

where, n is the number of observations, y_i is the actual value. $\widehat{y}_i$ is the predicted value, R^2 quantifies how much target variable variance the model explains. It measures model fit to data. Better fits have higher values (0–1). A 1 means the model predicts the target variable perfectly. In contrast, a number near to 0 indicates that the model does not capture much data variance.

R^2 equation:

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

where, $\bar{y}$ is the mean of the observed data.

MAE, RMSE, and R^2 can assess our predictive models' UHPC flowability accuracy. These metrics quantify how well models capture data patterns and relationships. More accurate forecasts have lower MAE and RMSE values and higher R^2 values. Based on these evaluation metrics, we may compare and choose the best UHPC flowability prediction model.

3. Results and discussion

3.1 Machine learning methods

Here are the prediction model evaluation results:

Table 3. Comparison of ML methods.

Regressor	RMSE	MAE	R^2
RF	1.074585	0.796568	0.861561
LR	1.337163	1.025122	0.785639
KNN	1.549953	1.349593	0.711986

Table 3 shows the RMSE, MAE, and R^2 assessment metrics for three UHPC flowability prediction ML algorithms: RF, LR, and KNN.

The RF model had the best RMSE, 1.074585. This means that the model's predictions are 1.07 units off, and it penalizes larger errors. LR and KNN have RMSEs of 1.337163 and 1.549953, respectively, indicating that their predictions are 1.34 and 1.55 units off.

When looking at the MAE in Fig. 5, the RF model again outperformed the other models with an average error of approximately 0.796568. This means that the model's predictions are, on average, 0.796568 away from the actual values, regardless of the direction of the errors. The MAE for the LR model is 1.025122, and for KNN, it is 1.349593.

Considering the R^2 metric, the RF model also had the highest score of 0.861561, indicating that the model can explain approximately 86.16% of the variation in UHPC flowability. The LR model and KNN model had R^2values of 0.785639 and 0.711986, respectively, implying that these models explain approximately 78.56% and 71.20% of the variance in the target variable. In conclusion, based on the RMSE, MAE, and R^2metrics, the RF model appears to be the most accurate and reliable model for predicting UHPC flowability among the three evaluated models.

3.2 Deep learning methods

The results of evaluating the DL methods are presented in the following Table 4:

Table 4. Comparison of DL methods.

Model	RMSE	MAE
ANN	2.469604	2.064296
LSTM	1.625995	1.243097
GRU	2.172126	1.682033
CNN	1.850219	1.595567

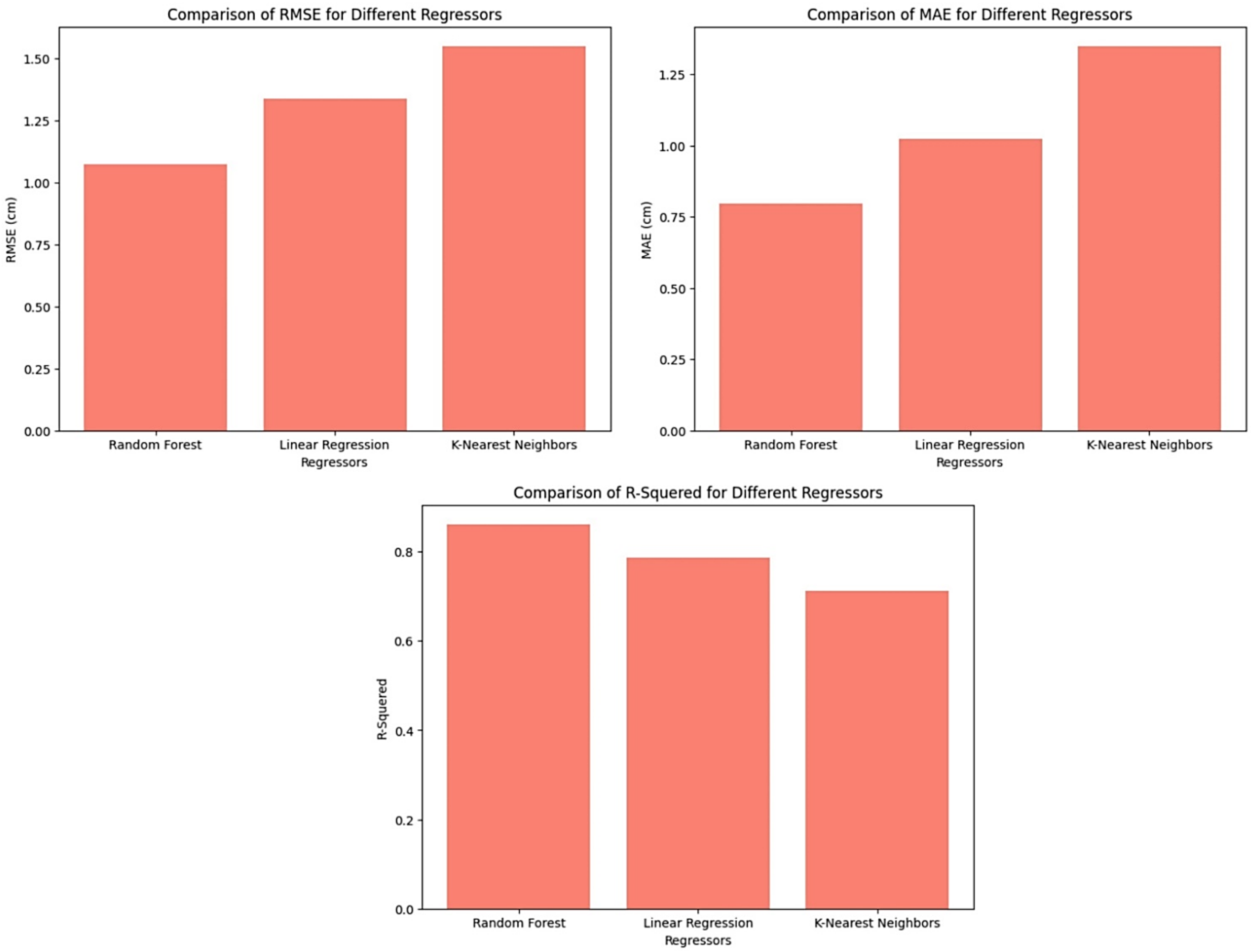

Fig. 5 Comparison of metrics for different ML methods.

Table 4 above outlines the evaluation metrics (RMSE, MAE) for four DL models: ANN, LSTM, GRU, and CNN, all used to predict UHPC flowability. The LSTM model achieved the lowest RMSE score of 1.625995, indicating an average deviation of around 1.63 units between its predictions and the actual data. The ANN model has the greatest (RMSE) of 2.469604, suggesting that its predictions differ by around 2.47 units. For the MAE metric, the LSTM model again demonstrated superior performance with an MAE of 1.243097. The LSTM model's predictions were 1.24 units off on average. The ANN model has the greatest MAE of 2.064296, indicating a 2.06-unit departure from actual data. Finally, the LSTM model predicts UHPC flowability best among DL models based on RMSE and MAE measures.

3.3 Hybrid-deep learning methods

Results of hybrid DL evaluation:

Table 5. Comparison of Hybrid DL methods.

Model	RMSE	MAE
CNN-LSTM	1.896753	1.582222
CNN-GRU	1.360066	1.036573

Table 5 compares the UHPC flowability predictions of two hybrid DL models. CNN-GRU surpassed CNN-LSTM in the RMSE statistic, which measures prediction error average. Median CNN-GRU model predictions were 1.36 units off with an RMSE of 1.360066. The CNN-LSTM model got an RMSE of 1.896753, indicating a 1.90-unit prediction error. The CNN-GRU model outperformed the CNN-LSTM model in MAE. The CNN-GRU model's predictions were 1.04 units off, regardless of error direction, with an MAE of 1.036573. The CNN-LSTM model had an MAE of 1.582222, indicating a 1.58-unit variation from actual data.

Based on the RMSE and MAE metrics, the CNN-GRU model appears to provide the most accurate and reliable predictions for UHPC flowability among the evaluated hybrid DL models.

3.4 Feature selection

FS-ML

Table 6 presents the evaluation metrics for the FS-ML models used to predict UHPC flowability.

The RF model achieved the lowest RMSE of 1.032841 and MAE of 0.767066 among the three models as shown in Fig. 6. This indicates that the (RF) model had a minor average difference between the predicted and actual flowability values,

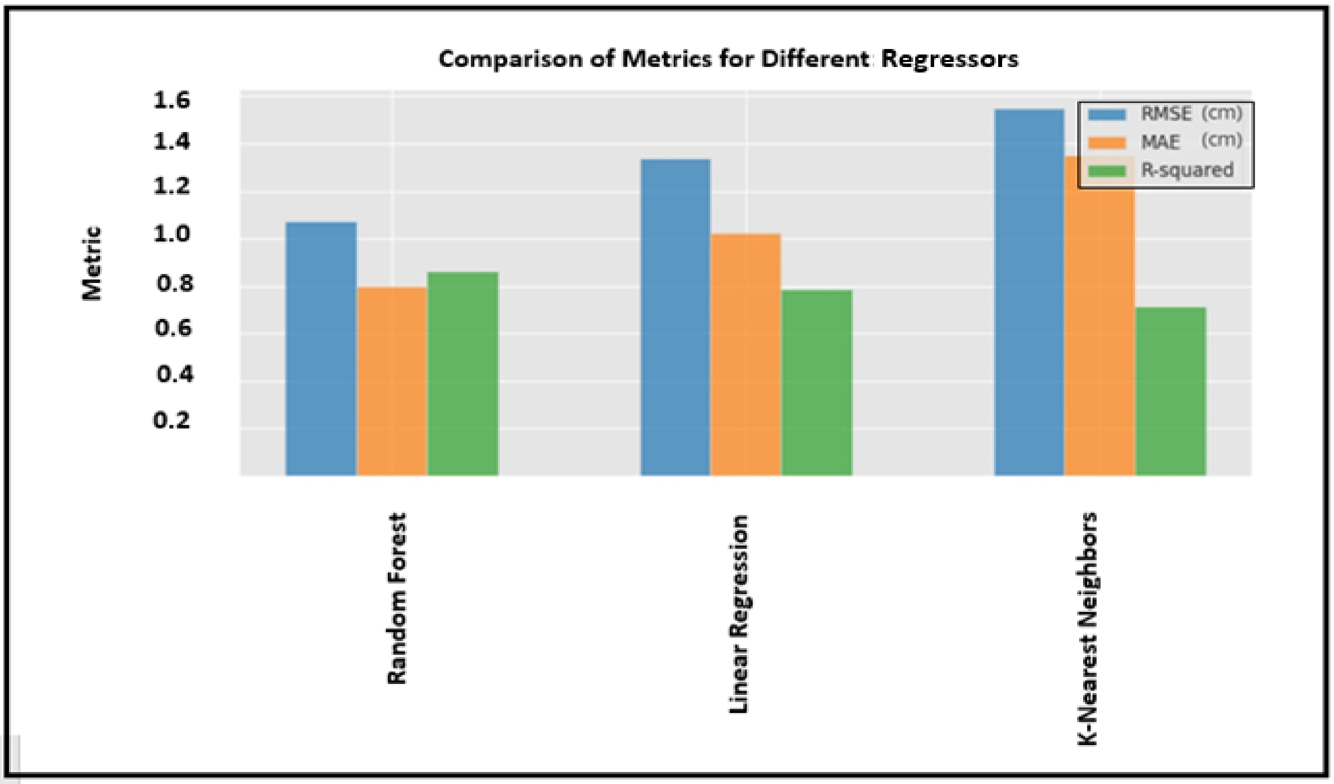

Fig. 6 FS-ML technique metrics comparison.

reflecting its high accuracy. Furthermore, the (RF) model exhibited a high R^2 value of 0.872108, suggesting that it explained approximately 87.2% of the variance in the UHPC flowability.

Table 6. Comparison of (FS-ML) methods.

Regressor	RMSE	MAE	R^2
RF	1.032841	0.767066	0.872108
KNN	1.549371	1.313218	0.712202
LR	2.046212	1.560119	0.498030

The KNN and LR models demonstrated comparatively higher RMSE values of 1.549371 and 2.046212, respectively. The KNN model had a higher MAE of 1.313218, while the (LR) model had an MAE of 1.560119. The (RF) model predicted UHPC flowability better than both models. The (kNN) and (LR) models fit the data less than the (RF) model, with values of 0.712202 and 0.498030, respectively.

FS-DL/FS-Hybrid DL

Table 7 shows the evaluation metrics for the UHPC flowability prediction models FS-DL and FS-Hybrid DL. The ANN model had the highest RMSE (4.258064) and MAE (3.098689). Fig. 7 shows that the ANN model predicted UHPC flowability less accurately due to a higher average discrepancy between projected and actual flowability values. The LSTM model performed better with a reduced RMSE of 1.387654 and MAE of 1.015121. A small average difference between projected and actual flowability values showed that LSTM predicted UHPC flowability better than the other models. UHPC flowability was predicted well by the GRU model with an RMSE of 2.172126 and MAE of 1.682033.

Table 7. Feature selection DL/Hybrid DL comparison.

Model	RMSE	MAE
ANN	4.258064	3.098689
LSTM	1.387654	1.015121
GRU	2.172126	1.682033
CNN	2.271201	1.840060
CNN-LSTM	1.896753	1.582222
CNN-GRU	1.650585	1.340550

The CNN model predicted UHPC flowability moderately well with RMSE of 2.271201 and MAE of 1.840060. CNN-LSTM and CNN-GRU hybrids outperformed DL models. CNN-LSTM recorded 1.896753 RMSE and 1.582222 MAE, whereas CNN-GRU had 1.650585 and 1.340550. CNN and LSTM/GRU strengths improved UHPC flowability prediction in these hybrid models.

The LSTM, CNN-LSTM, and CNN-GRU models outperformed the FS-DL and FS-Hybrid DL models in RMSE and MAE. Because they estimate UHPC flowability more accurately than other study models, these models are suggested.

The comparative analysis of our work with previous methods for predicting the CS of UHPC reveals several significant insights as shown in Table 8. Abuodeh *et al.* (2020) demonstrated the effectiveness of SFS and NID in determining essential material constituents for ANN, resulting in highly

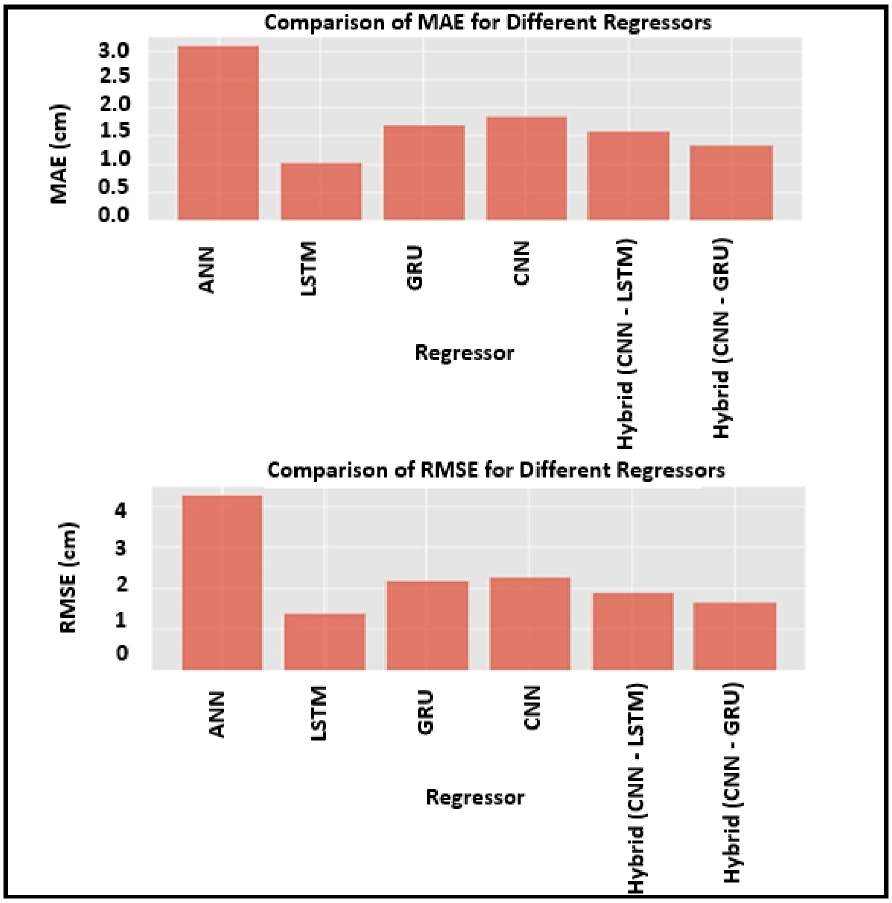

Fig. 7 Metric comparison for FS-DL/FS-Hybrid DL methods.

Table 8. Comparison with previous works.

Study Reference	Model	Dataset	Key finding	RMSE	MAE	R^2
Abuodeh *et al.* (2020) [12]	ANN	110 UHPC tests	SFS and NID helped determine important material constituents for ANN, resulting in more accurate predictions.	0.012	-	0.995
Choudhary *et al.* (2021) [13]	BPNN with SFS	110 data points, 8 materials	BPNN with selected features achieved higher accuracy compared to models with all features.	-	-	0.991
Shen *et al.* (2022)[18]	XGBoost	UHSC	XGBoost demonstrated higher accuracy in estimating UHSC CS compared to other methods. Curing time was identified as the most influential parameter.	6.4	7.6	0.90
Khan and Suthar (2023)[19]	RF	236 readings for UHPC	the RF model surpasses the M5P model in accurately predicting the 28-day CS of UHPC.	16.005	12.03	0.8568
Onyelowe *et al.* (2023c)[29]	AutoGluon	1133 net-zero concrete mix records	AutoGluon outperforms, achieving R2 of 92.6%; BFS significantly impacts strength, suggesting it as a sustainable cement alternative.	-	2.820	92.6%
Our Work	RF with FS	UHPC (135)	The CNN-GRU model exhibited the best predictive accuracy with a RMSE of 1.360066 and MAE of 1.036573. Additionally, FS techniques enhanced the performance of certain models, with the feature-selected RF model showing notable improvements in accuracy, achieving a RMSE of 1.032841 and MAE of 0.767066.	1.032841	0.767066	0.872108

accurate predictions. Choudhary *et al.* (2021) highlighted the superiority of BPNN with selected features over models with all features, emphasizing the importance of FS in enhancing predictive accuracy. Shen *et al.* (2022) found that XGBoost exhibited higher accuracy in estimating UHSC CS, with curing time identified as a crucial parameter. Khan and Suthar (2023) showcased the superiority of the RF model in accurately predicting UHPC CS compared to the M5P model. Onyelowe *et al.* (2023c) demonstrated the effectiveness of AutoGluon in achieving a remarkable R^2 of 92.6%, while also highlighting the significant impact of BFS on concrete strength, suggesting it as a sustainable alternative to cement. In our work, we found that our RF model, coupled with FS techniques, significantly enhanced predictive accuracy, achieving an impressive RMSE of 1.032841 and MAE of 0.767066. This underscores the importance of both model selection and feature engineering in optimizing predictive performance for UHPC CS estimation.

4. Conclusion

This work extensively compares ML, DL, and H-DL methods for UHPC flowability prediction. We created a data-driven pipeline from EDA and data preprocessing to model deployment and evaluation using systematic research. RF predicts UHPC flowability better than LR and KNN among non-feature-selected models (RMSE, MAE, and R^2). A greater R^2, lower RMSE, and MAE indicate a more accurate forecast. LSTM performed best among DL models (ANN, LSTM, GRU, CNN). CNN-GRU did best with the lowest RMSE and MAE scores among the Hybrid DL approaches. After FS integration, the (RF) classifier outperformed the other ML models. However, (LR) results worsened, suggesting overfitting or that the (FS) procedure may have excluded relevant factors.

In contrast, feature-selected DL models performed inconsistently. LSTM improved RMSE and MAE, making more accurate predictions, whereas ANN dropped significantly. Hybrid models (CNN-LSTM, CNN-GRU) performed somewhat worse than non-feature-selected models. The study emphasizes (FS) in model optimization and the need for a complete, systematic research approach to predictive model development. This study's comparative analysis can help future researchers choose UHPC flowability models. ML, DL, and H-DL models have demonstrated promising results, but they can be optimized by fine-tuning and investigating additional (FS) methods.

In conclusion, the selected ML, DL, and H-DL approaches may predict UHPC flowability, although (FS) and model choice greatly impact prediction accuracy. Future study should improve prediction models and apply more advanced methodologies in other construction scenarios.

4.1 Contributions of the Research

- This study shows that ML, DL, and H-DL can predict UHPC flowability, a crucial construction feature, helping civil engineering and construction. It has a robust EDA, preprocessing, model selection, assessment, and (FS) pipeline for similar tasks.
- This study's comparative analysis added to literature. This study tests several tactics in similar conditions to gain insights into model performance in real-world scenarios.
- The study emphasises FS's model performance impact. RFE was carefully integrated into each ML, DL, and H-DL approach to demonstrate its impact on prediction accuracy.

4.2 Limitations and future work

Despite its importance, the paper has some drawbacks. The assumption of linearity in (LR) models may not convey the complexity of UHPC flowability prediction. The (FS) procedure may also be limited. RFE may not have caught the greatest features for each model, as shown by their decreased performance.

Recommendations

- To accommodate non-linear input-output interactions, (FS) strategies and procedures could be researched and fine-tuned.
- Complex models like ensemble techniques or Transformer-based models may perform better.
- The investigation could also include other construction factors or concrete types. These models can also be used for structure health monitoring, materials selection, and safety prediction in civil engineering.

Conflict of Interest

There is no conflict of interest.

Supporting Information

Not applicable.

References

[1] D. Wang, C. Shi, Z. Wu, J. Xiao, Z. Huang, Z. Fang, A review on ultra-high-performance concrete: part II. Hydration, microstructure and properties, *Construction and Building Materials*, 2015, **96**, 368-377, doi: 10.1016/j.conbuildmat.2015.08.095.

[2] M. Zhou, W. Lu, J. Song, G. C. Lee, Application of ultra-high-performance Concrete in bridge engineering, *Construction and Building Materials*, 2018, **186**, 1256-1267, doi: 10.1016/j.conbuildmat.2018.08.036.

[3] X. Zhang, S. Zhao, Z. Liu, F. Wang, Utilization of steel slag in ultra-high-performance concrete with enhanced eco-friendliness, *Construction and Building Materials*, 2019, **214**, 28-36, doi: 10.1016/j.conbuildmat.2019.04.106.

[4] A. Arora, M. Aguayo, H. Hansen, C. Castro, E. Federspiel, B. Mobasher, N. Neithalath, Microstructural packing- and rheology-based binder selection and characterization for ultra-high-performance concrete (UHPC), *Cement and Concrete*

Research, 2018, **103**, 179-190, doi: 10.1016/j.cemconres.2017.10.013.
[5] A. Alsalman, C. N. Dang, W. Micah Hale, Development of ultra-high-performance concrete with locally available materials, *Construction and Building Materials*, 2017, **133**, 135-145, doi: 10.1016/j.conbuildmat.2016.12.040.
[6] D.-Y. Yoo, N. Banthia, Mechanical properties of ultra-high-performance fiber-reinforced concrete: a review, *Cement and Concrete Composites*, 2016, **73**, 267-280, doi: 10.1016/j.cemconcomp.2016.08.001.
[7] J. Abellán García, J. Fernández Gómez, N. Torres Castellanos, Properties prediction of environmentally friendly ultra-high-performance concrete using artificial neural networks, *European Journal of Environmental and Civil Engineering*, 2022, **26**, 2319-2343, doi: 10.1080/19648189.2020.1762749.
[8] H. El-Chabib, M. Nehdi, M. Sonebi, Artificial intelligence model for flowable concrete mixtures used in underwater construction and repair, *ACI Materials Journal*, 2003, **100**, 165-173, doi: 10.14359/12557.
[9] R. Jayaseelan, G. Pandulu, G. Ashwini, Neural networks for the prediction of fresh properties and compressive strength of flowable concrete, *Journal of Urban and Environmental Engineering*, 2019, 183-197, doi: 10.4090/juee.2019.v13n1.183197.
[10] Y. Qian, M. Sufian, O. Accouche, M. Azab, Advanced machine learning algorithms to evaluate the effects of the raw ingredients on flowability and compressive strength of ultra-high-performance concrete, *PLoS One*, 2022, **17**, e0278161, doi: 10.1371/journal.pone.0278161.
[11] S. Mahjoubi, W. Meng, Y. Bao, Auto-tune learning framework for prediction of flowability, mechanical properties, and porosity of ultra-high-performance concrete (UHPC), *Applied Soft Computing*, 2022, **115**, 108182, doi: 10.1016/j.asoc.2021.108182.
[12] O. R. Abuodeh, J. A. Abdalla, R. A. Hawileh, Assessment of compressive strength of ultra-high-performance concrete using deep machine learning techniques, *Applied Soft Computing*, 2020, **95**, 106552, doi: 10.1016/j.asoc.2020.106552.
[13] D. Choudhary, J. Keshari, I. A. Khan, Prediction of compressive strength of ultra-high-performance concrete using machine learning algorithms—SFS and ANN. Gao XZ, Tiwari S, Trivedi M, Mishra K, Advances in Computational Intelligence and Communication Technology. Singapore: Springer, 2021: 17-27.10.1007/978-981-15-1275-9_2.
[14] D. Fan, R. Yu, S. Fu, L. Yue, C. Wu, Z. Shui, K. Liu, Q. Song, M. Sun, C. Jiang, Precise design and characteristics prediction of ultra-high-performance concrete (UHPC) based on artificial intelligence techniques, *Cement and Concrete Composites*, 2021, **122**, 104171, doi: 10.1016/j.cemconcomp.2021.104171.
[15] Y. Asghari, G. Sadeghian, S. E. Mohammadyan-Yasouj, E. Mirzaei, Forecast of modern concrete properties using machine learning methods. Momeni E, Jahed Armaghani D, Azizi A, Artificial Intelligence in Mechatronics and Civil Engineering. Singapore: Springer, 2023.
[16] Z. Nurlan, A novel hybrid radial basis function method for predicting the fresh and hardened properties of self-compacting concrete, *Advances in Engineering and Intelligence Systems*, 2022, **1**, 50-64, doi: 10.22034/aeis.2022.148305.
[17] E. Sadrossadat, H. Basarir, A. Karrech, M. Elchalakani, Multi-objective mixture design and optimisation of steel fiber reinforced UHPC using machine learning algorithms and metaheuristics, *Engineering with Computers*, 2022, **38**, 2569-2582, doi: 10.1007/s00366-021-01403-w.
[18] Z. Shen, A. F. Deifalla, P. Kamiński, A. Dyczko, Compressive strength evaluation of ultra-high-strength concrete by machine learning, *Materials*, 2022, **15**, 3523, doi: 10.3390/ma15103523.
[19] M. Aqeel Khan, M. Suthar, Assessing the results of compressive strength of ultra-high-performance concrete using soft computing, *IOP Conference Series: Earth and Environmental Science*, 2023, **1110**, 012089, doi: 10.1088/1755-1315/1110/1/012089.
[20] A. Marani, A. Jamali, M. L. Nehdi, Predicting ultra-high-performance concrete compressive strength using tabular generative adversarial networks, *Materials*, 2020, **13**, 4757, doi: 10.3390/ma13214757.
[21] B. Xi, E. Li, Y. Fissha, J. Zhou, P. Segarra, LGBM-based modeling scenarios to compressive strength of recycled aggregate concrete with SHAP analysis, *Mechanics of Advanced Materials and Structures*, 2023, 1-16, doi: 10.1080/15376494.2023.2224782.
[22] Y. Qian, M. Sufian, A. Hakamy, A. Farouk Deifalla, A. El-said, Application of machine learning algorithms to evaluate the influence of various parameters on the flexural strength of ultra-high-performance concrete, *Frontiers in Materials*, 2023, **9**, 1114510, doi: 10.3389/fmats.2022.1114510.
[23] Al Sarfin, Md Abdullah, Srishti Banerji, Mohsen Zaker Esteghamati, and Andrew Sorensen. Prediction of Compressive Strength of UHPC Mix Designs Based on Mix Components Using Machine Learning Techniques. In International Interactive Symposium on Ultra-High Performance Concrete, vol. 3, no. 1. Iowa State University Digital Press, 2023.
[24] A. Ibrahim Bibi Farouk, J. Zhu, J. Ding, S. I. Haruna, Prediction and uncertainty quantification of ultimate bond strength between UHPC and reinforcing steel bar using a hybrid machine learning approach, *Construction and Building Materials*, 2022, **345**, 128360, doi: 10.1016/j.conbuildmat.2022.128360.
[25] P. Kumar, G. Nakkeeran, K. C. Onyelowe, L. Krishnaraj, Comparative study on net-zero masonry walls made of clay and fly ash bricks and grouts/mortars/stuccos with the effect of super fine fly ash blended cement-low carbon cement, *International Journal of Low-Carbon Technologies*, 2023, **18**, 1008-1014, doi: 10.1093/ijlct/ctad087.
[26] M. Ravi, B. Murugesan, K. C. Onyelowe, Performance evaluation of marine and industrial wastes in cement to envelope low carbon environment in manufacturing process, *International Journal of Low-Carbon Technologies*, 2023, **18**, 986-998, doi: 10.1093/ijlct/ctad082.
[27] K. C. Onyelowe, J. Jagan, D.-P N. Kontoni, A. Ali Baig Moghal, I. C. Onuoha, R. Viswanathan, D. K. Soni, Utilization of

GEP and ANN for predicting the net-zero compressive strength of fly ash concrete toward carbon neutrality infrastructure regime, *International Journal of Low-Carbon Technologies*, 2023, **18**, 902-914, doi: 10.1093/ijlct/ctad081.
[28] K. C. Onyelowe, A. M. Ebid, The influence of fly ash and blast furnace slag on the compressive strength of high-performance concrete (HPC) for sustainable structures, *Asian Journal of Civil Engineering*, 2024, **25**, 861-882, doi: 10.1007/s42107-023-00817-9.
[29] K. C. Onyelowe, D.-P N. Kontoni, S. R. M. Pilla, S. Hanandeh, A. M. Ebid, M. Razzaghian Ghadikolaee, L. U. Stephen, Runtime-based metaheuristic prediction of the compressive strength of net-zero traditional concrete mixed with BFS, FA, SP considering multiple curing regimes, *Asian Journal of Civil Engineering*, 2024, **25**, 1241-1253, doi: 10.1007/s42107-023-00839-3.
[30] K. C. Onyelowe, A. M. Ebid, S. Hanandeh, The influence of nano-silica precursor on the compressive strength of mortar using Advanced Machine Learning for sustainable buildings, *Asian Journal of Civil Engineering*, 2024, **25**, 1135-1148, doi: 10.1007/s42107-023-00832-w.