Metaheuristic-Optimized Decision Tree Models Using Catch Fish and Hummingbird Algorithms for Predicting Creep Coefficients in Ultra-High-Performance Concrete

Chao Wang, Hao Wang*
Department of Architecture and Design, Qinhuangdao Vocational and Technical College, Qinhuangdao City, Hebei Province, 066100, China
E-mail: Wanghao816p@163.com
*Corresponding author

Keywords: capability, machine learning, creep behavior, decision tree, ultra-high-performance concrete, catch fish optimization

Received: September 27, 2025

Machine learning (ML) was used to study the creep behavior of ultra-high-performance concrete ($UHPC$). This study employed Decision Tree (DT), to develop strong methods for forecasting for the creep coefficient (CC) utilizing a complete dataset acquired from literature. The reliability of DT is greatly affected by the choice of hyperparameters, that must be determined using metaheuristic optimization techniques. This is accomplished via the employment of the Catch Fish Optimization Algorithm ($CFOA$) and the Artificial Hummingbird Optimization Algorithm ($AHOA$). A complete dataset of 338 points of data was gathered from published research for the purpose of constructing the model. According to the statistics, $DT(CFOA)$ and $DT(AHOA)$ are both probably able to forecast CC with accuracyEstimated assessment of performance measures indicates that DT(CFOA) is a more precise and dependable methodology than DT(AHOA).In both the training and assessment stages, $DT(AHOA)$ receives lower R^2 values (0.98123 and 0.97342), but $DT(CFOA)$ obtains higher R^2 values (0.98978 and 0.98475).

Povzetek: Študija je pokazala, da optimizirano odločitveno drevo s CFOA natančneje napoveduje koeficient lezenja UHPC kot različica z AHOA, saj dosega višje vrednosti R².

1 Introduction, review, and novelty

$UHPC$ is an intriguing substance due to its outstanding Mechanical Features (MFs) and toughness features [1], [2]. Cr is crucial for forecasting and evaluating the MFs of concrete buildings under sustained loads over short and extended durations. Cr is a major agent in the extended period mechanical function of concrete structures, leading to considerable time-dependent engineering issues, namely, too much flexing and reduced tension in bridges [3]. A few scientists have merged the impacts of varying factors on the Cr performance of $UHPC$ and created Cr models for $UHPC$ [4], [5]. Research performed on $UHPC$ has demonstrated a feasible and functional method for analyzing Cr behavior [6]. Nonetheless, these approaches exhibit disadvantages such as elevated costs, protracted durations, and distinct substance attributes. Hence, other techniques like as ML have been examined for estimating the Cr actions of $UHPC$, which can procedure for larger and more diverse data collections than hypothetical frameworks. Huang et al. created and studied the influences of various elements, namely, fiber kind and amount, curing states, W/B, loading magnitude, and loading duration, on the Cr behavior of UHPC. An increase in the W/B, strength, elastic modulus, or sustained LT might result in a growth in Cr, but a rise in fiber and aggregate quantity or CT may hinder Cr development [7], [8]. ML techniques have been employed to forecast concrete variables, including resistance to compression, shear resistance, and defeat resistance [9], [10], [11], [12], [13]. Nunez et al. performed an extensive review of ML strategies for estimating the MFs of concrete. 7 ML computations were examined: Support Vector Machines ($SVMs$), Artificial Neural Network (ANN), Genetic Algorithms, Fuzzy Logic, Tree-based Ensembles, Deep Learning, and Hybrid Procedures. These techniques were utilized to estimate the durability of HPC, Recycled Aggregate Concrete (RAC), Self-Compacting Concrete (SCC), and also various other kinds of concrete [14]. Moreover, research employs ML algorithms to forecast the Cr characteristics of concrete. Present-day ML applications in predicting Concrete Creep (CC) mostly concentrate on Back Propagation (BP) neural networks. developed an ANN model that used the NU DB to estimate the creep behavior of concrete.[15]. Karthikeyan et al. foresaw the CC of UHPC by assessing the outcomes of the CEB 90 Cr model, and creating a Neural Network (NN) model [16]. Hodhod et al. formulated a precise CC equation and used genetic methods to reduce the difference between the estimated quantities from the

equation and the actual measurements, thus improving forecasting precision [17]. Gandomi et al. employed multi-objective genetic programming to create a clear formula for the CC by means of the provided DB [18]. To accurately identify the best HPs for models, namely, Least Squares Support Vector Machine (LS-SVM) and XGBoost in forecasting RAC Cr, Feng et al. used grid search and k-fold methodologies. The XGBoost-based model showed ACC and effectiveness in making multi-input predictions [19]. Xiao et al. created a NN model to forecast the Cr of recycled concrete, considering the impact of recycled Coarse Aggregate (CA) substitution percentages, using the RILEM B3 Cr model as a foundation [20]. Li et al. utilized a grid search approach to improve Cr forecasting models, thereby enhancing ACC in estimation. The support vector regression and fully connected neural network models both showed improved capabilities in predicting Cr levels. [21]. The ML model showed its possibilities in anticipating diverse actions of engineering problems [22]. Nonetheless, it continues to function as a black box model. Previous research included approaches such as attribute importance evaluation, grey relational evaluation, and variable sensitivity evaluation to elucidate the interpretability of diverse ML models. A study employing the GPR model and GJO metaheuristic the improvement yielded the most precise estimation of pile capacity for bearing (Pu) with a R^2 value of 0.994, indicating a viable option for optimal structural design and safety [23]. Another study introduced a reliable and cost-effective approach for predicting triaxial compressive strength from the physical-mechanical characteristics of frozen soils, utilizing the $ANFIS$ model alongside two ANN and SVM approaches [24]. A study presents a precise framework for estimating the compression resistance of SCC by integrating the SVM algorithm with sophisticated optimization techniques. The suggested model demonstrates superior performance relative to conventional techniques, with a R^2 of 97.3% and an $RMSE$ of 3.81 MPa [25].

The purpose of this research is to investigate the CC of $UHPC$. $UHPC$ combined with typical CA is a promising material because it has great MFs, is durable, and is less expensive than other materials. Three alternatives ML models were created using the data that was gathered during an experiment. The DT was used to calculate the information provided above. The DT HPs have a significant effect on its dependability, and they must be selected using MO approaches. This is accomplished by utilizing the $CFOA$ and the $AHOA$. The eight input components for the CC are generated using a mathematical DB that contains 338 examination findings from publications. The study's results enhance $UHPC$ mix designs for greater structural endurance, hence decreasing material expenses and maintenance requirements in bridges, skyscrapers, and precast components. Precise Cr forecasts improve performance-oriented building regulations, guaranteeing safer and more sustainable infrastructure. Furthermore, the incorporation of ML models with intelligent sensors allows predictive maintenance and digital twin simulations for sustained structural monitoring.

One of the attributes of the suggested model is its capability to handle changing and noisy data. The application of $CFOA$ and $AHOA$ metaheuristic algorithms establish a dynamic equilibrium between discovery and utilization, enabling the model to adapt to intricate and non-stationary data patterns while circumventing premature convergence. This property resembles discoveries in adaptive and nonlinear control, which demonstrate that optimization tactics can guarantee system robustness and stability amid uncertainty. The proposed framework demonstrates effective performance in accurately forecasting the creep of high-resistance concrete and possesses the capacity to generalize across various datasets with differing environmental situations, a crucial consideration for practical applications in real-world scenarios.

Enhancing the precision of creep forecasting in $UHPC$ directly influences the refinement of design codes, durability forecasting, and infrastructure optimization. Consistent with recent studies in adaptive control and neural-fuzzy learning that have achieved stability and diminished uncertainty through Lyapunov principles, the model presented in this research has illustrated superior precision and dependability compared to regional models by employing metaheuristic methods [26], [27], [28]. In comparison to regional techniques like gradient boosting and random forest, the boosted decision tree model in this work offers boosted transparency and interpretability while preserving accuracy, making it suitable for practical application in the construction and management of structures [29], [30], [31].

The primary issue of this study is the accurate estimation of the creep coefficient (CC) of ultra-high-strength concrete ($UHPC$) based on eight essential mix properties and loading situations, an issue exacerbated by data nonlinearity, a restricted sample size, and variable interactions. The objective of the project is to enhance estimation precision by creating decision tree models optimized with the Catch Fish and Artificial Hummingbird algorithms to address the shortcomings of traditional approaches.

Table 1. shows a summary of the works related to the present study.

Table 1: summary of the works related to the present study

Refs.	Dataset Size	Model Type	Optimization Approach	Performance Metrics	Limitations Identified	Contribution of Present Study
Bal & Buyle-Bodin (2014) [15]	200+	ANN	None	$R^2 \approx 0.92$	Black-box, low interpretability	
Karthikeyan et al. (2008) [16]	150	ANN	None	RMSE ≈ 0.15	Limited dataset, overfitting	
Hodhod et al. (2018) [32]	250	ANN + Genetic Programming	Evolutionary Optimization	$R^2 \approx 0.95$	Accuracy improved but interpretability lacking	
Gandomi et al. (2016) [17]	300	Genetic Programming	Multi-objective optimization	$R^2 \approx 0.96$	Complex models, poor transparency	
Feng et al. (2022) [18]	400	LS-SVM, $XGBoost$	Grid Search, k-fold	$R^2 \approx 0.97$	Strong accuracy but black-box, weak sensitivity analysis	
Zhu et al. (2024) [46]	500	$XGBoost$, Interpretable ML	Hyperparameter tuning	$R^2 \approx 0.97$	Limited focus on UHPC creep	
Present Study	338	Decision Tree	CFOA, AHOA	R^2 = 0.989 (train), 0.985 (test)	Improves accuracy, interpretable, includes sensitivity analysis	✓

Table. 1 summarizes those previous models ($ANN, SVM, XGBoost$, etc.) shown commendable estimation performance but typically encountered challenges related to restricted interpretability, difficulties in managing tiny and nonlinear datasets, or an absence of feature-level sensitivity examination. The $DT-CFOA$ and $DT-AHOA$ models rectify these shortcomings by integrating interpretable decision trees with advanced metaheuristic enhancement, yielding superior estimative precision, increased resilience, and greater clarity regarding attribute contributions.

2 Methodology

2.1 Collected data, properties, and preparation

Several techniques exist for forecasting the Cr action agent of $UHPC$, with ML emerging as a prevalent way due to its speed, ACC, and cost-effectiveness; thus, it is employed in this investigation. The employment of ML in estimating diverse civil engineering agents is underpinned by multiple scientific rationales, notably the intricate and nonlinear properties of the data, rendering this technique particularly apt for such predictions. In recent years, advancements in sensors and measuring tools have resulted in the generation of substantial data in building projects, which may be examined and utilized for estimative purposes through ML methodologies. Moreover, ML models typically exhibit superior precision compared to conventional analytical or mathematical approaches [33], particularly in empirical and data-driven contexts. ML is an effective instrument for estimating, examining, and decision-making in civil projects, enhancing the precision, speed, and sophistication of engineering operations. This strategy facilitates the transition from conventional engineering to data-driven engineering. Initially, a dataset is required to conduct prediction activities. This article utilized 338 data points gathered from reputable technical literature [34], [35]. Gathering data from scientific papers for estimating, particularly in civil engineering and ML, is a meticulous, systematic, and intentional procedure. Choosing appropriate data necessitates various considerations, including: precisely articulating the prediction aim, recognizing reliable and pertinent sources, formulating a systematic quest tactic, extracting data, normalizing and cleansing data, and estimating data quality, among others. The data must be segmented, followed by the execution of actions on it. Dividing the information into two segments, Train and Test (and occasionally Validation), is a fundamental notion in ML and forecasting. This procedure is conducted to assess the model's actual efficacy. It

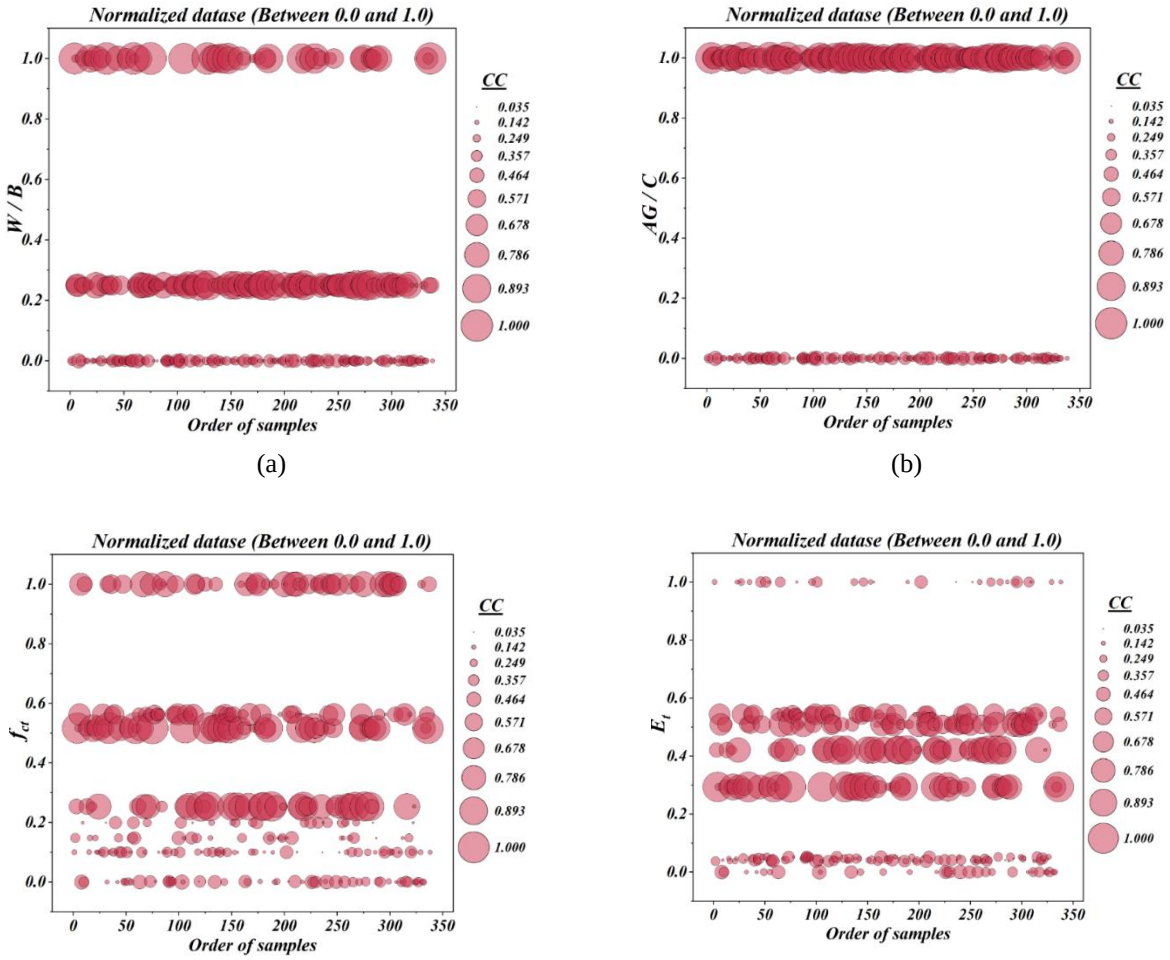
additionally mitigates overfitting and forecasts the simulation under actual circumstances. In this work, subsequent to the assessments, the dataset was partitioned into two segments: test & and train, comprising 25% & 75% of the total, correspondingly. This technique enhances the model's robustness and precision for real applications. The data normalization operation yields a more desirable and precise outcome. Data normalization, or scaling, is a crucial preprocessing step in ML and data evaluation, undertaken for many theoretical and practical purposes. The purposes of normalization are to standardize agent scales, boost ML technique efficiency, augment model correctness, and mitigate extreme fluctuations. Four-fold Cross-validation is an approach to statistics utilized to evaluate the efficacy of an ML model, aiding in the comprehension of the model's generalization to novel data. This approach involves partitioning the data into Train and Test numerous times, rather than doing a single division. This method involves partitioning the complete data collection into four roughly equal segments (folds). At every phase, one segment is designated as test data ($Test$), while the remaining three parts serve as training data ($Train$). This procedure is executed four times, ensuring that a distinct component is chosen for evaluation each time. The mean efficiency of these four phases is ultimately computed. This strategy offers enhanced assessment precision, diminished reliance on arbitrary partitioning, and improved data utilization. This strategy is employed when the data volume is moderate to substantial, when a fair and accurate model evaluation is required, and when the processing demands are manageable. The quantity of input parameters and their unique attributes, as determined by legitimate assessment metrics, are encapsulated in Table 2. This paper comprises eight input factors and one output factor, presented in the following sequence: W/B $(\frac{W}{B})$, AG/C $\left(\frac{AG}{C}\right)$, compression resistance at loading age (f_{ct}), EMLA (E_t), SFVC (SF), CT (T_c), RH, loading duration (t), and CC. The assessment of the table indicated that the Train data predominantly exhibit a rather normal distribution; nevertheless, multiple variables, such as W/B and SF, demonstrate considerable skewness. The factor t has considerable dispersion. The assessment of the Train indicated that multiple variables (such as W/B, SF, and t) exhibit a distribution distinct from that of the training data. The entire dataset has commendable diversity; yet, certain variables display an imbalanced distribution.

Table 2: Statistical summary of input and output factors for CC modeling of fiber-reinforced concrete

Index	W/B	AG/C	f_{ct}	E_t	SF	T_c	RH	t	CC
	–	–	MPa	GPa	%	$°C$	%	$Days$	–
	Water to binder	Aggregate to cement	CSLA	EMLA	SFVC	CT	RH	Loading duration	CC
	Data for learning model								
Min.	0.14	1.777	129.21	40	0	21	95	0.5197	0
Max.	0.22	2.463	170.6	60.43	2	75	100	179.55	0.7019
St. D.	0.0264	0.3425	13.166	5.8694	0.6851	26.959	2.4962	47.297	0.1826
Var.	0.0007	0.1173	173.34	34.449	0.4693	726.79	6.231	2237	0.0333
Skew.	1.5615	-0.11	0.8357	0.728	-1.577	0.1104	0.1104	0.9708	0.3921
Kurt.	1.1841	-2.004	-0.384	-0.018	0.9864	-2.004	-2.004	-0.003	-0.843
Avg.	0.1591	2.1389	143.95	46.718	1.6299	46.512	97.362	53.996	0.3022
Med.	0.16	2.463	139.7	46	2	21	95	42.146	0.2769
	Data for examining model								
Min.	0.14	1.777	129.21	40	0	21	95	0.4659	0
Max.	0.22	2.463	170.6	60.43	2	75	100	178.57	0.6843
St. D.	0.0242	0.3406	13.897	6.5807	0.6281	26.808	2.4822	46.099	0.1849
Var.	0.0006	0.116	193.14	43.306	0.3946	718.67	6.1614	2125.1	0.0342

Skew.	1.9103	0.2398	0.9943	0.781	-2	-0.24	-0.24	1.3246	0.6071
Kurt.	2.7949	-1.988	-0.236	-0.307	2.7064	-1.988	-1.988	0.8906	-0.825
Avg.	0.1552	2.0792	142.35	46.686	1.7143	51.214	97.798	50.824	0.2738
Med.	0.14	1.777	136.39	46	2	75	100	40.567	0.2252

Another method to boost comprehension of variable features is through visual tools such as statistics charts. They facilitate the revelation of concealed patterns and associations among variables and detect anomalies. They meticulously examine the distribution of data, offering enhanced comprehension of several variables concurrently. An illustration of this tool is shown in the collection of charts in Figure 1. The data in this figure is normalized. For instance, examining part (i) of this compilation revealed that this graph is a bubble scatter plot of normalized data illustrating the distribution of the desired parameter, CC, for each sample in the data collection. The X-Axis signifies the sequence of data entry, while the Y-Axis denotes the normalized value of the CC for every collection. The dimensions of each bubble represent the real number of CC. The evaluation indicates that the majority of the data exhibit median normalized CC. Significant bubbles are predominantly observed near the highest part of the $Y - axis$. Generally, significant diversity in the size and quantity of CC signifies abundant data for training a generalizable model.



(a) (b)

(c) (d)

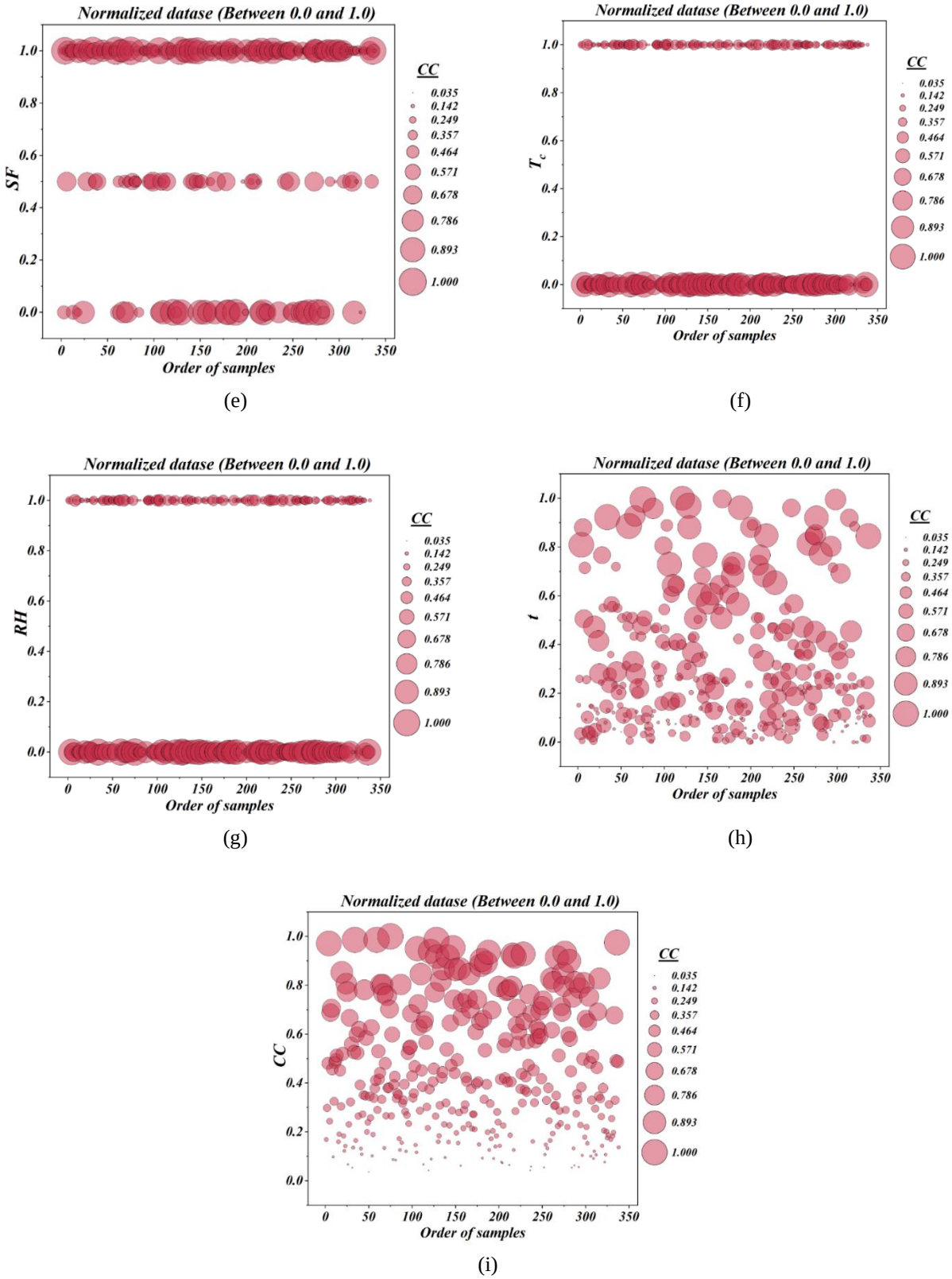


Figure 1: Visualization of data distribution in normalized form: emphasizing the dispersion of features within the dataset

$PAWN$ sensitivity evaluation is a contemporary and robust technique for assessing the relative impact of agents in intricate models, including ML models. This approach clarifies the elements that most significantly affect the result of the agent (CC in the CC model). Figure 2 illustrates an example of this sensitivity analysis. In

$PAWN$, the influence of every number on the output is assessed by contrasting the complete outcome distribution with the conditional outcome distribution when an agent is held constant. This method offers a number of benefits: it is not model-driven, it accommodates nonlinearity and non-parametricity, it employs a straightforward computation technique, and it is appropriate for costly models. An elevated $PAWN$ value (approaching 1) signifies that this agent is highly sensitive and influential on the outcome, while a low $PAWN$ number (approaching 0) denotes minimal or negligible impact. Typically, numbers between 0.1 and 0.9 are considered significant, warranting greater focus on the relative comparison among agents. The analysis of the figure revealed that each radius in the graph represents an input variable in the model, with the height or length of the radius signifying the $PAWN$ sensitivity index for that variable, and the number above each axis indicating the $PAWN$ sensitivity quantity for that agent (ranging from 0 to 1). The t agent has the maximum sensitivity, followed by f_{ct} and E_t, which both demonstrate relatively high sensitivity in the later levels. The T_c and RH agents exert a moderate influence, while W/B and AG/C also demonstrates a medium sensitivity level.

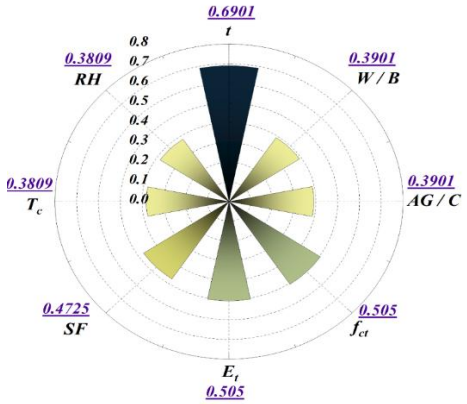


Figure 2: Evaluation of parameter influence using $PAWN$ sensitivity indices

Spearman's rank correlation serves as yet another instance of a graph that effectively illustrates the features of agents. Spearman's correlation coefficient quantifies the extent of a monotonic connection between two agents, without presuming linearity in their association. This graph is typically presented as a heatmap. This chart serves as an effective instrument for examining nonlinear yet monotonic correlations among agents. In contrast to Pearson, it does not necessitate a normal distribution and exhibits reduced sensitivity to outliers, rendering it the preferred option for experimental data in fields such as civil engineering, geotechnics, environmental studies, and intricate ML models. A representation of this chart is illustrated in the Figure 3. The examination of this chart reveals that darker hues signify strong positive correlations (up to +1), whereas lighter hues (yellow) denote significant negative correlations (up to -1). The parameters W/B and AG/C exhibit the strongest positive association at +0.953, whereas AG/C and T_c demonstrate the most significant negative correlation at -1. Moreover, AG/C in regard to f_{ct} exhibits a strong positive correlation of +0.871, while SF in relation to f_{ct} and E_t has significant negative correlations of -0.322 and -0.404, in the same order.

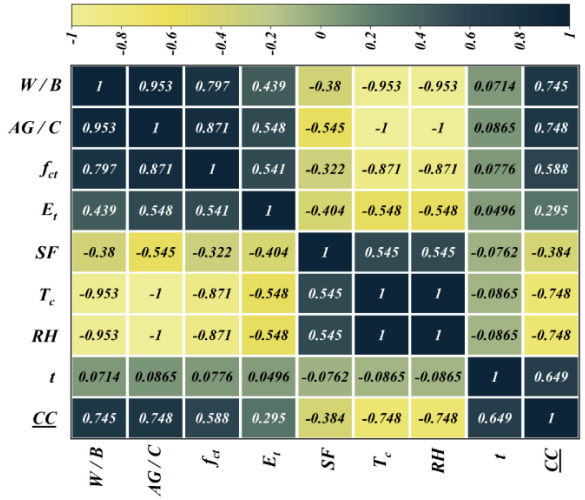


Figure 3: Spearman rank correlation analysis of model parameters

2.2 Estimation and optimization algorithms

2.2.1 Decision Tree (DT)

DTs forecast the worth of a factor via creating decision rules based on its attributes. A more complex tree leads to more elaborate decision rules and a more powerful model. DTs necessitate minimal data preparation in contrast to other methodologies that demand data standardization & several transformations. It employs a white box model, facilitating a clear elucidation of the circumstances through Boolean logic. Nonetheless, DTs can generate excessively intricate structures that overfit and fail to ensure effective generalization to the information [36]. According to [36], defining a maximum depth for the trees lessened the problem. The source data is consistently split to a pair of minor segments, referred to as nodes, according to an arrangement of decision criteria up to the point where a halt condition is fulfilled, specifically, the highest depth of the tree. Cases that have close target numbers are put in the same group, and the outcome is shown by the average of the y variable in each group. A tree can be seen as a consistent estimation across various segments. The choice of element and division placement is made to attain the optimal alignment. Denote the data at node m as R_m, comprising nm examples. For each applicant $split\ \theta = (j, t_m)$, comprising an attribute j and a threshold t_m, the data is divided into sub collections $R_m^{left}(\theta)$ and $R_m^{right}(\theta)$ such that:

$$R_m^{left}(\theta) = \{(x, y) \mid x_j \leq t_m\} \quad (1)$$

$$R_m^{right}(\theta) = \frac{R_m}{R_m^{left}(\theta)} \quad (2)$$

The standard of a candidate's division at the node m is subsequently determined employing a Loss Function (LF) H:

$$G(R_m, \theta) = \frac{n_m^{left}}{n_m} H\left(R_m^{left}(\theta)\right) + \frac{n_m^{right}}{n_m} H\left(R_m^{right}(\theta)\right) \quad (3)$$

The variables are chosen to minimize the LF.

$$\theta^* = \arg \min_{\theta} G(R_m, \theta) \quad (4)$$

This procedure is conducted repeatedly for the 2 sub-collections until the optimal depth is attained.

2.2.2 Catch fish Optimization algorithm ($CFOA$)

The $CFOA$ replicates the fishing practices of local fishermen. To facilitate the capture of fish, Fisherman (Fsh) select several fishing techniques. Like other metaheuristic algorithms (MAs), $CFOA$ comprises three separate phases: setting up, discovery, and utilization.

Setting up stage

The matrix F denotes the positional data of N quest factors within a d-dimensional zone, as seen in the formula as outlined below:

$$F = \begin{bmatrix} F_{1,1} & F_{1,2} & F_{1,n} \\ F_{2,1} & F_{2,2} & F_{2,n} \\ \dots & \dots & \\ F_{n,1} & F_{n,2} & F_{n,n} \end{bmatrix}_{N \times d} \quad (5)$$

$$F_{i,j} = lb_j + \left(ub_j - lb_j\right) \times rand \quad (6)$$

The matrix F denotes the positional data of N quest factors in a d-dimensional area. The formula for its setup is as the following: $F_{i,j}$ signifies the location of the i-th factor in the j-th dimension, with ub_j and lb_j indicating the upper and lower bounds of the j-th dimension, in the same order, and $rand$ is a stochastic factor within the range (0,1).

Utilizing the present positional data of each Fsh, a fitness assessment is performed on $fobj$ to determine their fitness rescores, which results in the subsequent fitness matrix.

$$f = fobj(F) = \begin{bmatrix} f_1 \\ f_2 \\ \dots \\ f_N \end{bmatrix} \quad (7)$$

In the aforementioned formula, f_2 indicates the fitness level of the second Fsh and so forth, f_1 signifies the fitness score of the first Fsh. The quantity 0.5 is used to fairly allocate the balance amongst utilization and discovery across repetitions. During the early part of the cycle (when $EFs/MaxEFs < 0.5$), group focus on universal discovery. Conversely, in the latter section (when $EFs/MaxEFs \geq 0.5$), they shift to utilization.

Discovery stage

Fishermen primarily investigate independently, utilizing group encirclement as an auxiliary method. As the study advances, the ecological benefits progressively transition from the aquatic species to the anglers. Moreover, ongoing capture will result in a decline in the number of fish and capture frequency. Fishermen will transition from autonomous discovery to predominantly depending on communal encirclement, utilizing individual resistances as support. The transformation in this mode is represented by the capture rate variable, denoted as δ.

$$\delta = \left(1 - \frac{3 \times EFs}{2 \times MaxEFs}\right)^{\frac{3 \times EFs}{2 \times MaxEFs}} \quad (8)$$

EFs & $MaxEFs$ donate the present and the greatest quantities of estimations, correspondingly.

Individual fishing (whenever $EFs/MaxEF < 0.5$)

Fishermen agitate the water to surface the fish, ascertain their location, and modify the course of their quest. The formula for the revision is as outlined below:

$$Exp = \frac{f_i - f_{pos}}{f_{max} - f_{min}} \quad (9)$$

$$R = Dis \times \sqrt{|Exp|} \times \left(1 - \frac{EFs}{MaxEFs}\right) \quad (10)$$

$$F_{i,j}^{T+1} = F_{i,j}^{T} + \left(F_{pos,j}^{T} - F_{i,j}^{T}\right) \times Exp + rand \times s \times R \quad (11)$$

In the aforementioned formula, Exp denotes the experimental examination number derived by the ith Fsh utilizing any other Fsh p (where pos = 1, 2 ... or N, p ≠ i) as the reference aim, with quantities spanning from -1 to 1. fmax and fmin denote the maximum & minimum fitness quantities, correspondingly, subsequent to the Tth full location adjustment. T represents the quantity of repetitions about the locations of fishermen. $F_{i,j}^{T}$ and $F_{i,j}^{T+1}$ represent the situation of the ith Fsh in the j-dimension following the Tth & $(T+1)th$ repetitions, accordingly. Dis represents the Euclidean distance among the i-th individual and the reference point, whereas s is a stochastic unit vector in d-dimensions.

Group fishing (when $EFs/MaxEFs \geq 0.5$)

Fishermen employ nets to augment their fishing efficacy and cooperate with one another. They form arbitrary groupings of 3 or 4 individuals to collaboratively surround prospective objectives. By utilizing their personal movement, they may investigate the region more thoroughly and precisely. The relevant formulas are described as follows:

$$Centre_c = mean(F_c^T) \quad (12)$$

$$F_{c,i,j}^{T+1} = F_{c,i,j}^{T} + r_2 \times \left(Centre_c - F_{c,i,j}^{T}\right) + \left(1 - \frac{2 \times EFs}{MaxEFs}\right)^2 \times r_3 \quad (13)$$

Here, c denotes a cluster of three to four individuals whose locations remain unchanged. $Centre_c$ signifies the focal point for the encirclement of the group C. $F_{c,i,j}^{T+1}$ and $F_{c,i,j}^{T}$ denote the location of the i-th Fsh in the group c within the j-dimension following the $(T+1)th$ and Tth updates, correspondingly. r_2 denotes the velocity at which a Fsh advances toward the center, differing among individuals and residing within the interval (0,1). r_3 denotes the move's offset, varying between $(-1, 1)$, and declines progressively as EFs escalate.

Utilization stage

All anglers employed the same method, deliberately congregating concealed fish in the same vicinity. The Fsh's places during the trapping procedure are revised as outlined below:

$$\sigma = \sqrt{\left(\frac{2\left(1 - \frac{EFs}{MaxEFs}\right)}{\left(1 - \frac{EFs}{MaxEFs}\right)^2 + 1}\right)} \quad (14)$$

$$F_i^{T+1} = Gbest + GD\left(0, \frac{r_4 \times \sigma \times |mean(F) - Gbest|}{3}\right) \quad (15)$$

Here, GD represents a Gaussian distribution function with an average μ of 0, and its total variance σ diminishes from 1 to 0 as the number of assessments rises. The location of the i-th Fsh subsequent to the $(T + 1)th$ revise. Average (F) is the matrix of average quantities for each dimension at the center of the Fsh's placements, whereas $Gbest$ denotes the universal ideal.

$$D^i = \begin{cases} 1 & if \quad i = p(j) \quad P = randperm(k), k \in [2, \lceil r_1 \ (d - 2)\rceil + 1] \\ 0 & else \end{cases} \quad (17)$$

The characteristics of omnidirectional flight include the following:

$$D^i = 1, \qquad i = 1, \ldots, d \quad (18)$$

Here, randi ([1, d]) produces an accidental integer within the range of one to d, randperm (k) provides an arbitrary permutation of values from one to k, and r_1 is an arbitrary quantity in the range $(0, 1]$. The diagonal trajectory in a d-dimensional region resides within a hyper-rectangle.

The $AHOA$ initially manufactures a collection of accidental responses and a visitation table. In each cycle, hunting is conducted through directed or local approaches 50% of the time. HBs can travel to their desired power supplies through directed hunting, which relies on nectar filling levels and a visitation schedule. Regional hunting enables HBs to efficiently traverse adjacent zones within their territory to discover potential fresh power supplies. Movement hunting occurs every two rounds. Every action and computation is conducted interactively until the stop condition is attained. The energy source that demonstrates the quickest rate of nectar Replenishment (RPL) is regarded as a generally ideal option.

1: A community of n HBs are arbitrarily assigned to n power supplies in the order of:

$$x_i = Low + r \times (Up - Low), \qquad i = 1, \ldots, n \quad (19)$$

$$x_i(t + 1) = \begin{cases} x_i(t), & f(x_i(t) \le f(v_i(t + 1))) \\ v_i(t + 1), & f(x_i(t) > f(v_i(t + 1))) \end{cases} \quad (23)$$

Here, $f(\cdot)$ denotes the fitness grade of the function. Eq. (23) indicates *that* if the nectar RPL rate of the prospective resource of nutrition surpasses that of the existing source, the HB will forsake the present power supply in favor of the possible resources, as determined by Eq. (21), for nutritional purposes.

3: Regional hunting: Upon arriving at a designated nourishment for nectar consumption, HBs may pursue additional nourishment resources. Therefore, an HB can effortlessly relocate to an adjacent area within its domain, where it may discover fresh nutrients that could serve as superior alternatives. The formula for modeling the regional exploration actions of HBs in regional hunting techniques and potential supplies of nourishment is as outlined below:

2.2.3 Artificial hummingbird optimization algorithm ($AHOA$)

$AHOA$ is a crowd-based metaheuristic technique that primarily emulates three gathering patterns of Hummingbirds (HBs): directed hunting, local hunting, and nomadic hunting. Three flight abilities are demonstrated during the hunting technique: omnidirectional, diagonal, & axial fligh. An access table emulating the remarkable memory capability of the HB is created to facilitate the HB's execution of universal improvement within the procedure. The three aerial competencies are delineated in the following manner:

The flight simulation capabilities are broadened to encompass the $d - D$ space, leading to the definition of axial flight as an outcome.

$$D^i = \begin{cases} 1 & if \quad i = rand\, i([1, d]), i = 1, \ldots, d \\ 0 & else \end{cases} \quad (16)$$

Diagonal flight is characterized by these features:

Here, r signifies a random vector in the range $[0, 1]$, Low and Up denote the minimum and maximum borders for a d-dimensional issue, accordingly, and x_i denotes the location of the i-th source of nourishment.

$$VT_{i,j} = \begin{cases} 0 & if \quad i \ne j \\ null & i = j \end{cases} \quad (20)$$

While $i = j$, $VT_{i,j} = null$ signifies that an HB is consuming nutrition at its designated resource, while $i \ne j$, $VT_{i,j} = 0$ denotes that the jth resource of nutrition has recently been accessed by the ith HB in the present repetition.

2: Directed hunting: Utilizing its flight abilities, the HB can reach its desired nourishment resource to acquire potential nourishment options. Thus, the numerical equation for simulating guided chasing action and potential food resources is:

$$v_i(t + 1) = x_{i,tar}(t) + a \times D \times \left(x_i(t) - x_{i,tar}(t)\right) \quad (21)$$

$$a \sim N(0, 1) \quad (22)$$

Here, $x_i(t)$ denotes the location of the i-th HB nourishment resource at the time t, $x_{i,tar}(t)$ represents the placement of the ith humming-bird goal power supply resource at the time t, and follows an ordinary distribution with a Standard Deviation (SD) of one and an average of zero.

$$v_i(t+1) = x_i(t) + b \times D \times x_i(t) \quad (24)$$

$$b \sim N(0,1) \quad (25)$$

Here, b follows an average distribution with a median of 0 and a SD of one.

4: If food is often limited in an area where HBs live, the bird typically goes to more distant food sources to find sustenance. A migration factor is specified in the AHA technique. HBs at the nutritional resource with the minimum filling rate will accidentally move to a different power supply across the whole quest region whenever the number of repetitions surpasses the established movement coefficient. At this juncture, the HB will forsake the initial resource and remain in a fresh situation for hunting. The action of an HB pursuing food from an area with a low nectar RPL rate to a newly available source can be described as follows:

$$x_{wor}(t+1) = Low + r \times (Up - Low) \quad (26)$$

Here, x_{wor} signifies the power supply exhibiting the minimum nectar RPL rate within the crowd.

The $CFOA$ and $AHOA$ were selected due to their complementary exploration–exploitation capabilities. Unlike conventional metaheuristics such as PSO and GA, which may suffer from premature convergence or require extensive parameter tuning, $CFOA$ enhances global search through adaptive fish-swarming behavior that maintains population diversity, while $AHOA$ employs intelligent foraging strategies that balance local exploitation and global exploration. In contrast to Bayesian optimization, which can be computationally expensive for high-dimensional or non-differentiable problems, these bio-inspired algorithms offer efficient and robust search dynamics. This theoretical reasoning justifies their selection and supports the expectation that the proposed hybrid model can achieve superior optimization performance.

The computational analyses were performed using Python 3.10 with the scikit-learn library (version 1.3). The DT model hyperparameters were optimized within the following ranges: max_depth (2–30), min_samples_split (2–20), and min_samples_leaf (1–10). The criterion 'MSE was used to evaluate split quality.

3 Statistical indices for effectiveness comparison

Several metrics were established to evaluate and compute the effectiveness of the created DT approach.

Coefficient of Determination:

$$R^2 = \left(\frac{\sum_{i=1}^{n}(X_i - \overline{X})(Y_i - \overline{Y})}{\sqrt{\left[\sum_{i=1}^{n}(X_i - X)^2\right]\left[\sum_{i=1}^{n}(Y_i - \overline{Y})^2\right]}} \right)^2 \quad (27)$$

Root Mean Square Error:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(Y_i - X_i)^2} \quad (28)$$

Normalized Root Mean Square:

$$NRMSE = \frac{RMSE}{\bar{Y}} \quad (29)$$

Relative absolute error:

$$RAE = \frac{\sum_{i=1}^{n}|X_i - Y_i|}{\sum_{i=1}^{n}|X_i - \bar{X}|} \quad (30)$$

Relative Square Error:

$$RRSE = \sqrt{\frac{\sum_{i=1}^{n}(X_i - Y_i)^2}{\sum_{i=1}^{n}(X_i - \bar{X})^2}} \quad (31)$$

variation Account Factor:

$$VAF = \left(1 - \frac{var(X_i - Y_i)}{var(X_i)}\right) * 100 \quad (32)$$

Mean Absolute Scaled Error:

$$MASE = \frac{\frac{1}{n}\sum_{i=1}^{n}|X_i - Y_i|}{\frac{1}{n-1}\sum_{i=1}^{n}|X_i - X_{i-1}|} \quad (33)$$

Mean Squared Logarithmic Error:

$$MSLE = \frac{1}{n}\sum_{i=1}^{n}\left(In(X_i + 1) - In(Y_i + 1)\right)^2 \quad (34)$$

Median Absolute Error:

$$MedAE = median(|X_i - Y_i|) \quad (35)$$

Y_i : the anticipated value of CC

$\bar{Y}$: the mean predicted value of CC

X_i : The measured value of CC

$\bar{X}$: the mean measured value of CC

n : the entire amount of data of CC

4 Findings and scrutinization

By combining the DT method with the $AHOA$ and $CFOA$ approaches, the CC was ascertained. These approaches are also known as $DT(CFOA)$ and $DT(AHOA)$. For the $DT(AHOA)$ and $DT(CFOA)$ approaches, the actual and expected CC values are shown in Figure 4. During the study's Train and Test, the values were acquired. The error % distribution between the observed and anticipated values of CC is also shown in Figure 5. Table 3 shows the results of the design assessments, which were gathered while studying and judging the product creation approach. To increase the ACC of the combined models, the current study further specified the variance percentage and alphabetical ordering for each model at each stage. ANN [37] was utilized to examine the resilience and dependability of the models. For doing so, the findings of the current investigation were contrasted with those of earlier studies.

According to the statistics, $DT(CFOA)$ and $DT(AHOA)$ are both probably able to forecast CC with ACC. Based on performance assessment measures, $DT(CFOA)$ is a distinct methodology that demonstrates greater accuracy and reliability compared to $DT(AHOA)$. In both the Train and Test, $DT(AHOA)$ receives less R^2 values (0.98123 and 0.97342), but $DT(CFOA)$ obtains greater R^2 values (0.98978 and 0.98475). The results showed a performance pattern similar to the R^2 value when the VAF were included in the

ideal model. Incorporating an extra evaluation focusing on measurement errors could increase the dependability of a procedure. Therefore, it is crucial to rigorously examine the efficacy of several metrics, namely, $RMSE$, $NRMSE$, $RRSE$, RAE, $MASE$, $MSLE$, and $MedAE$, in order to achieve this goal. Compared to an inflated number, a lower value in these measurements indicates a better degree of precise consistency. The results showed that the $DT(CFOA)$ architecture maintained outstanding dependability while achieving the lowest values. This result was reached by contrasting the $DT(CFOA)$ model with a different model. The two models' percentage variances for these parameters vary from at least 0.8% to a highest decrease of 47%, indicating the correctness and reliability of the $DT(CFOA)$. The learning and assessment phases have RAE index values of 0.0418 and 0.0491, respectively, based on $DT(CFOA)$. For the DT $(AHOA)$, comparable RAE values of 0.05786 and 0.06649 were found. After analyzing the evaluation criteria, logical inference, variance percentage, and alphabetical ranking, it has been concluded that each model is reliable and trustworthy. For its intended use, the $DT(CFOA)$ model is marginally better than the alternative model.

To assess the models' dependability, a detailed comparison with the existing literature, which includes the ANN model, is carried out [38]. This makes it easier to assess how reliable the models are. A close look at Table 3 shows that the $DT(CFOA)$ model employed in this study produced better outcomes than the one from the earlier study in this collection. Analogous measures from the training and evaluation data phases—R^2 and $RMSE$, respectively, were used to accomplish this result. ANN's findings are less reliable and robust than those generated by the better model ($DT(CFOA)$). The ACC of the $DT(CFOA)$ model is higher. R^2 decrease for ANN [38] dropped from 0.9717 to 0.98978 (-1.826% reduction) throughout learning and from 0.9702 to 0.98475 (-1.477% reduction) throughout evaluation. By analyzing the results of the instruction and assessing collections of data, a thorough comparison among $DT(CFOA)$ and ANN [38] can be carried out. Notable improvements were observed through the decline in the R^2 value.

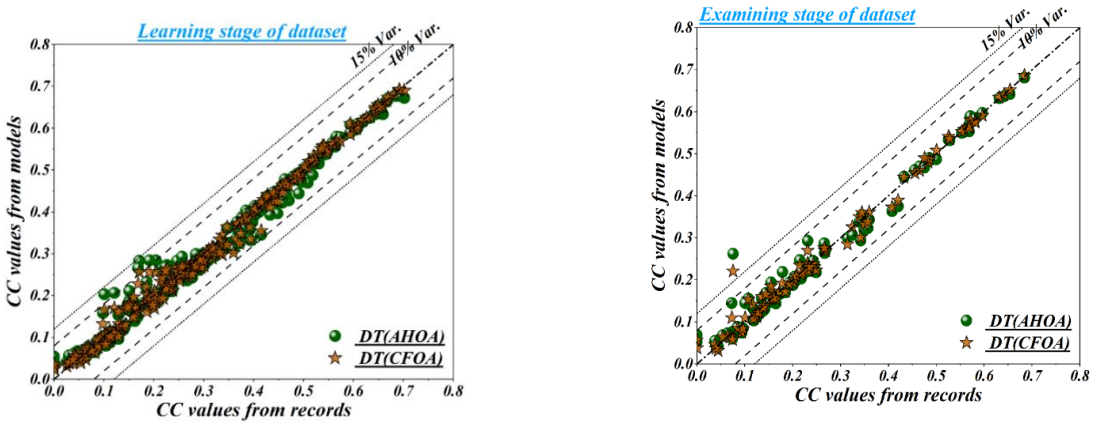
The amount of variation in Figure 5 that may be responsible for the model's erroneous predictions may be ascertained using plots of variance percentage in tree systems throughout the Train and Test. The error is computed by deducting the expected values from the actual values in order to get these charts. The variability of these errors over data points or epochs is then assessed. The variation percentage, which is often shown as a percentage, is calculated by dividing the error variance by the overall data variability. Plots tailored to the Train and Test are created to show how the variance percentage changes over time or between samples. While significant variation during training and low variance during testing indicate the presence of overfitting, low variance percentages indicate a high degree of stability and consistency in predictions. These graphs provide vital details on the stability and generalizability of the idea. The distribution information clearly indicates that the DT(CFOA) has regularly outperformed the DT(AHOA) at each point.

A comparative evaluation of the proposed models—DT-CFOA and DT-AHOA—demonstrates that both approaches outperform conventional models in terms of prediction accuracy and generalization. The superior performance can be primarily attributed to the integration of metaheuristic optimization, which effectively tunes the decision tree parameters to minimize overfitting and enhance predictive stability. Among the two, DT-AHOA shows slightly better results due to the adaptive foraging mechanism of AHOA, which provides a stronger balance between exploration and exploitation compared to CFOA's swarm-based search. Furthermore, the improved performance is not solely due to optimization but also influenced by the quality of the selected input features and the characteristics of the dataset, which enable the algorithms to capture complex nonlinear relationships more effectively. From a practical perspective, these improvements enhance the reliability of predictive modeling in structural engineering, leading to more accurate assessments, optimized design decisions, and improved safety margins in real-world applications.

Table 3: Statistical performance indicators for $DT(AHOA)$ vs. $DT(CFOA)$

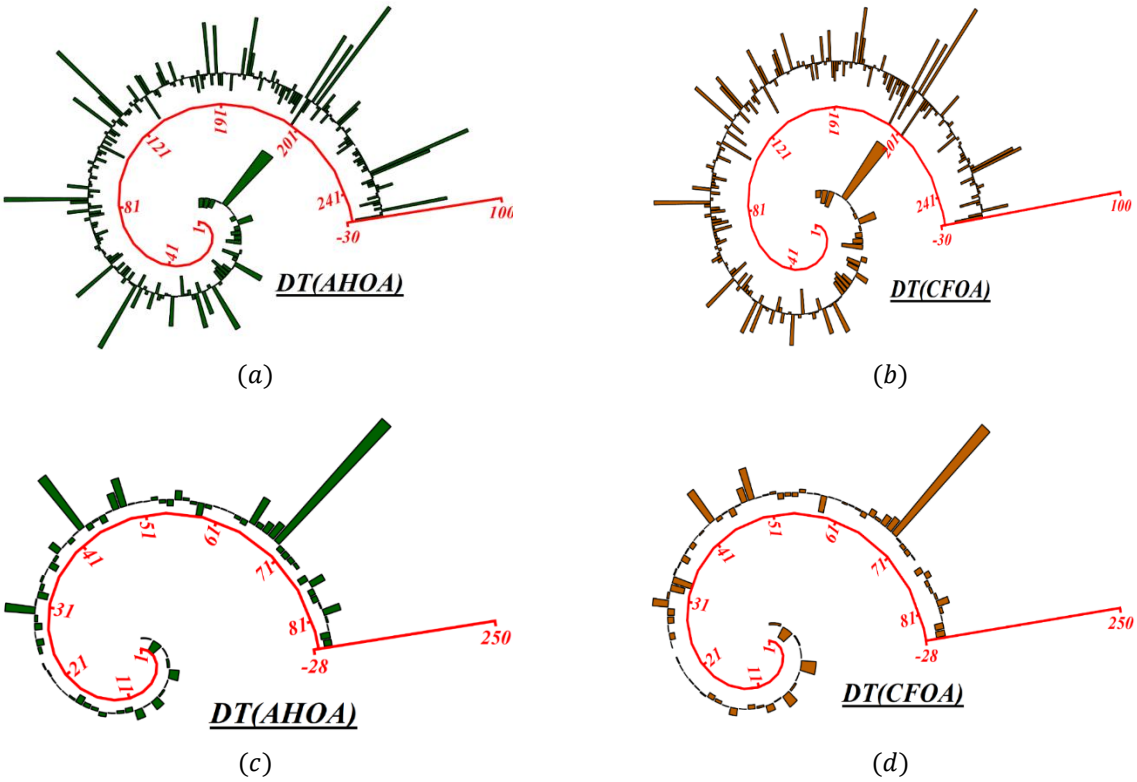
Indicators	Analysis portion	DT models					RF	Literature	
		$DT(AHOA$	Performance grade	$DT(CFOA$	Performance grade	Improvement percentage		ANN [38]	Improvement percentage
R^2	Learning	0.98123	A	0.98978	A+	0.8713	0.9138	0.9717	-1.826
	Examining	0.97342	A	0.98475	A+	1.1639	0.8636	0.9702	-1.477
$RMSE$	Learning	0.02535	A	0.01852	A+	-26.94		0.0397	114.36
	Examining	0.03046	A	0.02306	A+	-24.29		0.0391	69.557
$NRMSE$	Learning	0.03612	A	0.02639	A+	-26.9			
	Examining	0.04451	A	0.0337	A+	-24.28			
RAE	Learning	0.05786	A	0.0418	A+	-27.75			
	Examining	0.06649	A	0.0491	A+	-26.15			
$RRSE$	Learning	0.13884	A	0.10144	A+	-26.93			
	Examining	0.16468	A	0.1246	A+	-24.33			

VAF	Learning	98.0797	A	98.977	A+	0.9148			
	Examining	97.3098	A	98.4622	A+	1.1842			
$MASE$	Learning	0.08588	A	0.06204	A+	-27.75			
	Examining	0.0899	A	0.0664	A+	-26.14			
$MSLE$	Learning	0.00042	A	0.00022	A+	-47.61			
	Examining	0.00068	A	0.00039	A+	-42.64			
$MedAE$	Learning	0.01155	A	0.00801	A+	-30.64			
	Examining	0.01147	A	0.0080	A+	-30.25			
$U95\%$	Learning	0.0703	A	0.0513	A+				
	Examining	0.0845	A	0.0639	A+				



(a) (b)

Figure 4: Correlation between recorded and predicted CC values: evaluating the model's predictive ACC



(a) (b)

(c) (d)

Figure 5: Percentage error distribution between recorded and predicted CC values: analyzing model ACC and deviation trends

This research has a number of limitations despite its encouraging findings. The variety of $UHPC$ Cr behavior may not be well captured by the dataset size (338 data points), and relying only on data from the literature generates discrepancies because of varying experimental settings. It is yet unclear whether the model can be applied

to various loading and curing scenarios. Future studies should compare DT-based models with other ML techniques, investigate hybrid optimization approaches, and broaden the dataset with a variety of experimental data to enhance dependability. The credibility of the model would be further increased by including experimental validation, uncertainty quantification, and sophisticated feature interpretation (such as $SHAP$ analysis).

5 Conclusions

This research aims to create and evaluate the best ML and tree-based approaches for determining the CC of $UHPC$. To accomplish this goal, the DT was used. The $AHOA$ and the $CFOA$ were the MO techniques used in this investigation. To ascertain the ideal values for the variables employed in decision-making procedures, the DT studies were included into these approaches.

The $PAWN$ sensitivity study results indicated that the LT parameter (t) exerts the most significant influence on the result of the CC model, succeeded by the compressible strength (f_{ct}) and modulus of elasticity (E_t). Additional variables, including CT (T_c), RH, W/B, and AG/C, also exert a moderate influence.

According to the statistics, $DT(CFOA)$ and $DT(AHOA)$ were both probably able to forecast CC with ACC. Based on performance assessment measures, $DT(CFOA)$ is a distinct methodology that demonstrates greater accuracy and reliability compared to $DT(AHOA)$. In both the Train and Test, $DT(AHOA)$ received less R^2 values (0.98123 and 0.97342), but $DT(CFOA)$ obtained greater R^2 values (0.98978 and 0.98475).

The results showed that the $DT(CFOA)$ architecture maintained outstanding dependability while achieving the lowest values. This result was reached by contrasting the $DT(CFOA)$ model with a different model. The two models' percentage variances for these parameters vary from at least 0.8% to a maximum decrease of 47%, indicating the correctness and reliability of the $DT(CFOA)$.

The learning and examination stages had RAE index values of 0.0418 and 0.0491, respectively, based on $DT(CFOA)$. For the $DT\,(AHOA)$, comparable RAE values of 0.05786 and 0.06649 were found.

After analyzing the evaluation criteria, logical inference, variance percentage, and alphabetical ranking, it has been concluded that each model is reliable and trustworthy. For its intended use, the $DT(CFOA)$ model is marginally better than the alternative model.

ANN's findings from the literature were less reliable and robust than those generated by the better model ($DT(CFOA)$). The ACC of the $DT(CFOA)$ model was higher. R^2 decreased for ANN from 0.9717 to 0.98978 (-1.826% reduction) throughout learning and from 0.9702 to 0.98475 (-1.477% reduction) throughout evaluation. By scrutinizing the results of the instruction and assessing information collections, a thorough comparison among $DT(CFOA)$ and ANN can be carried out. Notable improvements were observed through the decline in the R^2 value.

The developed DT-CFOA and DT-AHOA models demonstrate strong potential for real-world integration within digital twin frameworks and predictive maintenance systems. In digital twin platforms, these models can serve as intelligent analytical components, continuously updating structural condition assessments based on real-time sensor data and providing early warnings of performance degradation. Similarly, in predictive maintenance schemes, the models can be used to forecast potential failures, optimize inspection schedules, and reduce maintenance costs while improving structural safety. The interpretable nature of the decision tree framework enables engineers to trace the influence of key features on model outputs, thus facilitating transparent decision-making. This interpretability ensures that the models are not treated as "black boxes," but as tools that enhance engineering insight, improve trust in data-driven predictions, and support informed decision-making in structural engineering practice.

Competing of interests

The authors declare no competing of interests.

Authorship contribution statement

Hao WANG: Writing-Original draft preparation, Conceptualization, Supervision, Project administration.
Chao WANG: Methodology, Software

Author statement

The manuscript has been read and approved by all the authors, the requirements for authorship, as stated earlier in this document, have been met, and each author believes that the manuscript represents honest work.

Ethical approval

All authors have been personally and actively involved in substantial work leading to the paper, and will take public responsibili

References

[1] C. Shi, D. Wang, L. Wu, and Z. Wu, "The hydration and microstructure of ultra-high-strength concrete with cement–silica fume–slag binder," *Cem Concr Compos*, vol. 61, pp. 44–52, 2015. https://doi.org/10.1016/j.cemconcomp.2015.04.013

[2] R. S. Benemaran, M. Esmaeili-Falak, and M. S. Kordlar, "Improvement of recycled aggregate concrete using glass fiber and silica fume," *Multiscale and Multidisciplinary Modeling, Experiments and Design*, vol. 7, no. 3, pp. 1895–1914, Jul. 2024, https://doi.org/10.1007/s41939-023-00313-2

[3] Z. P. Bazant and L. Panula, "Creep and shrinkage characterization for analyzing prestressed

concrete structures," *PCI journal*, vol. 25, no. 3, pp. 86–122, 1980.

[4] Y. Liu, L. Wang, Y. Wei, C. Sun, and Y. Xu, "Current research status of UHPC creep properties and the corresponding applications–A review," *Constr Build Mater*, vol. 416, p. 135120, 2024. https://doi.org/10.1016/j.conbuildmat.2024.135120

[5] B. Zhu, Y.; Huang, L.; Zhang, Z.; Bayrami, "Estimation of splitting tensile strength of modified recycled aggregate concrete using hybrid algorithms," *Steel and Composite Structures*, vol. 44, no. 3, pp. 389–406, 2022, doi: https://doi.org/10.12989/scs.2022.44.3.389.

[6] X. Zeng, S. Zhu, K. Deng, C. Zhao, and Y. Zhou, "Experimental and numerical study on cyclic behavior of a UHPC-RC composite pier," *Earthquake Engineering and Engineering Vibration*, vol. 22, no. 3, pp. 731–745, 2023. https://doi.org/10.1007/s11803-023-2185-9

[7] R. Ullah, Y. Qiang, J. Ahmad, N. I. Vatin, and M. A. El-Shorbagy, "Ultra-high-performance concrete (UHPC): A state-of-the-art review," *Materials*, vol. 15, no. 12, p. 4131, 2022. https://doi.org/10.3390/ma15124131

[8] Y. Huang, J. Wang, Q. Wei, H. Shang, and X. Liu, "Creep behaviour of ultra-high-performance concrete (UHPC): A review," *Journal of Building Engineering*, vol. 69, p. 106187, 2023. https://doi.org/10.1016/j.jobe.2023.106187

[9] M. Esmaeili-Falak and R. Sarkhani Benemaran, "Ensemble Extreme Gradient Boosting Based models to predict the Bearing Capacity of Micropile Group," *Applied Ocean Research*, 2024. https://doi.org/10.1016/j.apor.2024.104149

[10] K Zhang; Y Zhang; B Razzaghzadeh, "Application of the optimal fuzzy-based system on bearing capacity of concrete pile," *Steel and Composite Structures*, vol. 51, no. 1, pp. 25–41, 2024, doi: https://doi.org/10.12989/scs.2024.51.1.025.

[11] X. Sun, X. Dong, W. Teng, L. Wang, and E. Hassankhani, "Creation of regression analysis for estimation of carbon fiber reinforced polymer-steel bond strength," *Steel and Composite Structures*, vol. 51, no. 5, pp. 509–527, 2024. doi: 10.12989/scs.2024.51.5.509

[12] R. Sarkhani Benemaran, "Application of extreme gradient boosting method for evaluating the properties of episodic failure of borehole breakout," *Geoenergy Science and Engineering*, p. 211837, 2023, doi: 10.1016/j.geoen.2023.211837.

[13] N. K. Oghli and M. Esmaeili-Falak, "Predicting Liquefaction Triggering Potential Using Metaheuristic GMDH Approaches," *Geotechnical and Geological Engineering*, vol. 43, no. 6, p. 308, 2025, doi: 10.1007/s10706-025-03275-z.

[14] I. Nunez, A. Marani, M. Flah, and M. L. Nehdi, "Estimating compressive strength of modern concrete mixtures using computational intelligence: A systematic review," *Constr Build Mater*, vol. 310, p. 125279, 2021, doi: 10.1016/j.conbuildmat.2021.125279.

[15] L. Bal and F. Buyle-Bodin, "Artificial neural network for predicting creep of concrete," *Neural Comput Appl*, vol. 25, pp. 1359–1367, 2014. https://doi.org/10.1007/s00521-014-1623-z

[16] J. Karthikeyan, A. Upadhyay, and N. M. Bhandari, "Artificial neural network for predicting creep and shrinkage of high performance concrete," *Journal of advanced concrete technology*, vol. 6, no. 1, pp. 135–142, 2008. doi: 0.3151/jact.6.135.

[17] O. A. Hodhod, T. E. Said, and A. M. Ataya, "Prediction of creep in concrete using genetic programming hybridized with ANN," *Computers and Concrete, An International Journal*, vol. 21, no. 5, pp. 513–523, 2018.

[18] A. H. Gandomi, S. Sajedi, B. Kiani, and Q. Huang, "Genetic programming for experimental big data mining: A case study on concrete creep formulation," *Autom Constr*, vol. 70, pp. 89–97, 2016. https://doi.org/10.1016/j.autcon.2016.06.010

[19] J. Feng, H. Zhang, K. Gao, Y. Liao, W. Gao, and G. Wu, "Efficient creep prediction of recycled aggregate concrete via machine learning algorithms," *Constr Build Mater*, vol. 360, p. 129497, 2022. https://doi.org/10.1016/j.conbuildmat.2022.129497

[20] J. Z. Xiao, X. D. Xu, and Y. H. Fan, "Shrinkage and creep of recycled aggregate concrete and their prediction by ANN method," *Journal of Building Materials*, vol. 16, no. 5, pp. 752–757, 2013.

[21] K. Li, Y. Long, H. Wang, and Y.-F. Wang, "Modeling and sensitivity analysis of concrete creep with machine learning methods," *Journal of Materials in Civil Engineering*, vol. 33, no. 8, p. 4021206, 2021. https://doi.org/10.1061/(ASCE)MT.1943-5533.0003843

[22] M. Esmaeili-Falak and A. Letafat, "Prediction of Unconfined Compressive Strength in Pozzolanic Geopolymer-Stabilized Granular Materials Using Tree-Based Models," *Transportation Infrastructure Geotechnology*, vol. 12, no. 7, p. 247, 2025, doi: 10.1007/s40515-025-00701-w.

[23] T. Zhou, "Developing a Machine Learning-Driven Model that Leverages Meta-Heuristic Algorithms to Forecast the Load-Bearing Capacity of Piles," *Journal of Artificial Intelligence and System Modelling*, vol. 1, no. 01, pp. 1–14, 2023. https://doi.org/10.22034/jaism.2023.423296.1006

[24] M. Esmaeili-Falak, H. Katebi, M. Vadiati, and J. Adamowski, "Predicting triaxial compressive strength and Young's modulus of frozen sand using artificial intelligence methods," *Journal of Cold Regions Engineering*, vol. 33, no. 3, p. 4019007, 2019, doi: 10.1061/(ASCE)CR.1943-5495.0000188.

[25] F. Blanco and Y. J. Woo, "Modeling Compressive Strength of Self-Compacting Concrete (SCC) Using Novel Optimization Algorithm of AOA," *Advances in Engineering and Intelligence Systems*, vol. 3, no. 03, pp. 1–14, 2024. https://doi.org/10.22034/aeis.2024.470005.1206

[26] A. Boulkroune, F. Zouari, and A. Boubellouta, "Adaptive fuzzy control for practical fixed-time synchronization of fractional-order chaotic systems," *Journal of Vibration and Control*, p. 10775463251320258, 2025.

[27] A. Boulkroune, S. Hamel, F. Zouari, A. Boukabou, and A. Ibeas, "Output-Feedback Controller Based Projective Lag-Synchronization of Uncertain Chaotic Systems in the Presence of Input Nonlinearities," *Math Probl Eng*, vol. 2017, no. 1, p. 8045803, 2017. https://doi.org/10.1155/2017/8045803

[28] F. Zouari, K. Ben Saad, and M. Benrejeb, "Robust neural adaptive control for a class of uncertain nonlinear complex dynamical multivariable systems," *International Review on Modelling and Simulations*, vol. 5, no. 5, pp. 2075–2103, 2012.

[29] F. Zouari, K. Ben Saad, and M. Benrejeb, "Adaptive backstepping control for a class of uncertain single input single output nonlinear systems," in *10th International Multi-Conferences on Systems, Signals & Devices 2013 (SSD13)*, IEEE, 2013, pp. 1–6. https://doi.org/10.1109/SSD.2013.6564134

[30] F. Zouari, K. Ben Saad, and M. Benrejeb, "Adaptive backstepping control for a single-link flexible robot manipulator driven DC motor," in *2013 International Conference on Control, Decision and Information Technologies (CoDIT)*, IEEE, 2013, pp. 864–871. https://doi.org/10.1109/CoDIT.2013.6689656

[31] G. Rigatos, M. Abbaszadeh, B. Sari, P. Siano, G. Cuccurullo, and F. Zouari, "Nonlinear optimal control for a gas compressor driven by an induction motor," *Results in Control and Optimization*, vol. 11, p. 100226, 2023. https://doi.org/10.1016/j.rico.2023.100226

[32] O. A. Hodhod, T. E. Said, and A. M. Ataya, "Prediction of creep in concrete using genetic programming hybridized with ANN," *Computers and Concrete, An International Journal*, vol. 21, no. 5, pp. 513–523, 2018.

[33] N. K. Katariya, B. S. Choudhary, M. Esmaeili-Falak, and A. K. Raina, "Optimizing Open-Pit Iron ore Mine Waste Dump Stability with an increased height: A Geotechnical Perspective," *Geomechanics and Engineerin*, vol. 40, no. 1, pp. 69–78, 2025, doi: 10.12989/gae.2025.40.1.069.

[34] L. Zhu, J.-J. Wang, X. Li, G.-Y. Zhao, and X.-J. Huo, "Experimental and numerical study on creep and shrinkage effects of ultra high-performance concrete beam," *Compos B Eng*, vol. 184, p. 107713, 2020. https://doi.org/10.1016/j.compositesb.2019.107713

[35] Y. Xu, J. Liu, J. Liu, P. Zhang, Q. Zhang, and L. Jiang, "Experimental studies and modeling of creep of UHPC," *Constr Build Mater*, vol. 175, pp. 643–652, 2018. https://doi.org/10.1016/j.conbuildmat.2018.04.157

[36] T. Hastie, R. Tibshirani, J. H. Friedman, and J. H. Friedman, *The elements of statistical learning: data mining, inference, and prediction*, vol. 2. Springer, 2009.

[37] Y. Aydın, C. Cakiroglu, G. Bekdaş, and Z. W. Geem, "Explainable ensemble learning and multilayer perceptron modeling for compressive strength prediction of ultra-high-performance concrete," *Biomimetics*, vol. 9, no. 9, p. 544, 2024. https://doi.org/10.3390/biomimetics9090544

[38] P. Zhu, W. Cao, L. Zhang, Y. Zhou, Y. Wu, and Z. J. Ma, "Interpretable Machine Learning Models for Prediction of UHPC Creep Behavior," *Buildings*, vol. 14, no. 7, p. 2080, 2024. https://doi.org/10.3390/buildings14072080