Scilight

Bulletin of Computational Intelligence
https://www.sciltp.com/journals/bci

Article

Prediction of Chloride Resistance Level in Concrete Using Optimized Tree-Based Machine Learning Models

Ali Benzaamia [1], Mohamed Ghrici [1,*], Redouane Rebouh [1], Ahmed Abdelghafour Ghrici [1] and Panagiotis G. Asteris [2]

[1] Geomaterials Laboratory, Hassiba Benbouali University of Chlef, Chlef 02000, Algeria
[2] Computational Mechanics Laboratory, School of Pedagogical and Technological Education, 15122 Athens, Greece
* Correspondence: m.ghrici@univ-chlef.dz

How To Cite: Benzaamia, A.; Ghrici, M.; Rebouh, R., et al. Prediction of Chloride Resistance Level in Concrete Using Optimized Tree-Based Machine Learning Models. *Bulletin of Computational Intelligence* **2025**, *1*(1), 104–117. https://doi.org/10.53941/bci.2025.100007

Received: 15 July 2025
Revised: 8 September 2025
Accepted: 17 September 2025
Published: 28 September 2025

Abstract: The durability of reinforced concrete structures in chloride-rich environments remains a major concern in infrastructure design, particularly in coastal regions. While standardized laboratory procedures provide reliable quantification of chloride ingress resistance, they are often time-consuming, costly, and unsuitable for early-stage mix design. This study proposes a data-driven framework for predicting the chloride resistance level of concrete using tree-based machine learning (ML) classifiers. A comprehensive experimental dataset was utilized to train and validate three ML models: Decision Tree Classifier (DTC), Random Forest Classifier (RFC), and CatBoost Classifier (CatBC). Extensive hyperparameter tuning was performed using the Optuna framework with 2000 trials per model to enhance predictive performance. Among the tested models, CatBC outperformed its counterparts with a test accuracy of 0.95 and weighted F1-score of 0.85. Feature importance analyses using SHAP values, Prediction Values Change, and other CatBoost interpretability tools consistently identified the water-to-binder ratio, superplasticizer content, test age, and aggregate proportions as key predictors of chloride resistance. The findings demonstrate that machine learning offers a fast, cost-effective, and accurate alternative for classifying concrete's chloride resistance, supporting informed decision-making in mix design and service-life assessment.

Keywords: concrete durability; chloride resistance; machine learning; CatBoost; feature importance

1. Introduction

Reinforced concrete (RC) remains the backbone of modern infrastructure, particularly in marine and coastal regions where its resilience, versatility, and cost-effectiveness make it the material of choice. With growing urbanization and economic activities concentrated near shorelines, the exposure of RC structures to aggressive chloride-laden environments has become increasingly frequent. Chloride ions originating from marine aerosols or deicing salts penetrate concrete through its pore network, ultimately leading to the corrosion of embedded steel reinforcement. This corrosion compromises both the safety and service life of the structures, posing significant maintenance and economic challenges [1,2].

The ingress of chlorides into concrete is governed by complex transport mechanisms–including diffusion, permeation, and capillary suction–with diffusion generally considered the predominant process in saturated conditions. Conventionally, this process is mathematically described by Fick's second law, using a diffusion coefficient that serves as a fundamental durability indicator [3,4]. Several service-life prediction models, such as Life-365 [5] and DuraCrete [6], rely on the accurate determination of the chloride diffusion coefficient to assess structural longevity.

Among the available experimental techniques, the NT Build 492 [7] standard provides a practical approach for determining the non-steady-state migration coefficient (D_{nssm}), offering relatively rapid test results by applying an external electrical field to accelerate chloride ion migration. However, despite its efficiency compared to other test methods, NT Build 492 remains labor-intensive, time-consuming, and reliant on physical testing, often performed after 28 days of curing. These limitations hinder its integration into the early stages of mix design or durability assessments, particularly when multiple trials are required to achieve target performance levels [8].

To promote more efficient durability design, it is critical to develop predictive approaches that can estimate resistance to chloride penetration from compositional and early-age information. However, modeling chloride resistance is inherently challenging due to the multitude of interrelated factors influencing the transport behavior–ranging from binder type and dosage to aggregate content, admixtures, and curing age. The growing use of diverse supplementary cementitious materials (SCMs), such as fly ash (FA), slag (GGBFS), and silica fume (SF), further complicates the prediction process due to their nonlinear and interactive effects on concrete microstructure. Consequently, determining the optimal combination of constituents to meet specific durability requirements remains a significant challenge [9].

In this context, the need for fast, cost-effective, and accurate predictive tools to assess the chloride resistance of concrete is increasingly urgent. Traditional diffusion-based models rely on simplifying assumptions and fail to account for the nonlinear interactions among compositional variables. Moreover, many existing empirical models are limited to specific concrete types and exposure conditions, and do not generalize well across diverse datasets or mixtures [9].

To overcome these limitations, artificial intelligence (AI), particularly machine learning (ML), offers a promising alternative by leveraging data-driven algorithms capable of capturing complex, nonlinear relationships without presupposing functional forms [10–13]. Recent studies have applied such techniques across diverse civil engineering domains, including chloride erosion modeling [14], axial compression capacity prediction of GFRP-reinforced concrete columns [15], flexural capacity estimation of strengthened RC beams [16], chloride resistance modeling of recycled aggregate concrete [17], and bond strength prediction of externally bonded reinforcement on groove systems [18]. These applications highlight the versatility and effectiveness of AI-driven methods in capturing the behavior of complex concrete and structural systems.

Several studies have already demonstrated the efficacy of ML in predicting chloride transport properties [19–23]. For instance, artificial neural networks (ANN) and hybrid models have been applied to estimate the diffusion coefficient of high-performance and self-compacting concretes, often achieving high predictive accuracy. However, prior work is frequently constrained by small, homogeneous datasets, narrow feature selection, or limited algorithmic diversity. Most models are trained on specific concrete formulations, limiting their transferability to broader contexts. Addressing these gaps, Taffese and Espinosa-Leal [9] compiled a comprehensive multi-source dataset and proposed a machine learning approach to classify concrete based on its resistance to chloride ingress. Their work demonstrated the feasibility of ML-based classification using a diverse range of concrete types and mix designs. Nevertheless, their analysis was limited to a single algorithm (XGBoost), and further exploration of alternative ML models and optimization strategies was not conducted.

Building upon this foundation, the present study investigates the predictive potential of three tree-based classifiers—Decision Tree, Random Forest, and CatBoost—for estimating the chloride resistance level of concrete based on compositional and testing parameters. The objectives of this research are twofold:

(i) to develop optimized classification models using advanced hyperparameter tuning strategies to predict the chloride resistance level of concrete;
(ii) to assess the relative importance of different mix constituents in determining chloride transport resistance, thereby providing insights into mix design for enhanced durability.

This research aims to contribute to the development of a fast, interpretable, and cost-effective ML framework that supports the early-stage design and durability assessment of RC structures in chloride-exposed environments.

2. Research Significance

This study advances the application of machine learning in durability design by focusing on the prediction of chloride penetration resistance levels in concrete using optimized tree-based models. While previous studies have primarily modeled chloride transport through regression of diffusion coefficients, this work demonstrates the effectiveness of multiclass classification for resistance levels, thereby aligning predictive outputs directly with engineering decision-making needs. The incorporation of a systematically tuned CatBoost classifier—validated against Random Forest and Decision Tree baselines—highlights the value of advanced ensemble methods in handling heterogeneous, multi-source concrete datasets.

The findings reveal that critical mix design parameters such as water-to-binder ratio, superplasticizer dosage, and silica fume content, along with test age, strongly govern chloride ingress resistance. By integrating hyperparameter optimization and multiple interpretability analyses (Prediction Value Change, SHAP, and interaction importance), the study not only achieves high predictive accuracy but also provides transparent insights into the role of individual constituents. These contributions support the development of data-driven mix design strategies, reduce reliance on extensive laboratory testing, and enhance the efficiency of durability-oriented concrete design practices.

3. Database Description

The prediction of chloride penetration resistance in concrete necessitates the use of a comprehensive and high-quality dataset that encapsulates the complex interplay of constituent materials and testing parameters. In this study, the experimental database compiled by Taffese and Espinosa-Leal [9] was adopted as the foundation for model development. This dataset is notable for its breadth and heterogeneity, incorporating 843 concrete mix compositions collected from multiple international sources, including the Finnish DuraInt and LIFECON research projects, as well as various peer-reviewed publications. The database encompasses a wide variety of concrete types–normal-weight, lightweight, high-performance, high-strength, and self-compacting concretes–thereby ensuring a diverse representation of mix designs and performance characteristics.

Each entry in the dataset includes detailed information about the mix composition and testing conditions, alongside the measured non-steady-state migration coefficient (D_{nssm}), which serves as the basis for categorizing chloride resistance levels. Initially, the dataset contained 24 columns representing both input features and auxiliary properties. However, to align the model development with the objective of classifying chloride resistance, a subset of 14 features was selected based on their relevance and completeness. These included parameters related to the binder system (cement type, cement, slag, fly ash, silica fume, lime filler), aggregates (fine and coarse aggregate content), admixtures (plasticizer, superplasticizer, and air-entraining agents), the water-to-binder ratio (w/b), and the age of the specimen at the time of the chloride migration test. The categorical output feature, "Migration resistance", reflects the level of resistance to chloride ingress and was derived from the D_{nssm} values using the classification criteria reported in [9].

A rigorous data preprocessing protocol was implemented prior to model training. Missing data was identified as a significant issue, with certain features–such as fine and coarse aggregates, superplasticizer dosage, and admixture contents–exhibiting a considerable proportion of null entries. In order to preserve the integrity of the modeling process and avoid introducing bias through imputation, a complete case analysis approach was adopted. All entries containing at least one missing value across the selected features were excluded. This decision, while reducing the dataset size from 843 to 204 observations, ensured that the retained samples were both complete and consistent, allowing for robust training and evaluation of the machine learning models.

Table 1 summarizes the descriptive statistics of the retained dataset. The water-to-binder ratio ranged from 0.30 to 0.65, with a mean of approximately 0.42, reflecting a wide spectrum of mix workability and strength characteristics. Cement content varied significantly, from 52 to 525 kg/m^3, while the presence of supplementary cementitious materials also exhibited high variability. For instance, fly ash content reached up to 735 kg/m^3 and silica fume up to 468.5 kg/m^3, though many mixes included none of these materials, as indicated by the median values. Notably, certain features such as lime filler, plasticizer, and air-entraining agent remained constant across the retained subset and were recorded as zero in all cases, implying either their exclusion from the original mixes or the unavailability of reliable records. As these variables contained no variability, they were effectively excluded from contributing to model training. Aggregate contents also showed a broad range, with coarse aggregate varying from 0 to 1240 kg/m^3 and fine aggregate from 235 to 1574 kg/m^3. The age at migration testing ranged from as early as 3 days to 365 days, suggesting a diverse set of curing regimes and performance evaluation timelines.

The cleaned and structured dataset thus provides a suitable foundation for training predictive models capable of capturing the influence of compositional and temporal parameters on chloride resistance in concrete. Further preprocessing steps, including categorical encoding and data partitioning for model validation, are discussed in the subsequent section.

To provide a concise overview of the research workflow, a flowchart illustrating the overall methodology—from data collection to model evaluation—is presented in Figure 1. This diagram summarizes the key stages of the study, including dataset preparation, model development, hyperparameter tuning, and evaluation, thereby offering a visual guide to the subsequent detailed description.

Table 1. Descriptive statistics of the numeric input parameters.

Variable	Units	Statistics						
		Mean	STDEV	Min	25%	50%	75%	Max
Water-to-binder ratio	–	0.42	0.08	0.3	0.36	0.40	0.45	0.65
Cement	kg/m^3	361.85	96.90	52.0	297.88	350.00	444.38	525.00
Slag	kg/m^3	18.44	57.35	0.0	0.00	0.00	0.00	312.30
Fly ash	kg/m^3	44.12	133.70	0.0	0.00	0.00	0.00	735.00
Silica fume	kg/m^3	8.36	35.46	0.0	0.00	0.00	0.00	468.50
Lime filler	kg/m^3	0.00	0.00	0.0	0.00	0.00	0.00	0.00
Fine aggregate	kg/m^3	798.83	223.80	235.0	685.26	765.00	956.75	1574.10
Coarse aggregate	kg/m^3	797.21	306.18	0.0	451.50	915.53	1059.95	1240.00
Plasticizer	% by binder wt.	0.00	0.00	0.0	0.00	0.00	0.00	0.00
Superplasticizer	% by binder wt.	0.41	0.54	0.0	0.00	0.20	0.70	4.17
Air entraining	% by binder wt.	0.00	0.00	0.0	0.00	0.00	0.00	0.00
Migration test age	days	65.57	83.70	3.0	28.00	28.00	90.00	365.00

Figure 1. Flowchart summarizing the overall methodology from data collection to model evaluation.

4. Data Exploration and Preprocessing

Exploratory data analysis was carried out to examine the variability and distributional characteristics of the selected input features and the target output. Figure 2 illustrates the distribution of the numerical input features across the five chloride resistance levels. A clear downward trend is observed in the water-to-binder ratio (*w*/*b*), with lower values associated with higher resistance classes–confirming the known role of reduced *w*/*b* in enhancing durability. Similarly, the contents of slag, fly ash, and silica fume increase markedly in “Very high” and “Extremely high” resistance concretes, reflecting the effectiveness of supplementary cementitious materials (SCMs) in mitigating chloride penetration through pore refinement and improved microstructure. While cement

content generally decreases in the highest resistance class–likely due to increased SCM substitution–it remains relatively high for "High" and "Very high" classes.

The distributions of fine and coarse aggregates are highly variable, though a modest increase in fine aggregate content is noted in higher resistance classes. Superplasticizer dosage tends to be higher in durable mixes, possibly facilitating lower *w*/*b* ratios without compromising workability. Migration test age shows minor variation, with slightly older ages associated with improved resistance, potentially indicating extended curing or delayed testing. Overall, the figure underscores the combined influence of low *w*/*b* ratio and optimized binder composition on chloride resistance.

The distribution of the categorical target variable, chloride resistance level, is depicted in Figure 3. The target was constructed based on the non-steady-state migration coefficient (D_{nssm}), an established indicator of concrete's resistance to chloride ion ingress. Following the classification criteria proposed by Taffese and Espinosa-Leal [9], the resistance levels were discretized into five ordinal categories as follows:

- Low resistance: $D_{nssm} > 15.0 \times 10^{-12}$ m^2/s
- Moderate resistance: $10.0 < D_{nssm} \leq 15.0 \times 10^{-12}$ m^2/s
- High resistance: $5.0 < D_{nssm} \leq 10.0 \times 10^{-12}$ m^2/s
- Very high resistance: $2.5 < D_{nssm} \leq 5.0 \times 10^{-12}$ m^2/s
- Extremely high resistance: $D_{nssm} \leq 2.5 \times 10^{-12}$ m^2/s.

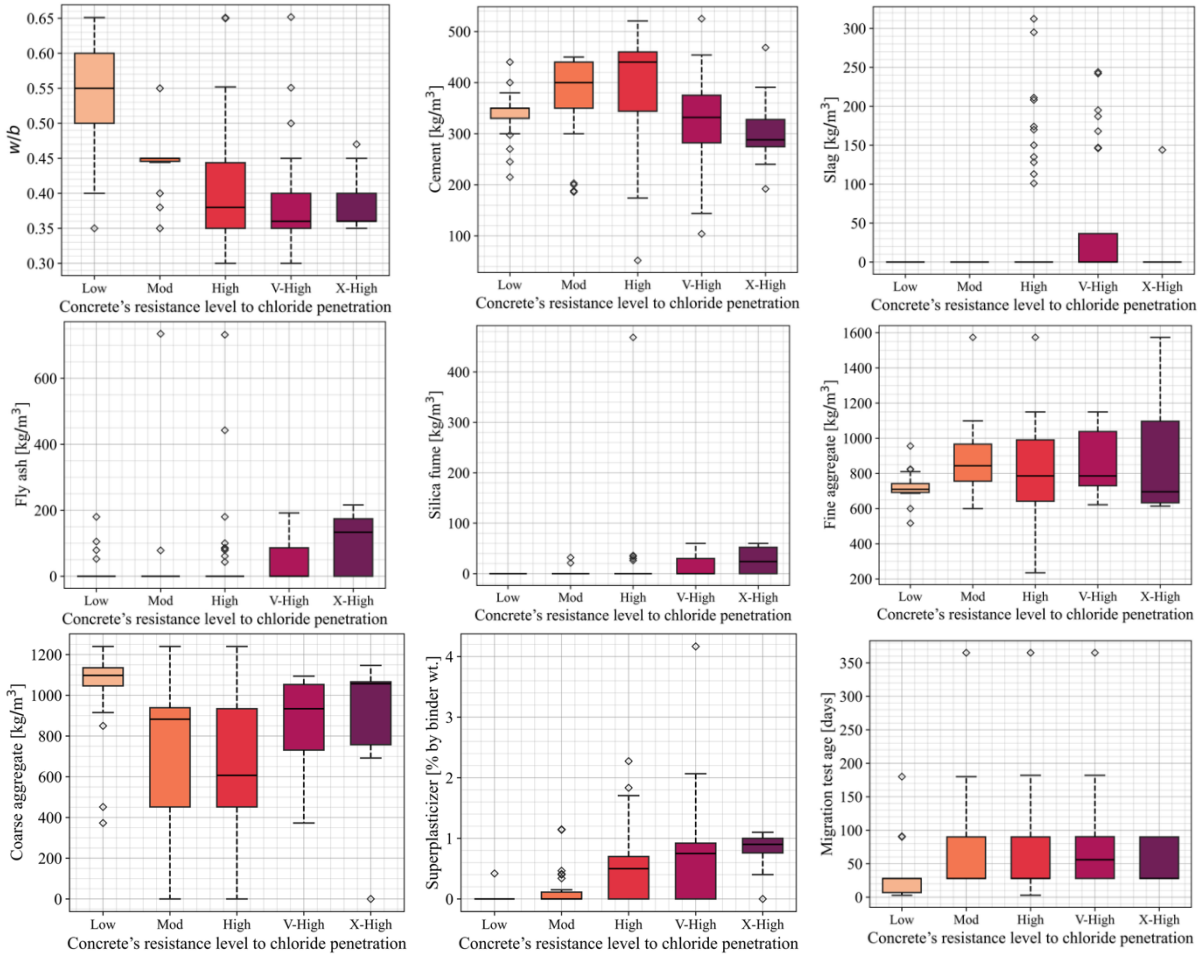


Figure 2. Variations in numerical input features across chloride resistance levels.

As observed in the figure, the majority of samples fall within the "High" resistance class, accounting for nearly half of the dataset, followed by the "Low" and "Very high" categories. The "Extremely high" and "Moderate" classes are underrepresented, introducing a class imbalance that must be carefully addressed to avoid model bias toward the dominant classes. To mitigate this, a stratified data splitting strategy was employed.

Prior to modeling, appropriate preprocessing steps were applied. The categorical variable "Cement type", which includes seven unique cement classifications (CEM I, CEM II/A-D, CEM II/A-V, CEM II/B-V, CEM II/B-S, CEM III/A, and CEM IV/A), was encoded using a one-hot encoding scheme for tree-based models such as decision trees and random forests. To avoid collinearity among dummy variables and ensure numerical stability,

the first category was dropped. For CatBoost, which supports native categorical feature handling, the raw categorical labels were retained and passed directly to the model.

Given the ordinal nature of the target variable, a label encoding approach was used to preserve the rank order of the chloride resistance levels. Each class was assigned a corresponding integer label from 0 (Low) to 4 (Extremely high), enabling the model to recognize the implicit progression in durability.

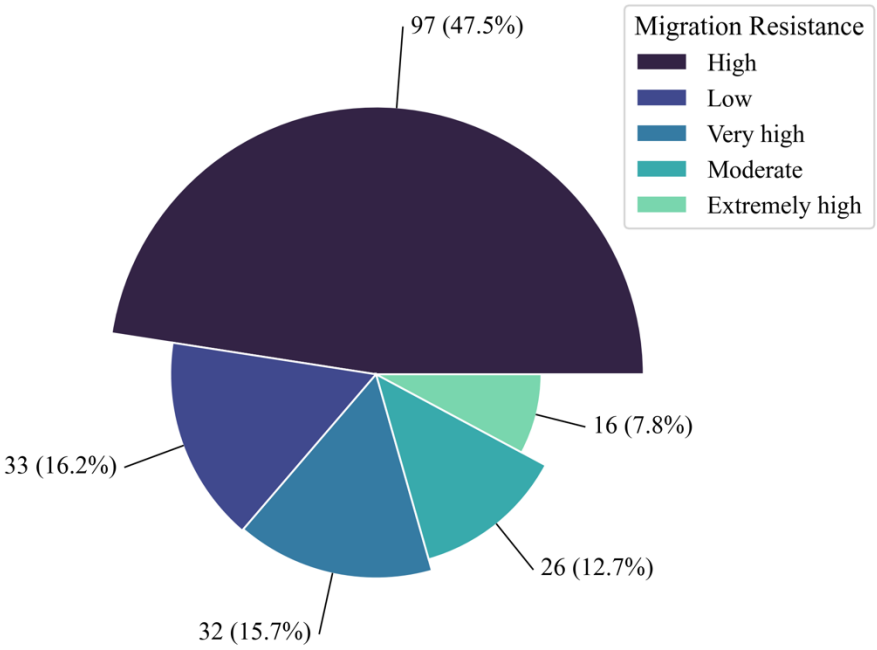


Figure 3. Class distribution of chloride resistance levels.

The final dataset was split into training and testing sets using an 80/20 partition, with stratification applied based on the target labels to ensure consistent class representation in both subsets. Categorical encoding was fitted exclusively on the training set and then applied to the test set to prevent information leakage and ensure proper generalization during evaluation.

5. Hyperparameter Tuning of Tree-Based Classifiers

In this study, three tree-based machine learning classifiers were developed to predict the chloride resistance level of concrete: the Decision Tree Classifier (DTC), the Random Forest Classifier (RFC), and the CatBoost Classifier (CatBC). These models were selected due to their proven ability to handle non-linear relationships and mixed data types [24,25], as well as their interpretability and robustness in classification tasks. The DTC algorithm is a simple yet effective model that splits data into subsets based on feature thresholds, allowing for transparent rule-based decision-making [26]. The RFC is an ensemble method that aggregates multiple decision trees trained on bootstrapped samples, thereby improving generalization and reducing overfitting [27]. CatBoost, developed by Yandex, is a gradient boosting framework optimized for categorical data and imbalanced classes. It incorporates sophisticated regularization techniques and efficient handling of categorical features, which enhances its predictive performance [28].

A major concern in the dataset was the class imbalance in the target variable Migration resistance (see Figure 3), which had the following distribution: High (97), Low (33), Very High (32), Moderate (26), and Extremely High (16). To mitigate potential bias towards majority classes, class weighting strategies were applied. For DTC and RFC, sample weights were computed using the "compute_sample_weight" function from Scikit-learn with the 'balanced' mode, and passed during model training. In contrast, CatBC inherently supports imbalanced classification via the auto_class_weights = 'Balanced' parameter, which adjusts the weight of each class inversely proportional to its frequency.

To ensure optimal performance, hyperparameter tuning was conducted using the Optuna optimization framework [29]. Hyperparameters play a crucial role in controlling model complexity and generalization, and their appropriate selection can significantly improve performance. Optuna is an efficient and flexible hyperparameter optimization tool that uses sequential model-based optimization (SMBO) with a Tree-structured Parzen Estimator

(TPE) as its default sampler. It automatically constructs and refines the search space based on prior evaluations, thereby accelerating convergence towards optimal configurations [30–33]. In this study, each model was tuned through 2,000 optimization trials, where the weighted F1-score obtained via stratified 4-fold cross-validation on the training data served as the objective metric. The search space included a combination of categorical, integer, and float parameters relevant to each model.

The best hyperparameter configurations for each model are summarized in Table 2. For the DTC, the optimal configuration used the Gini impurity criterion, a random splitter, and a tree depth of 11. The RFC achieved its best performance using the entropy criterion, a maximum depth of 20, and 39 estimators without bootstrapping. The CatBC benefited from a deep configuration (depth of 14) with 1140 boosting iterations, a learning rate of approximately 0.064, and a minimum of 12 data points per leaf. These results demonstrate the effectiveness of Optuna in navigating complex hyperparameter spaces to enhance model performance.

Table 2. Hyperparameter search spaces and optimal values obtained using Optuna.

Model	Hyperparameter	Search Space	Optimal Value
DTC	criterion	{gini, entropy, log_loss}	gini
	splitter	{best, random}	random
	max_depth	[2, 32]	11
	min_samples_splits	[2, 5]	5
	min_samples_leaf	[1, 5]	1
	max_features	{sqrt, log2, 0.6, 0.7, 0.8, 0.9, None}	None
RFC	criterion	{gini, entropy, log_loss}	entropy
	max_depth	[2, 32]	20
	min_samples_split	[2, 5]	5
	min_samples_leaf	[1, 5]	3
	max_features	{sqrt, log2, 0.6, 0.7, 0.8, 0.9, None}	0.9
	n_estimators	[20, 1000]	39
	bootstrap	{True, False}	False
CatBC	iterations	[50, 2000] (step=10)	1140
	learning_rate	[0.001, 0.1] (log-uniform)	0.0641
	depth	[4, 15]	14
	l2_leaf_reg	[1e-3, 10] (log-uniform)	0.0228
	bootstrap_type	{Bayesian, Bernoulli, MVS}	MVS
	random_strength	[0, 10]	0.674
	max_bin	[1, 255]	155
	od_type	{IncToDec, Iter}	IncToDec
	od_wait	[10, 200]	59
	leaf_estimation_method	{Newton, Gradient}	Newton
	grow_policy	{SymmetricTree, Depthwise, Lossguide}	SymmetricTree
	min_data_in_leaf	[1, 100]	12
	colsample_bylevel	[0.05, 1]	0.123
	bagging_temperature (if Bayesian)	[0, 10]	–
	subsample (if Bernoulli)	[0.6, 1.0]	–

6. Performance Assessment of Tuned Models

Following the hyperparameter optimization phase, each of the three classifiers–Decision Tree Classifier (DTC), Random Forest Classifier (RFC), and CatBoost Classifier (CatBC)–was retrained using the respective optimal configurations obtained through Optuna. The final evaluation aimed to assess the predictive capability and generalization of the tuned models on both the training and test sets.

A comprehensive suite of classification performance metrics was employed to evaluate the models. These include Accuracy Score (AS), Precision Score (PS), Recall Score (RS), Specificity Score (SS), and F1-score (F1S), in addition to the confusion matrix, which provides detailed insights into misclassification patterns across all resistance classes. All metrics were computed using the "Permetrics" Python library, which supports multi-class evaluation through customizable averaging schemes.

Accuracy Score (AS) represents the proportion of correctly classified instances, with three common averaging strategies used in multi-class problems: macro, micro, and weighted. Macro averaging treats all classes equally by averaging the accuracy of each class, while micro averaging aggregates the total number of correct predictions globally. Weighted averaging adjusts for class imbalance by assigning weights proportional to the number of instances in each class.

Precision (PS) quantifies the correctness of positive predictions, while Recall (RS) measures the model's ability to identify all relevant instances. Both are essential for understanding classification trade-offs, particularly

in imbalanced datasets. The F1-score (F1S), defined as the harmonic mean of precision and recall, offers a balanced assessment of classification effectiveness.

The Specificity Score (SS) is also included to measure how well each model identifies negative instances, which is especially important in multi-class settings where misclassification of minority classes may lead to inflated accuracy without capturing true model behavior. Like the other metrics, specificity was computed using weighted averaging to reflect the class distribution.

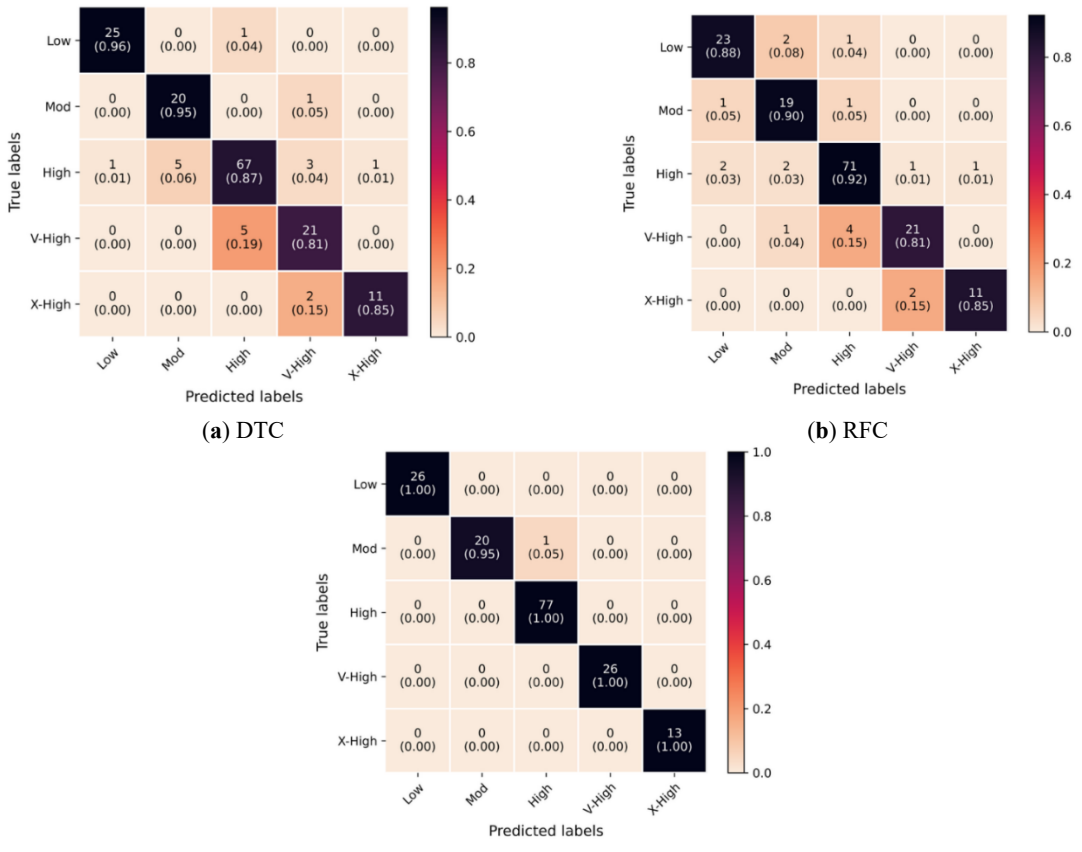
By evaluating these metrics on both training and test data, the robustness and generalization performance of each classifier can be thoroughly examined. The confusion matrices complement the numerical scores by highlighting the specific resistance levels that were most prone to misclassification, thereby supporting a deeper interpretation of model behavior.

The performance metrics obtained for the three tuned classifiers—DTC, RFC, and CatBC—on both the training and test datasets are summarized in Table 3, while the corresponding confusion matrices are presented in Figures 4 and 5. These results offer a comprehensive view of model behavior in terms of classification accuracy, error patterns, and generalization capability.

Table 3. Performance metrics of the Tree-Based models across training and test phases.

Model	Phase	AS	PS	RS	SS	F1S
DTC	Training	0.95	0.87	0.89	0.97	0.88
	Test	0.91	0.79	0.80	0.95	0.77
RFC	Training	0.96	0.88	0.87	0.97	0.87
	Test	0.90	0.74	0.73	0.94	0.72
CatBC	Training	1.00	1.00	0.99	1.00	0.99
	Test	0.95	0.86	0.86	0.96	0.85

During training, the CatBoost Classifier (CatBC) demonstrated superior learning performance, achieving perfect scores in accuracy (1.00), precision (1.00), specificity (1.00), and nearly perfect recall (0.99) and F1-score (0.99). Its confusion matrix confirms that CatBC correctly classified all training samples across the five resistance levels, with only a single misclassification in the "Very high" class. This result indicates a strong fit to the training data, though it also raises the question of potential overfitting.



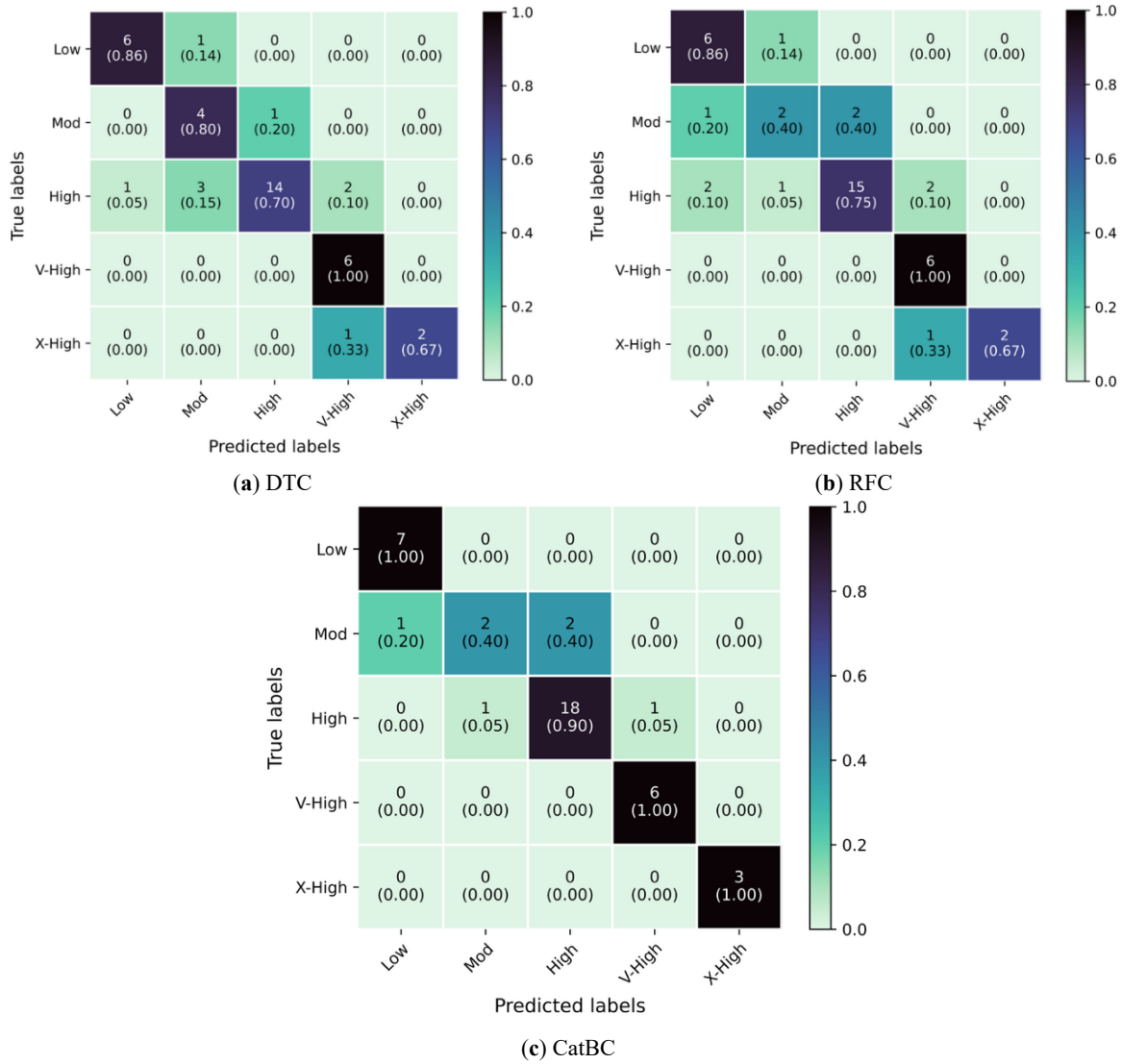
(**a**) DTC (**b**) RFC

(**c**) CatBC

Figure 4. Confusion matrix of the Tree-Based models for the training set.

The Decision Tree Classifier (DTC) and Random Forest Classifier (RFC) also performed well on the training data, with accuracy scores of 0.95 and 0.96, respectively. DTC yielded a weighted precision of 0.87, recall of 0.89, and F1-score of 0.88, while RFC recorded slightly higher precision (0.88) but a marginally lower recall (0.87) and F1-score (0.87). Specificity values were consistently high for both models (0.97), indicating reliable discrimination of negative instances. However, their confusion matrices revealed a few misclassifications, particularly in the "High" class—where DTC and RFC confused some instances with "Moderate" and "Very high" levels—reflecting the inherent difficulty in distinguishing between closely related resistance categories.

On the test set, CatBC maintained the strongest generalization performance, achieving an accuracy of 0.95 and balanced scores across all evaluation metrics: precision (0.86), recall (0.86), specificity (0.96), and F1-score (0.85). The test confusion matrix showed that CatBC correctly identified all instances of "Low", "High", "Very high", and "Extremely high" classes, with only three misclassifications involving the "Moderate" class—two misclassified as "High" and one as "Low". This highlights CatBC's ability to maintain both class balance and predictive consistency, even under imbalanced data conditions.



(**a**) DTC (**b**) RFC

(**c**) CatBC

Figure 5. Confusion matrix of the Tree-Based models for the test set.

In contrast, DTC achieved a test accuracy of 0.91 and an F1-score of 0.77. While it correctly predicted most "High" class instances (14 out of 20), it misclassified several "Low" and "Moderate" instances—specifically, one "Low" instance as "High", and two "Moderate" instances as "High" and "Extremely high". This drop in recall (0.80) and F1-score suggests some loss in generalization when compared to its training performance.

RFC recorded the lowest test performance among the three models, with an accuracy of 0.90 and F1-score of 0.72. Although the recall (0.73) and specificity (0.94) remained relatively high, the lower precision (0.74) indicates a higher number of false positives. The confusion matrix reveals that RFC struggled more than the other

models with the "Moderate" class, misclassifying two of its instances as "High" and one as "Low", while also showing minor errors in the "Low" and "Very high" classes.

In summary, CatBoost clearly outperformed the other classifiers, both in terms of raw metrics and error distribution. It was the only model to maintain consistent performance between the training and test sets, confirming the benefits of its native handling of class imbalance, regularization, and categorical encoding. While DTC and RFC achieved strong training metrics, their relatively lower test performance highlights their sensitivity to class overlap and the impact of data imbalance–despite the use of class weighting.

7. Feature Importance Analysis Using the CatBC model

To further interpret the decision logic of the best-performing model (CatBC), a comprehensive feature importance analysis was conducted using the model's built-in interpretation tools. Four types of importance measures were employed: Prediction Values Change, Loss Function Change, SHAP values, and Interaction importance. These complementary approaches offer different perspectives on the contribution of individual features and feature pairs to the model's predictive behavior.

The Prediction Values Change importance metric, which estimates the average change in prediction caused by varying each feature, ranked Superplasticizer as the most influential variable (13.29), followed closely by Coarse Aggregate (12.94), Fine Aggregate (12.19), Test Age (11.89), and the water-to-binder ratio (11.88). This indicates that the presence and dosage of chemical admixtures, aggregate proportions, and the maturity of concrete during testing substantially influenced the predicted resistance levels. Notably, Cement and Silica Fume also had moderate influence scores (11.52 and 8.70, respectively), while Fly Ash (5.93) and Slag (2.98) appeared to have lesser predictive impact in this configuration (see Figure 6).

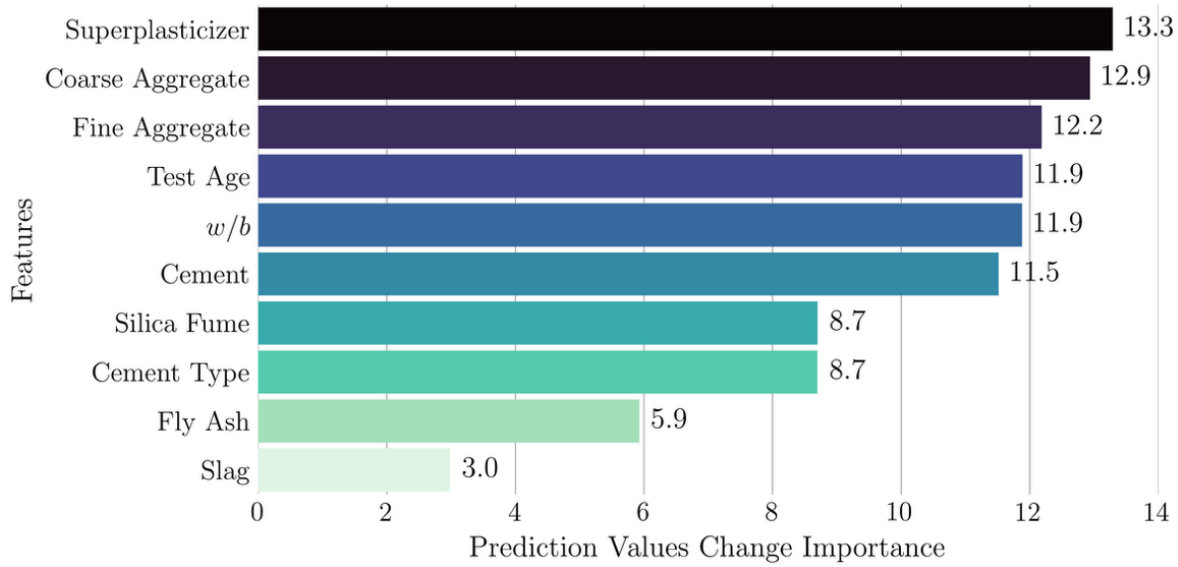


Figure 6. Feature Importance Based on Prediction Values Change.

When considering Loss Function Change–which assesses each feature's direct contribution to minimizing the model's loss function–Silica Fume emerged as the most critical factor (0.209), followed by *w*/*b* (0.112), Test Age (0.091), and Cement (0.083). Interestingly, Fly Ash and Cement Type exhibited negligible or even slightly negative importance scores (–0.002 and –0.017, respectively), suggesting that their inclusion offered limited improvement to the model's loss minimization objective (see Figure 7).

The SHAP value analysis, which estimates each feature's average marginal contribution to the model's prediction output, reinforced many of the earlier findings. Test Age was identified as the most influential variable (0.731), followed by *w*/*b* (0.666), Superplasticizer (0.650), and Silica Fume (0.618). SHAP values also highlighted Cement (0.547), Coarse Aggregate (0.517), and Fine Aggregate (0.444) as important drivers of prediction variability. These results suggest that durability-related properties (age and binder quality) and mixture composition play synergistic roles in determining chloride resistance (see Figure 8).

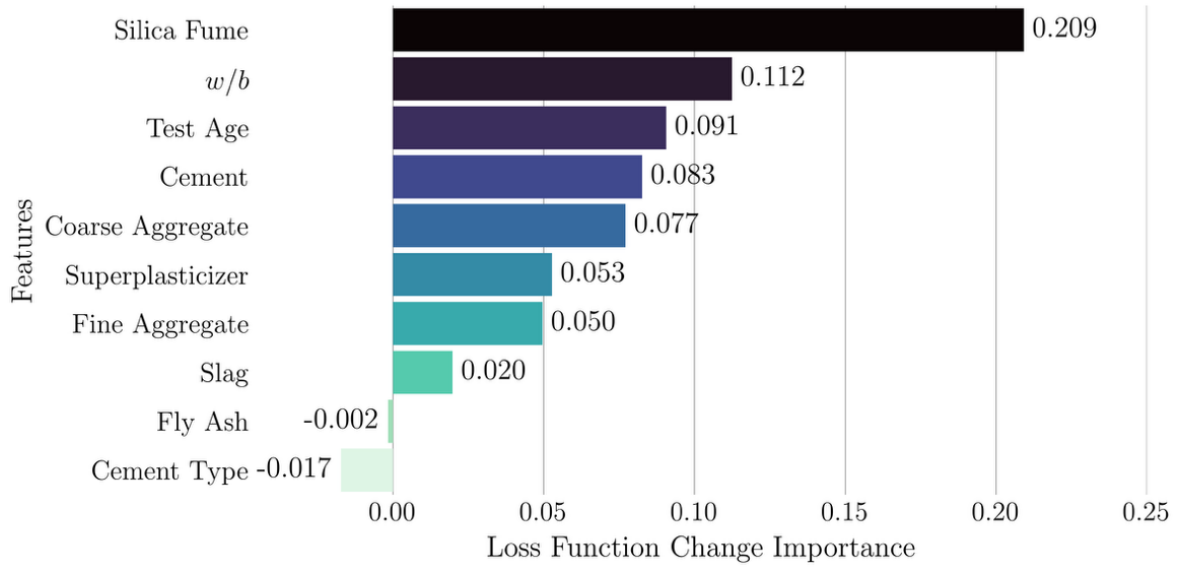


Figure 7. Feature importance based on loss function change.

Finally, the feature interaction analysis provided insight into how pairs of features influence predictions jointly. The most significant interaction was observed between *w*/*b* and Cement Type (importance = 3.91), followed by Cement Type & Cement (3.64) and Cement Type & Fine Aggregate (3.31). Other notable interactions included Fine Aggregate & Superplasticizer (3.09), *w*/*b* & Cement (3.02), and Cement & Coarse Aggregate (3.00). These interactions reveal that while individual features matter, their combined influence–especially involving cement type and mixture proportions–can be more critical in capturing the complex behavior of chloride resistance (see Figure 9).

In summary, all four importance analyses consistently emphasized the role of mix proportions (e.g., aggregates, cement, *w*/*b*), chemical admixtures, and curing age in determining chloride resistance. While the importance rankings varied slightly across methods, Test Age, *w*/*b*, and Superplasticizer repeatedly emerged as key predictors, supporting their central role in performance-based concrete durability design.

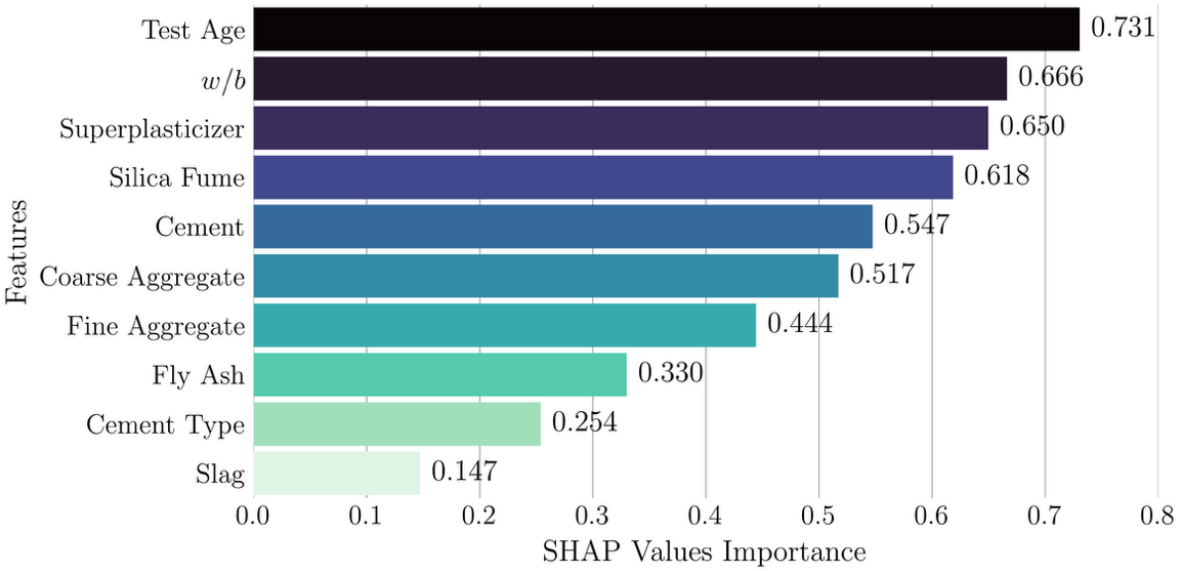


Figure 8. Feature importance based on SHAP values.

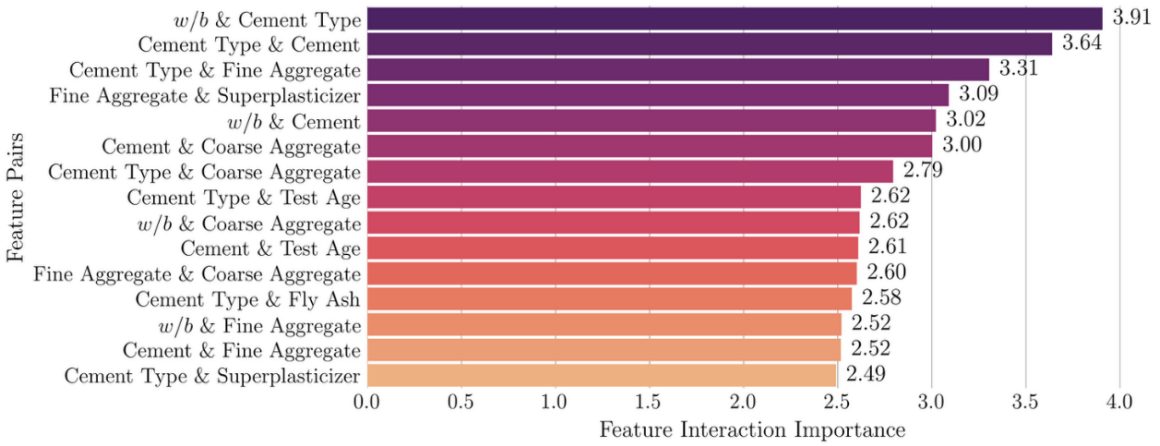


Figure 9. Top 15 pairwise feature interactions identified by the CatBC.

8. Limitations and Future Directions

This study demonstrates that optimized tree-based models, particularly CatBoost, can provide accurate and interpretable predictions of chloride resistance levels in concrete based on mix design and testing parameters. While the results are promising, several aspects merit further exploration to extend the scope of the present findings.

First, the retained dataset was relatively modest in size due to the adoption of a complete-case strategy. Although this ensured data consistency and prevented bias from imputation, future work could explore larger datasets and advanced imputation techniques to capture a broader range of concrete mix designs. Second, while multiclass classification provided valuable practical insights, future studies could investigate ordinal classification or regression approaches to leverage the inherent ordering of chloride resistance levels. Finally, external validation using independent datasets and uncertainty quantification methods would enhance the robustness and applicability of the proposed models for practical use in durability design.

By addressing these avenues, subsequent research can build upon the foundation established in this study, leading to more generalizable and widely applicable predictive tools for concrete durability assessment.

9. Conclusions

This study proposed an efficient and interpretable machine learning-based framework for predicting the chloride resistance level of concrete, aiming to reduce the dependence on time-consuming and resource-intensive laboratory tests. Leveraging a large and diverse experimental dataset, three tree-based classifiers—Decision Tree (DTC), Random Forest (RFC), and CatBoost (CatBC)—were developed and evaluated. To ensure optimal performance, each model underwent extensive hyperparameter tuning using the Optuna framework, with 2,000 optimization trials per model.

Among the three models, the CatBoost classifier demonstrated the most robust performance, achieving a test accuracy of 0.95 and a weighted F1-score of 0.85, outperforming both DTC and RFC in terms of generalization and stability. The use of native categorical feature handling, automatic class balancing, and advanced regularization techniques contributed to CatBC's superior predictive capabilities in an imbalanced multi-class classification setting. Furthermore, evaluation metrics and confusion matrices confirmed that CatBC was particularly effective in distinguishing between adjacent resistance levels, including underrepresented classes.

Beyond model performance, the study explored feature importance using multiple interpretability tools provided by CatBoost. All four importance measures–Prediction Values Change, Loss Function Change, SHAP values, and Interaction importance–consistently highlighted the significance of *w*/*b* ratio, test age, superplasticizer dosage, and aggregate content in influencing the chloride resistance level. These findings align with established durability design principles and offer data-driven support for optimizing concrete mix design in chloride-laden environments.

The proposed machine learning framework can serve as a valuable decision-support tool in the early stages of concrete mix design. By predicting chloride resistance levels directly from mix proportions and test age, engineers can rapidly screen alternative formulations before committing to resource-intensive experimental programs. This approach reduces testing costs, shortens design cycles, and supports the selection of mixes with improved durability performance at the outset. In practice, such a tool complements current design methods by providing data-driven insights that help prioritize promising mixtures, streamline laboratory validation, and enhance durability-focused design strategies in civil infrastructure projects. Future work may involve expanding

the dataset with more diverse environmental exposure conditions and further comparing ensemble learning approaches or deep learning techniques to explore additional performance gains.

Author Contributions

A.B.: Writing—original draft, Software, Methodology, Investigation; M.G., R.R., A.A.G. and P.G.A.: Writing—review & editing; M.G. and P.G.A.: Supervision, Conceptualization; R.R. and A.A.G.: Validation. All authors have read and agreed to the published version of the manuscript.

Funding

This research received no external funding

Institutional Review Board Statement

Not applicable.

Informed Consent Statement

Not applicable.

Data Availability Statement

Data will be made available on request.

Conflicts of Interest

The authors declare no conflict of interest. Given the role as Editor-in-Chief, Prof. Panagiotis G. Asteris had no involvement in the peer review of this paper and had no access to information regarding its peer-review process. Full responsibility for the editorial process of this paper was delegated to another editor of the journal.

References

1. Liu, J.; Ou, G.; Qiu, Q.; et al. Atmospheric Chloride Deposition in Field Concrete at Coastal Region. *Constr. Build. Mater.* **2018**, *190*, 1015–1022. https://doi.org/10.1016/J.CONBUILDMAT.2018.09.094.
2. Houska, C. Deicing Salt–Recognizing the Corrosion Threat. In *International Molybdenum Association*; TMR Consulting: Pittsburgh, PA, USA, 2007; pp. 1–10.
3. Pontes, J.; Bogas, J.A.; Real, S.; et al. The Rapid Chloride Migration Test in Assessing the Chloride Penetration Resistance of Normal and Lightweight Concrete. *Appl. Sci.* **2021**, *11*, 7251. https://doi.org/10.3390/APP11167251.
4. Elfmarkova, V.; Spiesz, P.; Brouwers, H.J.H. Determination of the Chloride Diffusion Coefficient in Blended Cement Mortars. *Cem. Concr.* **2015**, *78*, 190–199. https://doi.org/10.1016/J.CEMCONRES.2015.06.014.
5. Life-365™ Service Life Prediction Model ™ and Computer Program for Predicting the Service Life and Life-Cycle Cost of Reinforced Concrete Exposed to Chlorides. Available online: https://life-365.org/wp-content/uploads/2024/11/Life-365_v2.2.3_Users_Manual.pdf2020 (accessed on 24 September 2025).
6. Lindvall, A. Duracrete-Probabilistic Performance Based Durability Design of Concrete Structures. Available online: http://fib.bme.hu/proceedings/lindvall.pdf (accessed on 24 September 2025).
7. NT BUILD 492 Concrete, mortar and cement-based repair materials: Chloride migration coefficient from non-steady-state migration experiments. 1999. Available online: https://salmanco.com/wp-content/uploads/2018/10/NT-Build-492.pdf (accessed on 24 September 2025).
8. Tang, L.; Sørensen, H.E. Precision of the Nordic Test Methods for Measuring the Chloride Diffusion/Migration Coefficients of Concrete. *Mater. Struct.* **2001**, *34*, 479–485. https://doi.org/10.1007/BF02486496/METRICS.
9. Taffese, W.Z.; Espinosa-Leal, L. Prediction of Chloride Resistance Level of Concrete Using Machine Learning for Durability and Service Life Assessment of Building Structures. *J. Build. Eng.* **2022**, *60*, 105146. https://doi.org/10.1016/J.JOBE.2022.105146.
10. Sari-Ahmed, B.; Benzaamia, A.; Ghrici, M.; et al. Strength Prediction of Fiber-Reinforced Clay Soils Stabilized with Lime Using XGBoost Machine Learning. *Civ. Environ. Eng. Rep.* **2024**, *34*, 157–176. https://doi.org/10.59440/CEER/190062.
11. Benzaamia, A.; Ghrici, M.; Rebouh, R.; et al. Predicting the Compressive Strength of CFRP-Confined Concrete Using Deep Learning. *Eng. Struct.* **2024**, *319*, 118801. https://doi.org/10.1016/J.ENGSTRUCT.2024.118801.

12. Benzaamia, A.; Ghrici, M.; Rebouh, R.; et al. Shear Strength Modeling for Reinforced Concrete Beams Strengthened with Externally Bonded Fiber-Reinforced Polymer Using Machine Learning. *Structures* **2025**, *76*, 108954. https://doi.org/10.1016/J.ISTRUC.2025.108954.
13. Rebouh, R.; Benzaamia, A.; Ghrici, M. MLP Neural Networks for Compressive Strength Assessment of AFRP-Confined Concrete. *South Fla. J. Dev.* **2024**, *5*, e4765. https://doi.org/10.46932/sfjdv5n12-023.
14. Jifei, C.; Lin, B.; Pingping, R.; et al. Prediction Model of Chloride Erosion Concrete Based on Artificial Intelligence Algorithm. *Bull. Chin. Ceram. Soc.* **2024**, *43*, 439.
15. Nouri, Y.; Ghanizadeh, A.R.; Safi Jahanshahi, F.; et al. Data-Driven Prediction of Axial Compression Capacity of GFRP-Reinforced Concrete Column Using Soft Computing Methods. *J. Build. Eng.* **2025**, *101*, 111831. https://doi.org/10.1016/J.JOBE.2025.111831.
16. Raeisi, A.; Sharbatdar, M.K.; Naderpour, H.; et al. Flexural Capacity Prediction of RC Beams Strengthened in Terms of NSM System Using Soft Computing. *J. Soft Comput. Civ. Eng.* **2024**, *8*, 1–26. https://doi.org/10.22115/scce.2024.429316.1761.
17. Liu, K.H.; Zheng, J.K.; Pacheco-Torgal, F.; et al. Innovative Modeling Framework of Chloride Resistance of Recycled Aggregate Concrete Using Ensemble-Machine-Learning Methods. *Constr. Build. Mater.* **2022**, *337*, 127613. https://doi.org/10.1016/J.CONBUILDMAT.2022.127613.
18. Fakharian, P.; Nouri, Y.; Ghanizadeh, A.R.; et al. Bond Strength Prediction of Externally Bonded Reinforcement on Groove Method (EBROG) Using MARS-POA. *Compos. Struct.* **2024**, *349*, 118532. https://doi.org/10.1016/J.COMPSTRUCT.2024.118532.
19. Marks, M.; Glinicki, M.A.; Gibas, K. Prediction of the Chloride Resistance of Concrete Modified with High Calcium Fly Ash Using Machine Learning. *Materials* **2015**, *8*, 8714–8727. https://doi.org/10.3390/MA8125483.
20. Marks, M.; Jóźwiak-NiedzWiedzka, D.; Glinicki, M.A. Automatic Categorization of Chloride Migration into Concrete Modified with CFBC Ash. *Comput. Concr.* **2012**, *9*, 375–387. https://doi.org/10.12989/CAC.2012.9.5.375.
21. Hodhod, O.A.; Ahmed, H.I. Developing an Artificial Neural Network Model to Evaluate Chloride Diffusivity in High Performance Concrete. *HBRC J.* **2013**, *9*, 15–21.
22. Yao, L.; Ren, L.; Gong, G. Evaluation of Chloride Diffusion in Concrete Using PSO-BP and BP Neural Network. In *IOP Conference Series: Earth and Environmental Science*; IOP Publishing: Bristol, UK, 2021; Volume 687, p. 012037. https://doi.org/10.1088/1755-1315/687/1/012037.
23. Delgado, J.; Silva FA, N.; Azevedo, A.C.; et al. Artificial Neural Networks to Assess the Useful Life of Reinforced Concrete Elements Deteriorated by Accelerated Chloride Tests. *J. Build. Eng.* **2020**, *31*, 101445. https://doi.org/10.1016/J.JOBE.2020.101445.
24. Sari-Ahmed, B.; Ghrici, M.; Benzaamia, A.; et al. Assessment of Unconfined Compressive Strength of Stabilized Soil Using Artificial Intelligence Tools: A Scientometrics Review. *Stud. Syst. Decis. Control.* **2024**, *547*, 271–288. https://doi.org/10.1007/978-3-031-65976-8_15.
25. Benzaamia, A.; Ghrici, M.; Rebouh, R. Machine Learning Approaches for Predicting Compressive and Shear Strength of EB FRP-Reinforced Concrete Elements: A Comprehensive Review. In *New Advances in Soft Computing in Civil Engineering*; Springer: Berlin, Germany, 2024; pp. 221–249. https://doi.org/10.1007/978-3-031-65976-8_12.
26. Quinlan, J.R. Induction of Decision Trees. *Mach. Learn.* **1986**, *1*, 81–106. https://doi.org/10.1007/BF00116251.
27. Breiman, L. Random Forests. *Mach Learn* **2001**, *45*, 5–32. https://doi.org/10.1023/A:1010933404324/METRICS.
28. Prokhorenkova, L.; Gusev, G.; Vorobev, A.; et al. CatBoost: Unbiased Boosting with Categorical Features. In Proceedings of the Advances in Neural Information Processing Systems, Long Beach, CA, USA, 4–9 December 2017; pp. 6638–6648.
29. Akiba, T.; Sano, S.; Yanase, T.; et al. Optuna: A Next-Generation Hyperparameter Optimization Framework. In Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Anchorage, AK, USA, 4–8 August 2019; pp. 2623–2631. https://doi.org/10.1145/3292500.3330701.
30. Rebouh, R.; Benzaamia, A.; Ghrici, M. Bayesian-Optimized Tree-Based Models for Predicting the Shear Strength of U-Shaped Externally Bonded FRP-Strengthened RC Beams. *Asian J. Civ. Eng.* **2025**, *26*, 1465–1478.
31. Benzaamia, A.; Ghrici, M.; Rebouh, R.; et al. Predicting the Shear Strength of Rectangular RC Beams Strengthened with Externally-Bonded FRP Composites Using Constrained Monotonic Neural Networks. *Eng. Struct.* **2024**, *313*, 118192. https://doi.org/10.1016/J.ENGSTRUCT.2024.118192.
32. Kechroud, F.; Benzaamia, A.; Ghrici, M. Optimized Regression-Based Machine Learning Models for Predicting Chloride Diffusion in Concrete. *Asian J. Civ. Eng.* **2025**, *26*, 2513–2526. https://doi.org/10.1007/S42107-025-01326-7/METRICS.
33. Ali Aichouba, A.; Benzaamia, A.; Ezziane, M.; et al. TabNet-Based Prediction of Residual Compressive and Flexural Strengths in Hybrid Fiber-Reinforced Self-Compacting Concrete (HFR-SCC) Exposed to Elevated Temperatures. *Asian J. Civ. Eng.* **2025**, *26*, 3705–3724. https://doi.org/10.1007/S42107-025-01392-X/METRICS.