water

MDPI

Article

Aquaculture Water Quality Classification Using XGBoost ClassifierModel Optimized by the Honey Badger Algorithm with SHAP and DiCE-Based Explanations

S M Naim [1], Prosenjit Das [1], Jun-Jiat Tiang [2,*] and Abdullah-Al Nahid [1,*]

[1] Electronics and Communication Engineering Discipline, Khulna University, Khulna 9208, Bangladesh; prosenjitdas0929@gmail.com (P.D.)
[2] Centre for Wireless Technology, CoE for Intelligent Network, Faculty of Artificial Intelligence & Engineering, Multimedia University, Persiaran Multimedia, Cyberjaya 63100, Selangor, Malaysia
* Correspondence: jjtiang@mmu.edu.my (J.-J.T.); nahid.ece.ku@gmail.com (A.-A.N.)

Abstract

Water quality is an essential part of maintaining a healthy environment for fish farming. The quality of the water is related to a few of the chemical and biological characteristics of water. The conventional evaluation methods of the water quality are often time-consuming and may overlook complex interdependencies among multiple indicators. This study has proposed a robust machine learning framework for aquaculture water quality classification by integrating the Honey Badger Algorithm (HBA) with the XGBoost classifier. The framework enhances classification accuracy and incorporates explainability through SHAP and DiCE, thereby providing both predictive performance and transparency for practical water quality management. For reliability, the dataset has been randomly shuffled, and a custom 5-fold cross-validation strategy has been applied. Later, through the metaheuristic-based HBA, feature selections and hyperparameter tuning have been performed to improve and increase the prediction accuracy. The highest accuracy of 98.45% has been achieved by a particular fold, whereas the average accuracy is 98.05% across all folds, indicating the model's stability. SHAP analysis reveals Ammonia, Nitrite, DO, Turbidity, BOD, Temperature, pH, and CO_2 as the topmost water quality indicators. Finally, the DiCE analysis has analyzed that Temperature, Turbidity, DO, BOD, CO_2, pH, Ammonia, and Nitrite are more influential parameters of water quality.

Keywords: water; fish; machine learning; XGBoost; honey badger algorithm; SHAP; DiCE

check for updates

Academic Editor: Christos S. Akratos

Received: 9 September 2025
Revised: 15 October 2025
Accepted: 15 October 2025
Published: 16 October 2025

Citation: Naim, S. M.; Das, P.; Tiang, J.-J.; Nahid, A.-A. Aquaculture Water Quality Classification Using XGBoost Classifier Model Optimized by the Honey Badger Algorithm with SHAP and DiCE-Based Explanations. *Water* **2025**, *17*, 2993. https://doi.org/10.3390/w17202993

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Water quality (WQ) includes the chemical, physical, biological, and radiological characteristics of water [1]. Water is an important factor in the life of all aquatic species, influencing ecosystem health, biodiversity, and the sustainability of aquatic habitats. In this case, WQ is a crucial factor for healthy fish farming. The primary objective of the United Nations for 2030 is to ensure good health and a healthy environment for all forms of life [2]. Ensuring good water quality is crucial for enhancing fishery productivity [3] because poor WQ can affect the health and growth of fish. In this context, fish farmers pay attention to the ecology of WQ. Monitoring various WQ parameters and ensuring they remain within their ideal range is crucial for the growth and survival of fish [4]. Fish in all aquaculture systems engage in physiological processes, such as respiration, waste elimination, feeding, maintaining salt balance, and reproduction within the water environment. Therefore, the

overall effectiveness of any aquaculture system is, in part, influenced by its water quality parameters [5]. Parameters of WQ that are frequently observed in the aquaculture sector consist of temperature, dissolved oxygen (DO), biochemical oxygen demand (BOD), pH, alkalinity, hardness, ammonia, and nitrites [6]. Additionally, depending on the type of culture system, carbon dioxide, chlorides, and salinity may also be tracked [6]. The ranges of chemical and biological parameters are essential for the growth of fish in ponds and rivers.

In aquaculture, temperature is an important parameter that strongly influences the growth, fecundity, and overall health of fish species [7]. Temperature significantly affects all biological and chemical processes within an aquaculture operation [8]. Boyd (1982) has indicated that a water temperature range of 26.06 to 31.97 °C is appropriate for the culture of warm-water fish [9], such as Tilapia, Catfish, and Carp. Bolorunduro et al. have demonstrated that an ideal temperature range for tropical fish culture lies between 25 and 32 °C [10]. Siti-zahrah et al. (2004, 2008) found that water temperatures exceeding 30 °C lead to a significant increase in mortality rates for tilapia in cage culture at the Tasik Kenyir reservoir in Malaysia [11,12].

P^H is a crucial parameter, which plays a significant role for fish productivity. Hepher et al. (1981) have indicated that this range of 6.5 to 9.0 is suitable for fish farming [13].

DO is an important parameter for the assessment of water quality. Nsonga (2014) has reported that a dissolved oxygen (DO) level of 6.5 mg/L or higher, or at least 5 mg/L, is optimal for warm water fish species [14]. Daniel et al. (2005) have indicated that dissolved oxygen levels lower than 3.5 mg/L are not suitable for fish farming [15].

BOD is a crucial factor in assessing the pollution level of the water body. BOD signifies the portion of soluble organic material that is broken down and readily absorbed by bacteria, reflecting the measurable amount of biodegradable organic matter in water [8]. Elevated BOD levels decrease the dissolved oxygen, resulting in unpleasant smells and an unhealthy ecosystem [8]. Tamot et al. (2008) have proposed that a higher value of BOD level occurs because the high temperatures create favorable conditions for microbial activity [16].

Nitrate is produced during the nitrification process, where aerobic bacteria oxidize nitrite (NO_2) into nitrate (NO_3). Boyd (1998) has stated that the preferred nitrate concentration for aquaculture ranges from 0.2 to 10 mg/L [9]. In addition, ammonia is a WQ parameter that can affect aquatic processes. Ammonia is the primary nitrogenous waste generated through the metabolism of aquatic animals and is mainly excreted across the gills. Daniel et al. (2005) concluded that an ammonia concentration of more than 0.2 mg/L is undesirable for fish farming[15]. Other parameters of WQ, such as hardness, alkalinity, carbon dioxide, etc., are also related to aquatic fish farming. In this context, early detection and monitoring of WQ is a very important part for good fish farming and produces healthy fish. For this purpose, machine learning (ML) is an important tool for assessing WQ. ML is suitable for identifying the water condition for fish farming. Some previous studies have applied machine learning and IoT-based approaches to aquaculture water quality analysis. For instance, Zulfikhry et al. developed an IoT-enabled water-quality-monitoring system, where key parameters, including temperature, turbidity, dissolved oxygen, and water level, were analyzed using a Random Forest Classifier for predictive analytics [17].A. A. Nayan et al. conducted research on the quality of river water for agricultural and fishing purposes and have identified fish diseases resulting from alterations in water quality through the use of machine learning [18,19]. They evaluated water quality based on pH, DO, BOD, chemical oxygen demand (COD), total suspended solids (TSS), total dissolved solids (TDS), electrical conductivity (EC), phosphate (PO_4^{3-}), nitrate nitrogen (NO_3-N), and ammonia nitrogen (NH_3-N), and employed a boosting technique to predict the outcomes. S. Sen et al. implemented and tested several machine learning models to predict water quality [20]. They achieved satisfactory results for various attributes, including pH, hardness,

solids, chloramines, sulfate, conductivity, organic carbon, trihalomethanes, turbidity, and potability.

In recent years, ML has become popular in WQ classification because it allows for effective monitoring, prediction, and decision-making regarding environmental management. Recent studies have used ML algorithms, such as Decision Tree (DT), Support Vector Machines (SVMs), Random Forests (RFs), Gradient Boosting (GB), Extreme Gradient Boosting (XGB), and Categorical Boosting (CB), for classifying WQ. A number of studies have been conducted for this purpose. However, Mahmoud et al. used the Indian WQ dataset and applied four classification algorithms (RF, XGBoost, AdaBoost, and GB) to classify WQ [21]. Among four classifiers, the GB classifier achieved the best results, with an accuracy of 99.50%, F1-score of 99.40%, precision of 99.50%, and recall of 99.50% [21]. In addition, Nasir et al. performed CatBoost classifier, attaining an accuracy of 94.51% [22]. Nur et al. selected the Malaysian WQ dataset and utilized the GB algorithm, achieving the highest accuracy of 94.90% [23]. Khan et al. chose the Bangladeshi WQ dataset and applied a GB classifier algorithm, attaining a classification accuracy of 100% [24]. Ho et al. applied DT classifier and achieved an accuracy of 81% [25]. Uddin et al. used the XGBoost classifier, where the result indicates that it outperformed the others, achieving accuracy of 100% [26]. Overall, these previous studies focused on classifying water quality with machine learning techniques. Finally, all the studies' results have been presented in Table 1. According to previous studies, we observed that the ML model improved classification performance. In addition, ML provides effective tools for early detection of water quality issues and continuous monitoring in aquaculture systems. In this case, metaheuristic algorithms integrated with ML classifiers can be applied, and deep learning and other techniques can be used for classification cases. Previous studies have not integrated metaheuristic optimization with the XGBoost classifier nor employed SHAP and DiCE for interpretability.

Table 1. Summary of the previous studies of WQ classification.

Author	Dataset Information	Classification Model	Accuracy (%)
Mahmoud et al. [21]	Indian Water Quality Dataset	GB	99.50
Nasir et al. [22]	Indian Water Quality Dataset	CB	94.51
Nur et al. [23]	Department of Environment, Malaysia	GB	94.90
Khan et al. [24]	Gulshan Lake Water Quality Dataset	GB	100
Ho et al. [25]	Klang River Water Quality Dataset	DT	81
Uddin et al. [26]	Lee, Cork Harbor, and Youghal Bay	XGB	100

In this research, we used the aquaculture WQ dataset to classify the WQ. In addition, we applied a metaheuristic algorithm for hyperparameter tuning and feature selection. We analyzed feature importance using SHapley Additive exPlanations (SHAP) to identify which parameters are mainly associated with WQ. Finally, we applied Diverse Counterfactual Explanations (DiCEs) to observe the influential parameters of WQ. SHAP and DiCE were used in this research; neither of these were directly considered in the prior studies. In this study, our research question is how effectively can an XGBoost classifier optimized by the Honey Badger Algorithm classify aquaculture water quality and identify key influencing parameters using SHAP and DiCE explanations?

2. Materials and Methods

This research has executed a multi-class classification task to differentiate Excellent (E), Good (G), and Poor (P) classes based on the WQ dataset. The WQ dataset is a labeled dataset, so we have considered a supervised machine learning procedure to solve this problem. Firstly, we have randomly shuffled our dataset. We have used a custom 5-fold cross-validation to partition our dataset, in which each fold contains 70% data for training cases and 30% data for test cases. In addition, we have applied the metaheuristic-based Honey Badger Algorithm (HBA) with the Extreme Gradient Boosting (XGBoost) classifier for optimization and feature selection purposes. We have used the HBA_XGB classifier for performance analysis. For interpretability, we have employed the SHAP analysis to assess feature importance and utilized DiCE analysis for counterfactual explanations. The overall method is represented in Figure 1. In this study, we have used metaheuristic-based HBA for optimization and feature selection purposes, as previously demonstrated in biomedical applications [27]. In addition, Weiling et al. have applied XGBoost-SHAP to analyze key influencing factors of water quality [28]. Simona et al. have used DiCE to improve sales and e-commerce strategies [29]. These examples show that each technique used in our study is established and has previously been used in other studies.

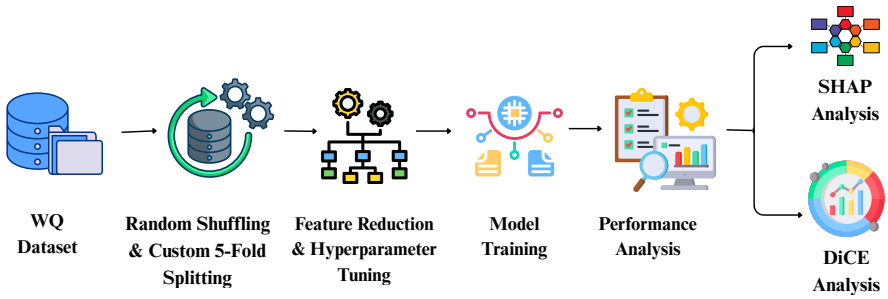


Figure 1. Overall methodology of this study.

2.1. Dataset

In this study, we have used the aquaculture WQ dataset. This dataset can be found in Mendeley Data and at the following URL: https://data.mendeley.com/datasets/y78ty2g293/1, accessed on 20 August 2025. This dataset contains 4300 samples with 14 input features and one output label column. The label column of the WQ dataset consists of 3 classes, in which class E (0), G (1), and P (2). This dataset includes a total of 1500 water samples of poor quality, 1400 water samples of excellent quality, and 1400 water samples of good quality. The features of the dataset are Temperature, Turbidity, DO, BOD, CO_2, pH, Alkalinity, Hardness, Calcium, Ammonia, Nitrite, Phosphorus, H_2S, and Plankton.

According to the Figure 2, the water quality is classified into three categories (Excellent, Good, and Poor), showing significant variation across several parameters. DO and BOD show the most distinct differences, where DO remains comparatively stable, and BOD is minimal in the Excellent category, but both show substantial variation and higher values in the Poor water-quality category. Similarly, turbidity, CO_2, calcium, ammonia, nitrite, and H_2S are noticeably higher in the Poor water-quality category, showing signs of pollution. In contrast, pH and temperature stay fairly consistent across all three levels. This suggests that these variables alone are not strong signs of water quality. Plankton counts are comparatively higher in Excellent and Good water, whereas they decline in Poor water, reflecting ecological imbalance. In general, DO, BOD, Turbidity, Ammonia, and Plankton become the most significant water quality indicators, whereas pH and Temperature are slightly influential indicators.

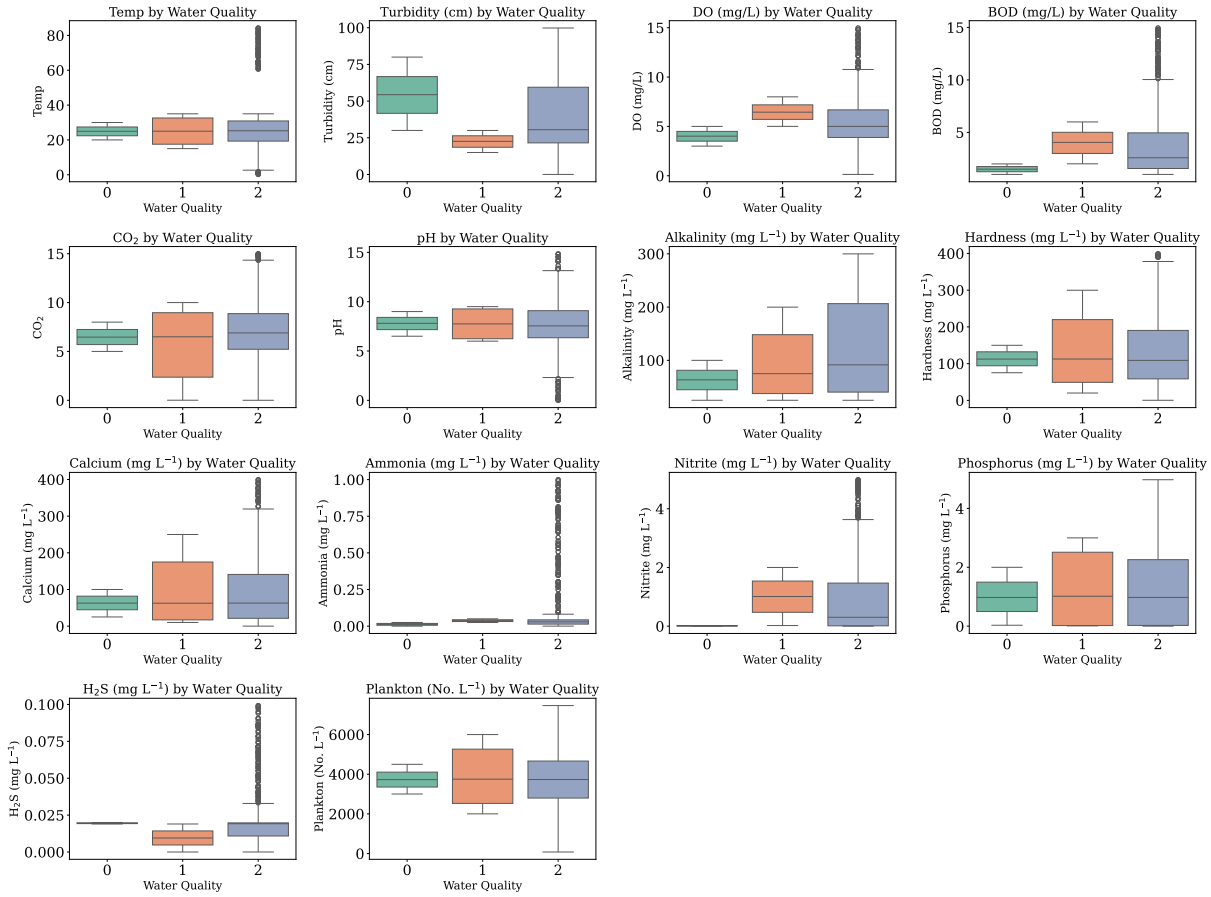


Figure 2. Boxplot representation of different physico-chemical and biological parameters across three water quality classes (0 = Excellent, 1 = Good, 2 = Poor).

In addition, the mean, maximum, and minimum values of each feature for each target class are provided in Table 2.

Table 2. Summary of the dataset based on the target class (Excellent, Good, and Poor).

Feature	Index	Excellent (0)			Good (1)			Poor (2)		
		Mean	Max	Min	Mean	Max	Min	Mean	Max	Min
Temp (°C) (F1)	0	24.92	29.99	20.00	25.02	34.99	15.00	27.04	84.25	0.19
Turbidity (cm) (F2)	1	54.46	79.98	30.02	22.53	29.99	15.00	40.08	99.79	0.05
DO (mg/L) (F3)	2	4.01	5.00	3.00	6.46	7.99	5.00	5.43	14.97	0.13
BOD (mg/L) (F4)	3	1.50	2.00	1.00	4.02	5.99	2.01	3.81	14.94	1.00
CO2 (mg/L) (F5)	4	6.48	7.99	5.00	5.72	9.99	0.01	6.90	14.98	0.00
pH (F6)	5	7.79	8.99	6.50	7.75	9.49	6.00	7.61	14.85	0.00
Alkalinity (mg/L) (F7)	6	63.02	99.99	25.03	93.19	199.90	25.03	122.86	299.91	25.01
Hardness (mg/L) (F8)	7	112.78	149.92	75.25	135.45	299.75	20.01	132.54	398.80	0.26
Calcium (mg/L) (F9)	8	62.69	99.99	25.08	96.93	249.94	10.08	94.32	399.32	0.02
Ammonia (mg/L) (F10)	9	0.012	0.025	0.000	0.037	0.050	0.025	0.092	0.999	0.000
Nitrite (mg/L) (F11)	10	0.010	0.020	0.000	1.012	2.000	0.020	0.889	4.990	0.000
Phosphorus (mg/L) (F12)	11	0.998	1.999	0.031	1.264	2.999	0.010	1.251	4.974	0.000
H2S (mg/L) (F13)	12	0.019	0.020	0.019	0.010	0.019	0.000	0.020	0.099	0.000
Plankton (No./L) (F14)	13	3728.80	4498.68	3002.30	3888.95	5999.20	2002.15	3799.23	7460.42	78.60

2.2. Custom Cross-Validation

In this study, we have used a custom cross-validation strategy. The custom cross-validation strategy has been designed such that, in every fold, 70% of the data is allocated for training and 30% for testing.

2.3. Classifier Model

In this study, we used the XGBoost classifier to classify the target variable. The XGBoost algorithm has been widely used in recent years as an effective classifier in various fields of application [30]. XGBoost is an efficient and fully scalable tree-boosting system, developed by Tianqi Chen and Carlos Guestrin [31], which has been widely used by data scientists to attain cutting-edge performance in numerous machine learning tasks. The algorithm is constructed based on the gradient-boosting framework, whereby a number of weak learners, usually decision trees, are pooled together to create a robust and highly accurate predictor model [32].

XGBoost (XGB) classifier has some hyperparameters that include the number of estimators, learning rate, maximum tree depth, and minimum child weight, which regulate the learning process and complexity of the model. The XGB model could be represented mathematically as expressed in the formula below (Equation (1)): the goal is to minimize the loss function $L(\phi)$ [31].

$$\hat{y}_i = \phi(x_i) = \sum_{k=1}^{K} f_k(x_i), \quad f_k \in \mathcal{F} \tag{1}$$

Here, f_k denotes an independent decision tree, $\mathcal{F}$ represents the functional space of regression trees, x_i corresponds to the input variables, and K is the number of additive functions. The overall objective function of the XGBoost algorithm is defined as follows:

$$L(\phi) = \sum_i l(\hat{y}_i, y_i) + \sum_k \Omega(f_k) \tag{2}$$

where $L(\phi)$ is the regularized loss function, $l(\hat{y}_i, y_i)$ measures the difference between the predicted value $\hat{y}_i$ and the actual target y_i, and $\Omega(f_k)$ is the regularization term that penalizes model complexity by accounting for both the number of leaves and their corresponding scores.

2.4. Metaheuristic-Based HBA

HBA is a nature-inspired population-based metaheuristic optimization algorithm originally proposed by Hashim et al. in 2022 [33]. HBA is inspired by the advanced foraging behaviors of the honey badger, that is, its honey-finding and digging behaviors, to reconcile exploration and exploitation in the search process. Controlled randomization and dynamic search mechanics ensure population diversity, which is important in navigating complex optimization landscapes. We have chosen the HBA because it effectively solves complex optimization problems, outperforming other methods in convergence speed and exploration–exploitation balance [33].

HBA models the search process as agents (honey badgers) moving in a multidimensional search space, guided by smell intensity (I), distance (d), and density factor (α), which have been mathematically represented below:

$$I_i = r^2 \times 4\pi d_i^2 S, \tag{3}$$

$$S = (x_i - x_{i+1})^2, \tag{4}$$

$$d_i = x_{\text{prey}} - x_i, \tag{5}$$

$$\alpha = C \times (1 - T_t) \tag{6}$$

where I_i represents the intensity or influence of each agent, reflecting its hunting or searching capability. In this context, r is a random scalar associated with distance, d_i denotes the distance from the agent's current position to the prey or target, and S represents the

difference between two neighboring solutions in the search space, where x_i is the current position, and x_{i+1} is the neighboring position. C is a constant weight, and T_t is the normalized iteration count.

2.5. Custom-Defined Problem

A custom-defined problem (CDP) with given boundary conditions has been solved by MHAs in combination with a classifier and feature selector. This stepwise procedure has been enabling the identification of optimal features and hyperparameters for the XGB classifier, as shown in Table 3.

Table 3. Hyperparameter ranges for model tuning.

Hyperparameter	Type	Range
Number of Estimators (n_estimators)	Integer	100 to 300
Learning Rate (learning_rate)	Float	0.001 to 0.2
Max Depth (max_depth)	Integer	3 to 7
Minimum Child Weight (min_child_weight)	Integer	1 to 10

2.6. Performance Analysis

This part examines Accuracy, F-score, Precision, and Recall to assess the effectiveness of the model. We have also evaluated the Confusion Matrix to gain valuable insights into the classification performance. The Confusion Matrix can illustrate the counts of True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN) for the models. TP, TN, FP, and FN are explained below:

- **TP** : Correctly predicted positive samples;
- **TN**: Correctly predicted negative samples;
- **FP**: Predicted positive but actually negative;
- **FN**: Predicted negative but actually positive.

The metrics are computed in the following manner:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \tag{7}$$

$$\text{Precision} = \frac{TP}{TP + FP} \tag{8}$$

$$\text{Recall} = \frac{TP}{TP + FN} \tag{9}$$

$$\text{F-score} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{10}$$

2.7. SHAP

SHAP is widely recognized as a popular tool in Explainable Artificial Intelligence (XAI) for understanding the results produced by machine learning models. It calculates the impact of each feature on a particular prediction by giving it an importance score [34]. In this research, we have used SHAP to analyze the feature importance of WQ parameters.

2.8. DiCE

DiCE is a technique that can be applied to any model, generating a variety of credible counterfactual instances to aid users in understanding how changes to input features can result in varying model predictions [35]. The optimization goal of DiCE is expressed in Equation (11), where $C(x)$ represents the collection of k counterfactuals created for a specific input x:

$$C(x) = \arg\min_{c_1,\ldots,c_k} \frac{1}{k}\sum_{i=1}^{k} y_{\text{loss}}(f(c_i), y) + \lambda_1 \frac{1}{k}\sum_{i=1}^{k} \text{dist}(c_i, x) - \lambda_2 \,\text{dpp_diversity}(c_1, \ldots, c_k) \quad (11)$$

In this approach, the term $y_{\text{loss}}(f(c_i), y)$ represents the classification loss, guiding each counterfactual c_i towards the intended target class y. The proximity element $\text{dist}(c_i, x)$ prevents the generation of counterfactuals that deviate significantly from the original instance x, thus preserving realism. Lastly, $\text{dpp_diversity}(c_1, \ldots, c_k)$ utilizes Determinantal Point Processes (DPPs) to promote variety among the generated counterfactuals, ensuring that the resulting set provides several distinct and practical alternatives. The equation of dpp_diversity is represented below:

$$\text{dpp_diversity} = \det(K), \quad (12)$$

where

$$K_{i,j} = \frac{1}{1 + \text{dist}(c_i, c_j)}$$

and $\text{dist}(c_i, c_j)$ and $\text{dist}(c_i, c_j)$ denote a distance metric between the two counterfactual examples. The hyperparameters λ_1 and λ_2 regulate the balance between proximity and diversity.

3. Results

In this section, we divide our results into seven parts: Custom k-fold validation-based evaluation of the HBA_XGB classifier, Performance of HBA_XGB model, Confusion matrix analysis, Misclassifications output, ROC curve analysis, SHAP analysis, and DiCE analysis. Custom k-fold validation–based evaluation of the HBA_XGB classifier section has explained optimal feature set selected by the HBA and the corresponding hyperparameter values for the classifier model of each fold. Performance of HBA_XGB model section has evaluated model output for training and test cases. In addition, confusion matrix and roc have been analyzed. Finally, SHAP analysis have used for analyzing the feature importance.

3.1. Custom k-Fold Validation-Based Evaluation of the HBA_GB Classifier

Figure 3 illustrates the cost function behavior of the optimization process over 50 epochs for all five folds of the custom cross-validation experiment. "Epochs" refer to the total number of times the learning algorithm goes through the entire training dataset during model training. One epoch means that every training sample has been used once to update the model's parameters. The cost function plots across all five folds demonstrate that the fitness score remained consistently at 1.000 over 50 epochs on the training data using the F1-score. This indicates that the optimization algorithm rapidly achieved an optimal solution without fluctuation. The uniformity across folds highlights the stability and robustness of the model.

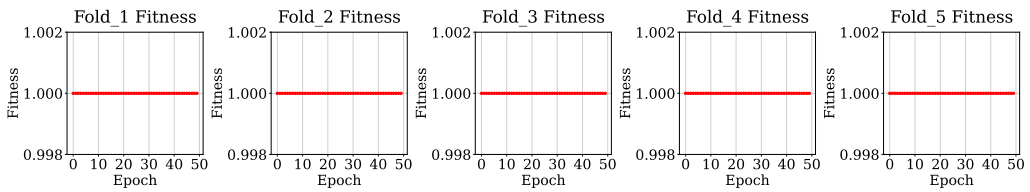


Figure 3. Graphical representation of the cost function of each fold.

Table 4 shows the hyperparameters and the best feature results of each fold. In this case, HBA has demonstrated that the best hyperparameters are identical for all five folds,

with n_estimators, learning_rate of 0.2, max_depth of 7, and min_child_weight of 10. In addition, selected feature index (SFI) of fold = 0, 1, 2, 3, 4, 5, 9, 10.

Table 4. Optimal features and hyperparameters for classifiers of each fold.

Fold No.	n_estimators	Learning_rate	Max_depth	Min_child_weight	SFI
1	300	0.2	7	10	0, 1, 2, 3, 4, 5, 9, 10
2	300	0.2	7	10	0, 1, 2, 3, 4, 5, 9, 10
3	300	0.2	7	10	0, 1, 2, 3, 4, 5, 9, 10
4	300	0.2	7	10	0, 1, 2, 3, 4, 5, 9, 10
5	300	0.2	7	10	0, 1, 2, 3, 4, 5, 9, 10

3.2. Performance of HBA_XGB Model

The results of the proposed HBA_XGB classifier have been shown in Table 5. It is important to note that, at the training stage, the model reported 100% accuracy, F1-score, precisions, and recalls in all five folds in training the model and should, therefore, be expected to perform similarly on the test data.

Table 5. HBA_XGB classifier output.

Fold	Train				Test			
	Acc. (%)	F1-Score. (%)	Pre. (%)	Rec. (%)	Acc. (%)	F1-Score (%)	Pre. (%)	Rec. (%)
Fold-1	100	100	100	100	97.67	97.71	97.62	97.89
Fold-2	100	100	100	100	98.45	98.46	98.49	98.46
Fold-3	100	100	100	100	98.22	98.22	98.23	98.26
Fold-4	100	100	100	100	97.52	97.56	97.50	97.68
Fold-5	100	100	100	100	98.37	98.37	98.38	98.38
Mean	100	100	100	100	98.05	98.07	98.04	98.13

In fold 1, the HBA_XGB has performed with a testing accuracy of 97.67% and an F1-score of 97.71%. Precision and recall values of 97.62% and 97.89% have been obtained, respectively. This fold has shown that the model recorded a good balance between precision and recall.

In the case of fold 2, the accuracy of the classifier was 98.45% on a test. The F1-score achieved the value of 98.46%, precision was a bit higher, 98.49%, and recall was the same, 98.46%. The model has consequently been showing a very consistent and dependable performance.

In fold 3, the testing accuracy has become 98.22%, with a 98.22% F1-score. Precision and recall have been reported as 98.23% and 98.26%, respectively. This fold has proved that the classifier has been sustaining almost similar values in all the evaluation metrics.

Fold 4 attained a testing accuracy of 97.52%, which has been the worst of all folds, but still very high. F1-score has achieved 97.56%, precision of 97.50% and recall of 97.68%. It is slightly lower, but the results have shown that the classifier has been performing well and has generalized quite well.

In fold 5, the classifier achieved a testing accuracy of 98.37%. The F1-score has as well, whereas the precision and recall achieved 98.38% each. This fold has demonstrated that the classifier has been maintaining good and balanced results according to metrics.

Finally, the mean test result of all folds' accuracy, F1-score, precision, and recall have been noted as 98.05%, 98.07%, 98.04%, and 98.13%, respectively. These findings have

confirmed that the integration of HBA for feature optimization with the XGB classifier has significantly enhanced the classification performance.

3.3. Confusion Matrix Analysis

In this case, we analyzed the confusion matrix. The fold-wise confusion matrix of the HBA_XGB classifier under custom 5-fold cross-validation is represented in Figure 4.

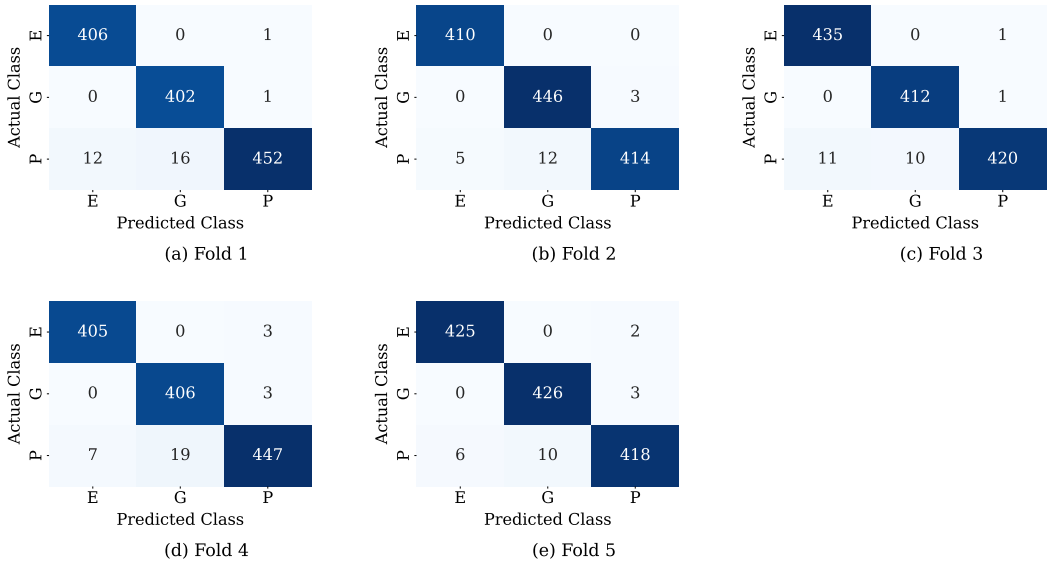


Figure 4. Fold-wise visualization of confusion matrices for the HBA_XGB classifier (E = Excellent, G = Good, and P = Poor water quality class).

In fold 1, the HBA_XGB correctly predicted 406 instances of class E, 402 instances of class G, and 452 instances of class P. Nonetheless, it falsely classified 1 case of E as P, 1 case of G as P, 12 cases of P as E, and 16 cases of P as G. Therefore, the overall accuracy of the model has been up to the mark, but it has continued to fail at identifying several samples of the class P. In fold 2, HBA_XGB has obtained the right classification of 410 on E, 446 on G, and 414 on P. On the one hand, it has failed to correctly locate three cases of G as P, five cases of P as E, and twelve cases of P as G. The model has, as such, retained good results on classes E and G but has still not been perfect in the classification of P. In fold 3, HBA_XGB classifier correctly recognized 435 E, 412 G, and 420 P instances. However, it falsely clustered 1 E with P, 1 G with P, 11 P with E, and 10 P with G. The same has been reflected consistently in the misclassifications of the class P, and this needs to be addressed, as it should have been easy to identify a class. In fold 4, HBA_XGB correctly classified 405 instances of E, 406 of G, and 447 of P. Still, it has misclassified three cases of E as P, three cases of G as P, seven cases of P as E, and nineteen cases of P as G. This fold has exhibited a relatively greater misclassifications in class P further evidencing that P was the most problematic class in correct classification. Lastly, in fold 5, the HBA_XGB classifier has correctly predicted 425 instances of class E, 426 instances of class G, and 418 instances of class P. Misclassifications have included two E instances predicted as P, three G instances predicted as P, six P instances predicted as E, and ten P instances predicted as G. Although most of the predictions were quite accurate, the persistence of the problem with the class P has remained consistent throughout all folds.

3.4. Misclassification Output

A summary of misclassification instances is been presented in this section. Table 6 shows the misclassified instances in each fold. The number of misclassified samples and their indexing of instance for each fold have been reported. In fold 1, 30 samples have

been misclassified. In addition, fold 2 and fold 3 have obtained, respectively, 21 and 20 samples misclassified. However, fold 4 and fold have presented two different subsets of misclassifications of 32 and 21 cases, respectively.

Table 6. Misclassified instances across folds.

Fold No.	Number of Misclassified	Misclassified Instances
1	30	103, 203, 205, 227, 238, 257, 258, 263, 278, 283, 287, 290, 299, 309, 318, 321, 325, 331, 342, 386, 505, 618, 712, 761, 785, 790, 1180, 1374, 2801, 3028
2	20	208, 230, 237, 243, 323, 344, 347, 365, 374, 382, 383, 917, 1249, 1364, 1446, 1452, 1459, 1826, 2648, 2718
3	23	117, 148, 248, 267, 302, 317, 333, 339, 363, 373, 384, 385, 388, 395, 396, 398, 730, 777, 824, 943, 1013, 1821, 3068
4	32	103, 134, 149, 200, 205, 206, 227, 238, 240, 245, 301, 309, 318, 321, 325, 397, 505, 618, 656, 761, 790, 971, 1180, 1210, 1267, 1374, 2250, 2524, 2565, 3537, 3777, 4145
5	21	221, 230, 237, 243, 257, 258, 283, 347, 382, 386, 547, 1132, 1249, 1446, 1452, 1459, 1826, 2648, 2748, 2949, 4168

In addition, Figure 5 illustrates the percentage of misclassification for each fold. Fold 4 has the worst misclassification rate at 2.48%, and fold 2 has the best at 1.55%. In general, the model has been relatively consistent with misclassification rates in all folds, which are relatively low.

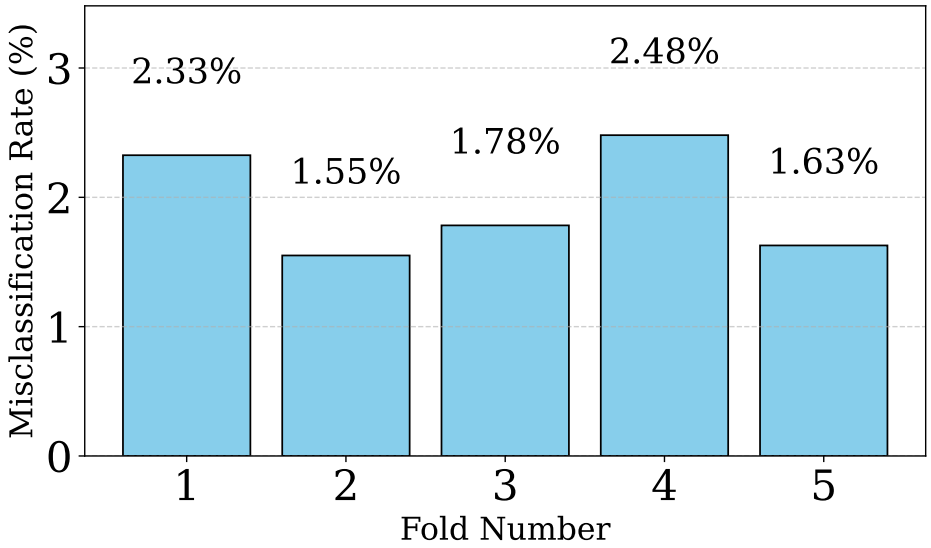


Figure 5. Misclassifications rate of each fold.

3.5. ROC Curve Analysis

Figure 6 presents the ROC curve of each fold using HBA_XGB classifier. The ROC curves of the HBA_XGB classifier on five folds have indicated high consistency and near-perfect performance. In fold 1 and fold 3, the model has shown a perfect AUC of 1.00 in all classes, implying perfect discrimination of classes. In fold 2 and fold 5, classes E and G have recorded an AUC of 1.00 in both folds and class P has the least deviation of 0.99. Likewise, fold 4 has also given an AUC score of 1.00 of each class. The curves in all folds have been lying very close to the top-left corner, confirming a strong balance between sensitivity and specificity. The model has indicated stable results with the AUC value between 0.99 and 1.00 since the beginning of the data split, indicating model robustness. The outcomes have

shown that the HBA_XGB classifier is a generalization and classifies separate target classes efficiently. On the whole, the ROC analysis has been demonstrating the effectiveness and validity of the suggested approach.

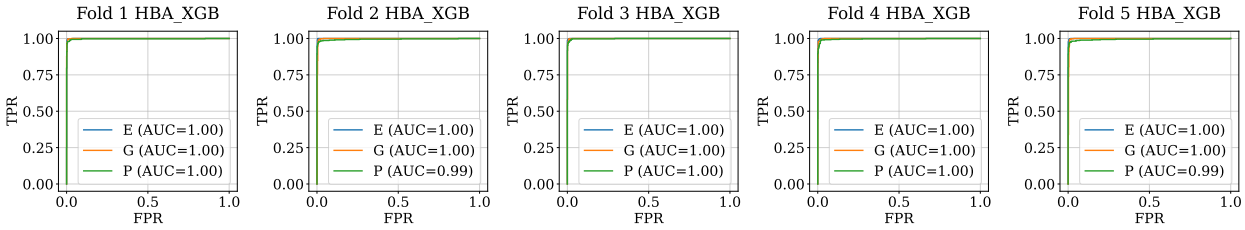


Figure 6. ROC curve visualization for each fold of HBA_XGB.

3.6. Controlling the Overfitting

Each fold of HBA_XGB achieved a training accuracy of 100%, and the test accuracy is slightly lower than the training accuracy. In this case, we have conducted an additional analysis to better understand and control overfitting using an XGBoost classifier on the fold 1 result, where only the maximum depth value has been changed, ranging from 1 to 10, to observe its effect on model generalization.

Table 7 shows that, at lower maximum depths (1–3), the training accuracy is high but not perfect, and the test accuracy remains close to the training accuracy, resulting in a small gap, indicating that the model is relatively simple and generalizes well. In contrast, at higher maximum depths (4–10), the training accuracy reaches 100% (perfect fit), but the test accuracy slightly decreases, leading to a slightly higher training-test gap. This behavior suggests a clear indication of a slight overfitting. When the maximum depth has increased, the training accuracy has reached 100%, while the test accuracy has decreased slightly, indicating a minor overfitting. The overall visualization of overfitting issues has been represented in Figure 7.

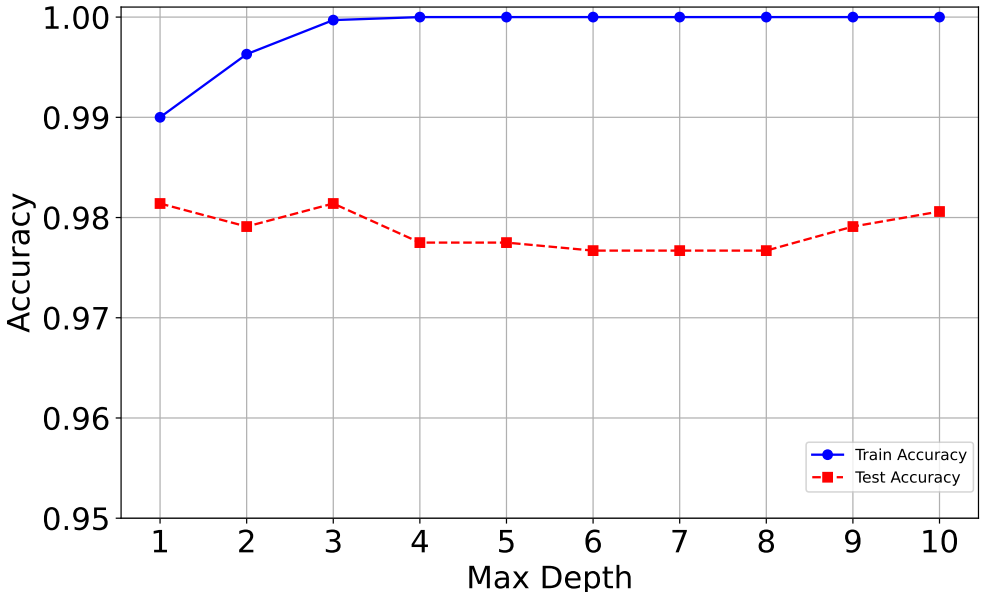


Figure 7. Training and test accuracy at different maximum depths.

Table 7. Fold-1 performances under different maximum depth settings (F1 = F-score; Pre. = Precision; Rec. = Recall; Gap = Train-Test Accuracy Gap).

Max Depth	Train				Test				Gap (%)
	Acc. (%)	F1 (%)	Pre. (%)	Rec. (%)	Acc. (%)	F1 (%)	Pre. (%)	Rec. (%)	
1	99.00	99.01	99.01	99.02	98.14	98.17	98.09	98.31	0.86
2	99.63	99.64	99.64	99.64	97.91	97.95	97.86	98.09	1.72
3	99.97	99.97	99.97	99.97	98.14	98.17	98.09	98.31	1.83
4	100.00	100.00	100.00	100.00	97.75	97.79	97.70	97.95	2.25
5	100.00	100.00	100.00	100.00	97.75	97.79	97.70	97.96	2.25
6	100.00	100.00	100.00	100.00	97.67	97.71	97.62	97.89	2.33
7	100.00	100.00	100.00	100.00	97.67	97.71	97.62	97.89	2.33
8	100.00	100.00	100.00	100.00	97.67	97.71	97.62	97.88	2.33
9	100.00	100.00	100.00	100.00	97.91	97.94	97.86	98.09	2.09
10	100.00	100.00	100.00	100.00	98.06	98.10	98.01	98.22	1.94

3.7. SHAP Analysis

In this section, we observe the mean SHAP values result for excellent class of each fold, which indicates the feature importance. Each fold's mean SHAP values have been represented in Figure 8. The findings have shown that the parameters that have contributed the most have varied a little across folds but some common patterns have been identified.

Ammonia (mg/L) has been established as the most important in fold 1, 3, and 4, with a close consideration to Dissolved Oxygen (DO), Turbidity, and BOD. On the other hand, Nitrite (mg/L) has prevailed as the most predictive factor in fold 2 and fold 5. Regardless of these variations, the general ranking has revealed that Ammonia, Nitrite, DO, Turbidity, and BOD have been the five most influential parameters across the folds, which confirms their essential importance in model predictions.

Conversely, attributes such as Temperature, pH, and CO_2 have always recorded low mean SHAP values, indicating that their contribution to the classification performance has been smaller, as compared to the ones mentioned above. In addition, the analysis has been continuously pointing out that Ammonia, Nitrite, DO, Turbidity, BOD, Temperature, pH, and CO_2 features are the most important WQ indicators for fish farming.

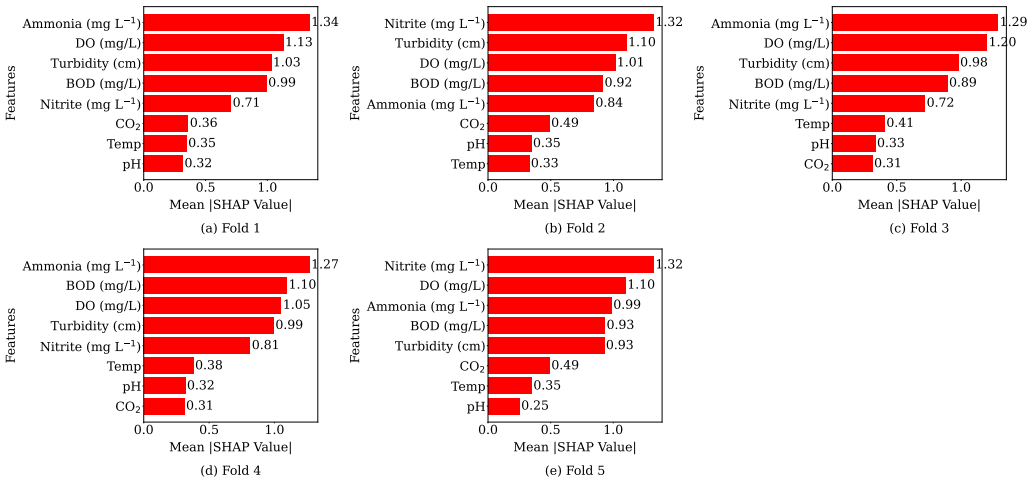


Figure 8. Each fold mean SHAP value representation for excellent class.

3.8. DiCE Analysis

Table 8 presents the counterfactual explanations for misclassified instances, and Table 9 represents correctly classified instances. In the counterfactual analysis, Tables 8 and 9, the Index refers to the sample identification number from the dataset. The term Actual refers to the true class label of the sample. Predicted indicates the class label assigned by the trained model. We have made a comparison between the misclassified and correctly classified cases to determine the effect of features on model predictions. In the case of misclassified samples (i.e., index 57 actual 2, predicted 1), (index 299 actual 1, predicted 2), and (index 927 actual 0, predicted 2), counterfactual explanations demonstrate that the differences in Temp (F1), Turbidity (F2), DO (F3), BOD (F4), CO_2 (F5), pH (F6), Ammonia (F10), and Nitrite (F11) are influential to correct the predictions. In the correctly labeled cases (indices 0, 1 and 3), we have identified which features we would have to alter to cause the prediction to be in the other class. As an example, index 0 (actual 1, predicted 1) involves changes in BOD (F4), CO_2 (F5), and Ammonia (F10); index 1 (actual 0, predicted 0) depends on Temp (F1), DO (F3), BOD (F4), CO_2 (F5), pH (F6), Ammonia (F10), and Nitrite (F11).

In general, the model has been continuing to provide high accuracy scores on correctly classified cases, and counterfactual analysis has been helpful in understanding what water quality features affect prediction and where these cases can be improved.

Table 8. Counterfactual explanations for misclassified instances.

Index	Class (Actual→Pred)	Feature	Original	CF1	CF2	CF3
57	2→1	F2	17.85	89.51	N/A	N/A
57	2→1	F4	3.11	13.60	14.09	N/A
57	2→1	F6	9.38	N/A	N/A	13.95
299	1→2	F4	2.39	4.37	N/A	N/A
299	1→2	F5	8.31	N/A	3.16	N/A
299	1→2	F6	6.46	N/A	5.92	N/A
927	0→2	F3	3.00	4.88	3.78	3.57
927	0→2	F5	6.63	4.79	N/A	3.48

Notes: N/A indicates that the feature value did not change in that counterfactual.

Table 9. Counterfactual explanations for correctly classified instances.

Index	Class (Actual→Pred→Desired)	Feature	Original	CF1	CF2	CF3
0	1→1→2	F4	2.47	8.76	N/A	N/A
0	1→1→2	F5	4.32	N/A	10.42	N/A
0	1→1→2	F1	30.33	N/A	N/A	9.73
1	0→0→2	F3	4.39	12.66	N/A	N/A
1	0→0→2	F5	6.68	N/A	N/A	9.62
1	0→0→2	F6	6.75	12.54	N/A	N/A
1	0→0→2	F10	0.002	N/A	0.60	N/A
3	2→2→1	F1	28.30	32.69	32.24	30.60
3	2→2→1	F3	4.08	5.84	5.76	5.47
3	2→2→1	F5	3.09	N/A	N/A	5.39
3	2→2→1	F6	6.71	5.72	5.64	5.35
3	2→2→1	F10	0.02	0.38	0.37	0.35
3	2→2→1	F11	4.92	N/A	N/A	1.79

Note: N/A indicates that the feature value did not change in that counterfactual.

In this study, we applied SHAP for understanding the feature importance of the WQ dataset. The mean SHAP values convey information about feature importance, whereas

DiCE provides what changes in the input are needed to alter the prediction. In this case, each fold has been shown that Ammonia, Nitrite, DO, Turbidity, BOD, Temperature, pH, and CO_2 are the most important feature according to mean SHAP values. In addition, Ammonia, Nitrite, DO, Turbidity, BOD, Temperature, pH, and CO_2 are the most influential to alter the prediction outcomes. Finally, we can conclude that Ammonia, Nitrite, DO, Turbidity, BOD, Temperature, pH, and CO_2 are the key drivers of the model, both most important for prediction and most sensitive to changes.

4. Discussion

In this study, our primary research question was to determine how effectively an XGBoost classifier optimized by the Honey Badger Algorithm can classify aquaculture water quality and identify key influencing parameters using SHAP and DiCE explanations. According to our research question, HBA_XGB classified WQ, achieving a highest accuracy of 98.45%, which shows how effectively it can classify WQ. The SHAP analysis identified that Ammonia, Nitrite, DO, Turbidity, BOD, Temperature, pH, and CO_2 are the most influential in driving classification decisions. DiCE-based counterfactuals have explained how variations in these parameters could impact predictions.

Table 10 presents a comparison of classification outcomes for various WQ-related datasets alongside the results from our proposed approach. Though Khan et al. [30] and Uddin et al. [26] have obtained perfect accuracy through gradient boosting methodologies, our model has shown comparable performance with maximum fold accuracy of 98.45% and average accuracy of 98.05% through 5-fold cross validation. When compared to other gradient boosting-based models, e.g., Mahmoud et al. [21] and Nur et al. [23], HBA_XGB has lower peak accuracy but higher fold-to-fold consistency. Ho et al.'s decision tree model [25] has achieved very low accuracy (81%), which is significant in establishing the dominance of ensemble methodologies for classification in WQ. Overall, these results show the robustness and reliability of HBA_XGB WQ classification. In addition, we have reduced the feature of our selected dataset. We have also analyzed SHAP results to observe the impact of WQ parameters. However, none of the previous studies are directly comparable with our work because the dataset is different. In this work, we used the aquaculture water quality dataset, which is a recently published dataset. In the SHAP and DiCE analyses, Ammonia, Nitrite, DO, Turbidity, BOD, Temperature, pH, and CO_2 were identified as the most important and influential water quality parameters. We conclude that deviations from the optimal range of these parameters can adversely affect aquaculture species.

In future research, our precise models can be successfully applied in aquaculture, as well as in overseeing water quality because our work has achieved a promising accuracy. Reducing the number of water quality (WQ) features offers several practical benefits in real-world applications. Fewer WQ features imply a reduced number of parameter measurements, which can significantly lower the cost and effort required for water quality monitoring. However, various metaheuristic optimization methods and classifiers might be incorporated to improve the prediction accuracy.

The dataset is well balanced across classes, with 1500 Poor-quality, 1400 Excellent-quality, and 1400 Good-quality water samples, ensuring that class representation does not bias the model. All samples of this dataset are collected from specific regions. In this context, our work has some limitations. First, the HBA_XGB model has been applied to a specific dataset, which may limit its generalizability to other geographic regions or water bodies with different characteristics. In future work, we plan to integrate multi-region datasets to improve the generalizability of the HBA_XGB model across diverse aquaculture environments. Second, the model was trained using only selected features, and some

important chemical and biological water parameters have been omitted, which could further affect its generalizability.

Table 10. Comparison of our work with previous studies.

Author	Dataset Information	Classification Model	Accuracy (%)
Mahmoud et al. [21]	Indian Water Quality Dataset	GB	99.50
Nasir et al. [22]	Indian Water Quality Dataset	CB	94.51
Nur et al. [23]	Department of Environment, Malaysia	GB	94.90
Khan et al. [24]	Gulshan Lake Water Quality Dataset	GB	100
Ho et al. [25]	Klang River Water Quality Dataset	DT	81
Uddin et al. [26]	Lee, Cork Harbour, and Youghal Bay	XGB	100
Our method-1	Aquaculture Water Quality Dataset	HBA_XGB (Fold 2)	98.45 (highest accuracy among 5 folds)
Our method-2	Aquaculture Water Quality Dataset	HBA_XGB (5-fold)	98.05 (mean accuracy of custom 5 folds)

5. Conclusions

This work has classified aquaculture water quality into Excellent, Good, and Poor classes using a robust machine learning model. A metaheuristic-based HBA approach has been employed for relatively efficient feature optimization and selection. Additionally, HBA has been integrated with the powerful XGB classifier to enhance the prediction performance. The dataset has randomly shuffled, and a custom 5-fold cross-validation strategy has been implemented, with the HBA_XGB model. Among the folds, fold 2 has achieved the highest accuracy of 98.45%, while the mean accuracy across all folds are 98.05%, demonstrating the model's consistency and reliability. Through the comprehensive SHAP analysis, the model identified Ammonia, Nitrite, DO, Turbidity, BOD, Temperature, pH, and CO_2 as the most influential and critical WQ indicators for sustainable fish farming. Finally, the DiCE analysis identified Temperature, Turbidity, DO, BOD, CO_2, pH, Ammonia, and Nitrite as the most influential parameters that affect the water quality. In this study, Ammonia, Nitrite, DO, Turbidity, BOD, Temperature, pH, and CO_2 have been identified as the key water-quality parameters based on the SHAP and DiCE contexts, and straying from the ideal range of these parameters can negatively impact aquaculture species.

Author Contributions: Conceptualization, S.M.N., P.D., J.-J.T., and A.-A.N.; data curation, S.M.N. and P.D.; formal analysis, S.M.N., P.D., and A.-A.N.; investigation, S.M.N., P.D., and A.-A.N.; methodology, S.M.N., P.D., J.-J.T., and A.-A.N.; resources, S.M.N., P.D., J.-J.T., and A.-A.N.; software, S.M.N. and P.D.; supervision, J.-J.T. and A.-A.N.; validation, S.M.N., P.D., J.-J.T., and A.-A.N.; visualization, S.M.N., P.D., J.-J.T., and A.-A.N.; writing—original draft, S.M.N. and P.D.; writing—review and editing, J.-J.T. and A.-A.N. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: Dataset is available on Mendeley Data Repository. Dataset Name: Aquaculture—Water Quality Dataset. URL: https://data.mendeley.com/datasets/y78ty2g293/1, accessed on 20 August 2025.

Acknowledgments: We have used some of the large Language Models such as Chat GPT-5, and DeepSeek to enhance the structure of sentences.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Eruola, A.O.; Ufoegbune, G.C.; Awomeso, J.A.; Abhulimen, S.A. Assessment of cadmium, lead and iron in hand dug wells of Ilaro and Aiyetoro, Ogun State, South-Western Nigeria. *Res. J. Chem. Sci.* **2011**, *2231*, 606X.
2. Matta, G.; Kumar, P.; Uniyal, D.P.; Joshi, D.U. Communicating water, sanitation, and hygiene under sustainable development goals 3, 4, and 6 as the panacea for epidemics and pandemics referencing the succession of COVID-19 surges. *ACS ES&T Water* **2022**, *2*, 667–689.
3. FAO. *The State of World Fisheries and Aquaculture 2018: Meeting the Sustainable Development Goals*; Food and Agriculture Organization of the United Nations: Rome, Italy, 2018.
4. MAAIF. *Essentials of Aquaculture Production, Management and Development in Uganda*; Ministry of Agriculture, Animal Industry and Fisheries (MAAIF): Entebbe, Uganda, 2018.
5. Abd El-Hamed, N. Environmental studies of water quality and its effect on fish of some farms in Sharkia and Kafr. *El-Sheikh Governorates*, 2014.
6. Cline, D. *Water Quality in Aquaculture*; Alabama Cooperative Extension System, Auburn University: Auburn, AL, USA, 2019. Available online: https://freshwater-aquaculture.extension.org/water-quality-in-aquaculture/ (accessed on 26 August 2019).
7. Palma, J.; Correia, M.; Leitão, F.; Andrade, J.P. Temperature effects on growth performance, fecundity and survival of *Hippocampus guttulatus*. *Diversity* **2024**, *16*, 719.
8. Devi, P.A.; Padmavathy, P.; Aanand, S.; Aruljothi, K. Review on water quality parameters in freshwater cage fish culture. *Int. J. Appl. Res.* **2017**, *3*, 114–120.
9. Boyd, C.E. *Water Quality Management for Pond Fish Culture*; Elsevier Scientific Publishing Co.: Amsterdam, The Netherlands, 1982.
10. Bolorunduro, P.I.; Abdullah, A.Y. Water quality management in fish culture, national agricultural extension and research liaison services, Zaria. *Ext. Bull.* **1996**, *98*.
11. Siti-Zahrah, A.; Misri, S.; Padilah, B.; Zulkafli, R.; Kua, B.C.; Azila, A.; Rimatulhana, R. Pre-disposing factors associated with outbreak of Streptococcal infection in floating cage-cultured red tilapia in reservoirs. In Proceedings of the 7th Asian Fisheries Forum, Penang, Malaysia, 1–2 December 2004; Volume 4, pp. 129.
12. Siti-Zahrah, A.; Padilah, B.; Azila, A.; Rimatulhana, R.; Shahidan, H. Multiple streptococcal species infection in cage-cultured red tilapia but showing similar clinical signs. In *Diseases in Asian Aquaculture VI*; Fish Health Section, Asian Fisheries Society: Manila, Philippines, 2008; pp. 313–320.
13. Hepher, B.; Pruginin, Y. *Commercial Fish Farming. A Wiley-Interscience Publication*; John Wiley and Sons: New York, NY, USA, 1981.
14. Nsonga, A. Indigenous fish species a panacea for cage aquaculture in Zambia: A case for *Oreochromis macrochir* at Kambashi out grower scheme. *Int. J. Fish. Aquat. Stud.* 2014, *2*, 102–105.
15. Daniel, S.; Larry, W.D.; Joseph, H.S. Comparative oxygen consumption and metabolism of striped bass (*Morone saxatilis*) and its hybrid. *J. World Aquac. Soc.* **2005**, *36*, 521–529.
16. Praveen, T.; Mishra, R. Somdutt. Water quality monitoring of Halali reservoir with reference to Cage aquaculture as a modern tool for obtaining enhanced fish production. *Proc. Taal* **2007**, 318–324.
17. Razali, R.M. Predictive Water Quality Monitoring in Aquaculture Using Machine Learning and IoT Automation. *Adv. Comput. Intell. Syst.* **2025**, *1*, 10–17.
18. Nayan, A.A.; Kibria, M.G.; Rahman, M.O.; Saha, J. River water quality analysis and prediction using GBM. In Proceedings of the 2020 2nd International Conference on Advanced Information and Communication Technology (ICAICT), Dhaka, Bangladesh, 28–29 November 2020; IEEE: Piscataway, NJ, USA, 2020; pp. 219–224.
19. Nayan, A.A.; Mozumder, A.N.; Saha, J.; Mahmud, K.R.; Al Azad, A.K. Early detection of fish diseases by analyzing water quality using machine learning algorithm. *Walailak J. Sci. Technol.* **2020**, *18*.
20. Sen, S.; Maiti, S.; Manna, S.; Roy, B.; Ghosh, A. Smart Prediction of Water Quality System for Aquaculture using Machine Learning Algorithms. *TechRxiv* **2023**. https://doi.org/10.36227/techrxiv.22300435.v1.
21. Shams, M.Y.; Elshewey, A.M.; El-Kenawy, E.-S.M.; Ibrahim, A.; Talaat, F.M.; Tarek, Z. Water quality prediction using machine learning models based on grid search method. *Multimed. Tools Appl.* **2024**, *83*, 35307–35334.
22. Nasir, N.; Kansal, A.; Alshaltone, O.; Barneih, F.; Sameer, M.; Shanableh, A.; Al-Shamma'a, A. Water quality classification using machine learning algorithms. *J. Water Process Eng.* **2022**, *48*, 102920.
23. Malek, N.H.A.; Yaacob, W.F.W.; Nasir, S.A.M.; Shaadan, N. Prediction of water quality classification of the Kelantan River Basin, Malaysia, using machine learning techniques. *Water* **2022**, *14*, 1067.
24. Khan, M.S.I.; Islam, N.; Uddin, J.; Islam, S.; Nasir, M.K. Water quality prediction and classification based on principal component regression and gradient boosting classifier approach. *J. King Saud Univ.—Comput. Inf. Sci.* **2022**, *34*, 4773–4781.
25. Ho, J.Y.; Afan, H.A.; El-Shafie, A.H.; Koting, S.B.; Mohd, N.S.; Jaafar, W.Z.; Sai, H.L.; Malek, M.A.; Ahmed, A.N.; Mohtar, W.H.M.; et al. Towards a time and cost effective approach to water quality index class prediction. *J. Hydrol.* **2019**, *575*, 148–165.

26. Uddin, M.G.; Nash, S.; Rahman, A.; Olbert, A.I. Performance analysis of the water quality index model for predicting water state using machine learning techniques. *Process Saf. Environ. Prot.* **2023**, 169, 808–828.
27. Reddy, A.P.; Sophia, P.E.; Kirubakaran, S.S. Automated Cardiovascular Disease Diagnosis using Honey Badger Optimization with Modified Deep Learning Model. *Biomed. Mater. Devices* **2025**, 1–8. https://doi.org/10.1007/s44174-025-00495-y.
28. Li, W.; Deng, M.; Liu, C.; Cao, Q. Analysis of Key Influencing Factors of Water Quality in Tai Lake Basin Based on XGBoost-SHAP. *Water* **2025**, *17*, 1619.
29. Oprea, S.-V.; Bâra, A. Diverse Counterfactual Explanations (DiCE) Role in Improving Sales and e-Commerce Strategies. *J. Theor. Appl. Electron. Commer. Res.* **2025**, *20*, 96.
30. Dong, W.; Huang, Y.; Lehane, B.; Ma, G. XGBoost algorithm-based prediction of concrete electrical resistivity for structural health monitoring. *Autom. Constr.* **2020**, *114*, 103155.
31. Chen, T.; Guestrin, C. XGBoost: A scalable tree boosting system. In Proceedings of the KDD'16: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, 13–17 August 2016; pp. 785–794.
32. Abualdenien, J.; Borrmann, A. Ensemble-learning approach for the classification of Levels Of Geometry (LOG) of building elements. *Adv. Eng. Inform.* **2022**, *51*, 101497.
33. Hashim, F.A.; Houssein, E.H.; Hussain, K.; Mabrouk, M.S.; Al-Atabany, W. Honey Badger Algorithm: New metaheuristic algorithm for solving optimization problems. *Math. Comput. Simul.* **2022**, *192*, 84–110.
34. Lundberg, S.M.; Lee, S. A unified approach to interpreting model predictions. *Adv. Neural Inf. Process. Syst.* **2017**, 30.https://doi.org/10.48550/arXiv.1705.07874.
35. Mothilal, R.K.; Sharma, A.; Tan, C. Explaining machine learning classifiers through diverse counterfactual explanations. In Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency, Barcelona, Spain, 27–30 January 2020; pp. 607–617.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.