Advances in Nonlinear Variational Inequalities
ISSN: 1092-910X
Vol 27 No. 3 (2024)

Designing and Developing novel methods for Enhancing the Accuracy of Water Quality Prediction for Aquaponic Farming

[1*]Govinda Khandelwal, [2]Bhushankumar Nemade, [3]Namdeo Badhe, [4]Darshan Mali, [5]Krishna Gaikwad, [6]Namrata Ansari

[1,6]K. J. Somaiya Institute of Engineering and Information Technology, Mumbai, India
[2]Shree L R Tiwari College of Engineering, Mumbai, India.
[3,4,5] Thakur College of Engineering and Technology, Mumbai, India
[1]govindak8177@gmail.com, [2]bnemade@gmail.com, [3]namdeobadhe1982@gmail.com, [4]id.darshanmali@gmail.com, [5]Krishna.gaikwad@thakureducation.org, [6]ngharat@somaiya.edu

Article History:

Received: *16-02-2024*

Revised: *26-04-2024*

Accepted: *18-05-2024*

Abstract

The Internet of Things (IoT)-based automated water surveillance system is crucial for fish farming, as it helps control risks and improve output and productivity. However, the industry faces challenges such as a lack of understanding of organism selection based on water purity indicators, a shortage of premium seeds and species, and imbalanced datasets. Existing systems also lack optimal feature selection methods, as well as modern machine learning and deep learning approaches. To address these issues, a novel system is proposed that integrates an aquaponic ecosystem containing fish, plants, and bacteria to balance water quality parameters and boost productivity. The system collects data from IoT devices, performs a data cleaning process using missing values, and outliers handling. Then, it performs a feature extraction process to select the optimum features. Next, it employs the novel H-SMOTE approach to tackle the issue of imbalanced datasets. The system uses multi-model categorization to cultivate fish in cold waters, warm-water aquatic plants, and bacteria in an aquaponic environment. The system uses a voting principle to identify the most effective prediction model. The proposed system achieves 99.50% accuracy for water quality prediction for aquaponic farming, addressing the limitations of existing systems and improving prediction accuracy and overall aquaponic farming productivity.

Keywords: Water Quality Prediction, Aquaponic Farming.

1. Introduction

Aquaculture, or the cultivation of aquatic organisms, has emerged as a critical alternative for combating global food scarcity, improving food security, and driving economic growth. Over the last three decades, fish farming has become widely used, considerably increasing fish production and stability [1]. Notably, China, a key contributor to global aquaculture output, demonstrates the critical role aquaculture plays in addressing food emergencies and preserving livelihoods [2,3]. However, the increasing aquaculture business faces a wide range of issues, including resource depletion and environmental damage [4]. High breeding density and enhanced yields of intensive industrial aquaculture operations can lead to incorrect feeding procedures and degradation of water quality [5,6]. The delicate balance of aquatic ecosystems emphasizes the importance of maintaining appropriate water quality parameters to protect aquatic animals' health and production [7]. To solve these difficulties, novel solutions are required to reduce environmental impact and improve operating efficiency. One significant challenge is water quality management, which is critical to the health and growth of cultured species [8,9]. Traditional water quality management techniques are frequently reactive and labor-intensive, relying on repeated sampling and analysis that may fail to capture the dynamic nature of aquatic habitats [10,11]. To deal with this problem we would move toward machine Learning.

A Machine learning is a kind of intelligent technology that offers predictive skills with data-driven insights. ML algorithms can process massive volumes of data from several sensors in real time, detecting patterns and abnormalities that human operators may miss [12,13]. Farmers who integrate machine learning into aquaculture can shift from reactive to proactive management, detecting possible hazards before they become serious [14]. For example, ML can predict oxygen depletion events, allowing for proactive aeration to maintain ideal levels for fish survival [15]. Similarly, ML may improve feeding schedules based on the stock's expected growth rates and health indicators, decreasing waste and improving feed conversion rates [16,17]. Furthermore, ML models can aid in disease detection and prevention by analyzing behavioral and physiological data for early signs of stress or illness [18,19]. The use of ML in aquaculture is consistent with the broader goals of precision farming, which strives to increase productivity while minimizing environmental impact [16-19]. Aquaculture operations that use machine learning for water quality prediction can improve resource usage and animal welfare, as well as contribute to the worldwide endeavor of sustainable food production [20].

To address the complexity of water quality management in aquaculture, researchers have tried a variety of approaches, including traditional prediction methods and machine learning (ML) models [20,21]. Traditional approaches are simple to adopt, but they fail to capture the nonlinear and dynamic nature of water quality data [22]. In contrast, machine learning methods show promise for precise and efficient water quality prediction. These models include a variety of techniques [23].

We're using support vector Machine, decision tree, and random forest algorithms. Studies have demonstrated that ML models can reliably predict crucial water quality parameters for aquaculture operations, such as pH, dissolved oxygen, turbidity, and salinity [21-23]. Notably, ML algorithms such as LSTM and ARIMA have been particularly effective for real-time water quality boundary prediction, allowing aqua farmers to take preemptive actions [24]. The implementation of machine learning (ML) in aquaculture water quality management represents a substantial departure from traditional prediction methodologies [25]. Traditional techniques, while simpler to execute, frequently fail to account for the intricate and changing patterns found in aquatic environments [26,27]. In contrast, ML models flourish in such a complicated environment [28]. They excel at deciphering the nonlinear correlations and temporal dynamics that characterize water quality data, resulting in a more nuanced knowledge and control over the aquaculture environment [29]. The robustness and flexibility of support vector models (the SVM), decision trees (the DT), and random forest algorithms (RF) distinguish them from other machine learning (ML) systems [30,31]. These models have contributed to improving the precision and accuracy of water quality forecasts while also adjusting to the specific needs of diverse aquaculture systems [32]. We've made tremendous progress with LSTM and ARIMA. Empirical research has shown that machine learning models can predict crucial quality of water indicators such as alkalinity, oxygen dispersion, turbidity, as well as salinity—parameters required for aquatic species survival and growth [33,34].

These models include ARIMA (Autoregressive Integrated Moving Average) and LSTM (Long Short- Term Memory) systems [35]. These networks, which are a form of neural network with recurrent connections, are particularly excellent at capturing long-term relationships, making them ideal for aquaculture time- series prediction applications [36,37]. ARIMA models, known for their ability to anticipate time series, improve LSTMs by giving accurate short-term projections [38]. The combination of LSTM and ARIMA models is a powerful tool for real-time water quality boundary prediction [39]. This competence enables aquaculture producers to forecast and respond swiftly to possible changes in water conditions, maintaining the sustainability and productivity of their aquaculture operations [40,41]. Farmers may employ powerful machine learning algorithms to optimize resource allocation, increase fish welfare, and ultimately contribute to the global goal of sustainable food production [42].

Aquaculture is vital to supplying the increasing demand for aquatic goods, and fish farming is an important part of the industry [43]. Aquaculture includes a wide range of marine creatures farmed in various aquatic settings, such as food fish and ornamental species [44]. However, maintaining suitable water quality requirements is critical to the success of aquaculture operations since it ensures the health and production of aquatic species [45]. Failure to maintain proper water conditions may have a severe influence on fish health and production, reducing the overall viability of aquaculture operations. Industrial aquaculture activities, although important to fulfill global demand, provide significant challenges in terms of water quality management [42-45]. High breeding rates and intensive agricultural techniques may lead to nutrient buildup and water quality deterioration, necessitating adequate monitoring and control strategies. Traditionally, aqua farmers used manual evaluation techniques, which are time-consuming and prone to mistakes [46]. However, using contemporary technology, particularly machine learning (ML) models, provides a realistic option for enhancing water quality monitoring and prediction in aquaculture. support vector algorithms (SVM) are two machine learning technologies that have shown great potential in accurately forecasting water quality indicators [45,46].

These models use previous data to estimate future patterns, allowing aquaculture farmers to make more educated decisions and proactively manage risks [47]. Furthermore, advances in machine learning technology,

such as LSTM and ARIMA models, offer real-time prediction of water quality dynamics, allowing for prompt action in critical situations [48].

The implementation of machine learning (ML) algorithms in aquaculture has considerable potential for enhancing productivity and assuring the sector's sustainability [49]. Machine learning algorithms allow for proactive management of aquaculture operations by providing aquafarmers with accurate and timely information on water quality indicators, reducing the danger of disease outbreaks and environmental harm [50]. Aquafarmers may also improve resource allocation and overall operational efficiency by using ML- based systems, which are a low-cost and effective alternative to traditional monitoring approaches [46-50]. Despite their potential benefits, adopting and deploying machine learning models in aquaculture presents additional obstacles [51]. The scarcity of data, especially on small-scale industrial farms, creates a considerable hurdle to constructing accurate prediction models [52]. Furthermore, the intricate nature of aquatic ecosystems and the way that many environmental factors interact make it necessary to use strong modelling tools that can capture relationships that are not linear and change over time [53]. Addressing these difficulties would necessitate cross-disciplinary collaboration as well as novel data collection and analysis methods [54,55]. Researchers can improve the accuracy and reliability of ML models for aquaculture water quality prediction by combining data from diverse sources, such as sensors, satellites, and environmental monitoring stations [56]. Furthermore, ongoing research efforts focused on enhancing modeling methodologies and exploring innovative algorithms hold potential for overcoming current limitations and achieving real- world outcomes. Aquaculture can fully utilize the potential of machine learning [57,58].

Given the growing importance of aquaculture in global food security and economic development, ongoing research and innovation in water quality management are required [59]. Drawing on current research publications, this study intends to contribute to the continuing discussion about water quality prediction in aquaculture and provide ways for increasing productivity and sustainability in the sector [60,61]. We a thorough assessment and analysis of the available literature, we hope to identify important issues and possibilities in water quality management and investigate the potential of machine learning (ML) models in tackling these challenges [59-61]. In the future, increasing collaboration among researchers, industry stakeholders, and policymakers will be critical to developing effective solutions for water quality management in aquaculture [62,63].

Leveraging advances in machine learning technology and incorporating data-driven techniques into aquaculture practices can improve the resilience and sustainability of aquaculture operations while also ensuring the continuing supply of high-quality aquatic products for global consumers [64,65]. Together, we can use innovation to create a more resilient and sustainable future for aquaculture. We cannot over emphasize the importance of water quality control in aquaculture [66]. As the aquaculture industry expands to meet the growing demand for aquatic products, maintaining appropriate water conditions becomes critical to the health and production of aquatic species [67,68].

Machine learning (ML) models present a possible option for improving water quality prediction and management, allowing aquaculture farmers to make more informed decisions and effectively manage risks [65-68]. Drawing on existing research publications, this work adds to the continuing discussion about water quality management in aquaculture [69]. It presents solutions to increase productivity and sustainability in the sector [70]. Through interdisciplinary collaboration and new ideas, we can create a more resilient and sustainable aquaculture future, ensuring the continuous supply of high-quality aquatic goods for global consumers [71-72].

2. Literature Survey & Research Gaps

Bhushankumar Nemade and **Deven Shah [73]** developed an IoT-based water quality prediction system for aquaponics farming, using IoT sensors for data collection and the novel M-SMOTE algorithm to address class imbalance. The system achieved 99.13% accuracy with XGBoost and Random Forest classifiers. This system is crucial for risk management and productivity in aquaponics, with significant environmental benefits.

Chiang Liang Kok, **I Made Bagus Pradnya Kusuma, Yit Yan Koh, Howard Tang, and Ah Boon Lim [74]** developed an automated water quality management system for aquaponics using PID calculations. The system controlled a BLDC motor, aeration pump, and heater with PWM signals. Independent testing confirmed functionality, with PID settings ensuring stable water levels in vegetable and fish tanks. Future work could integrate machine learning for dynamic adjustments and compare control strategies for efficiency. Long-term studies are needed to assess the system's sustainability and productivity in urban aquaponics.

Andres Felipe Zambrano, Luis Felipe Giraldo, Julian Quimbayo, Brayan Medina, and **Eduardo Castillo [75]** used machine learning to predict water quality in fish farming. The study utilized a random forest model to accurately estimate ammonium and ammonia levels based on pH and temperature, comparing its performance with linear regression and neural networks. The research highlights the potential to reduce comprehensive measurements by accurately forecasting key water quality variables, even with limited data. It emphasizes the importance of optimizing data collection frequency and suggests a need for developing new machine learning techniques or hybrid models for sparse data. The study also points out the necessity for accessible tools for fish farmers, especially in resource-limited settings, and the need for models capable of long-term forecasting.

Aldhyani et al. [76] conducted a study utilizing artificial intelligence to predict water quality indices (WQI), splitting the dataset into 70% for training and 30% for testing. They evaluated several algorithms, including NARNET, LSTM, SVM, KNN, and Naive Bayes. The NARNET model with 12 hidden layers and the LSTM model with 200 hidden layers achieved high accuracy, with an R-value above 93.93%. The SVM algorithm excelled in water quality classification. The study emphasizes the potential of AI in environmental monitoring, highlighting the importance of predictive modeling in water quality management. However, the research's focus on a single region limits its broader applicability. This summary covers the study's key aspects and acknowledges its limitations.

A. Najah et al. **[77]** ANN models, including MLP-NN and RBF-NN, effectively predicted water quality in the Johor River basin, with RBF-NN showing superior error distribution. RBF-NN models outperformed LRMs and MLP-NN in speed and efficiency. The study recommends investigating advanced ANNs for better accuracy. There's potential for more dynamic ANN models and optimization techniques to further improve water quality predictions.

Avila et al. [78] The study evaluated statistical models for water quality prediction in the Oreti River, focusing on E. coli levels and rainfall impact on water modes. E. coli concentrations showed variability, with mostly acceptable levels for recreation. Rainfall influenced water modes, with more Amber and Red modes during higher rainfall. Bayesian networks excelled in predicting red mode days, with low error rates. Bayesian networks are optimal for real-time water quality prediction, with potential for future enhancements and policy-making applications. While the Bayesian network model shows promise for real-time water quality prediction in the Oreti River, further research is needed to assess its effectiveness in different environmental conditions and geographic locations. Additionally, integrating spatial components into the model could enhance its predictive accuracy, offering a more comprehensive tool for environmental management and policy-making.

3. Proposed Methodology

Farmers' lack of skill and understanding presents hurdles to aquaponic farming. Adopting current technology is critical for growing the aquaponics sector. This includes monitoring water quality, selecting disease-free seeds and species, and conducting effective breeding operations. Our suggested method is designed to predict suitable cold- and warm-water fish, plants, and beneficial microbes for freshwater fish farming. Furthermore, it offers tactics and treatments for the efficient cultivation of fish while taking into account additional aquatic creatures which include plants and microbes maintain the health of the ecosystem. The system has six components: Real-time data collection: Sensor gadgets collect crucial data. Data cleaning comprises removing missing data and outliers. Feature extraction is the process of identifying relevant qualities from sensory input. unbalanced Data Handling: The H-SMOTE approach is designed to handle imbalanced datasets. Preferred Classifiers: We use some models to categorize. Optimal Prediction: A voting-based method guarantees accurate projections. Aquaponic farmers may enhance their operations by using technology and prediction tools.

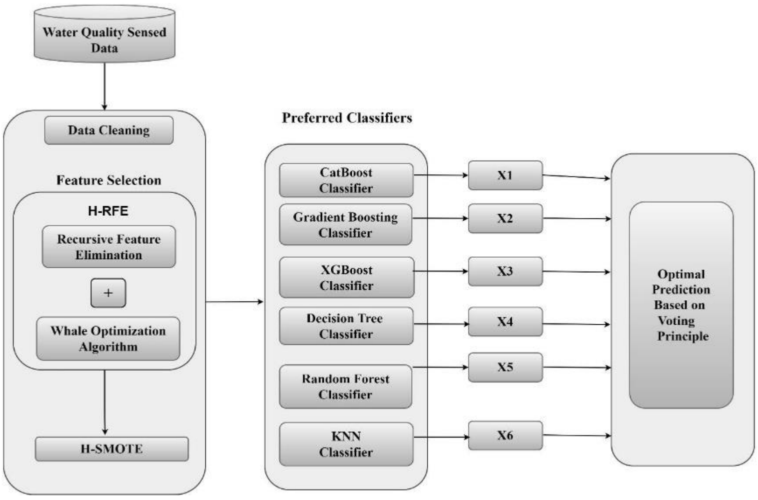

Fig. 1 Water Quality Prediction Framework for Aquaponic Farming.

3.1 Real-Time Data Collection

The water's pH indicators were initially obtained using devices with Internet of Things sensors in different places. The collected data is statistically stated as follows:

Ss= {S1, S2, S3, …Sm}.

where Ss represents the dataset and Sn denotes the quantity of data sensor S.

3.2 Data Cleaning Process for Handling Missing Values and Outliers

In our suggested system, we use data cleaning to improve the quality of our dataset. We specifically tackle two typical issues: missing numbers and outliers. When encountering missing values, we take a basic technique. For instance, if the twentieth value is absent, we calculate the average of the two values before it (18th and 19th) and following it (21st and 22nd). This average replaces the missing 20th value. Correcting outliers can improve dataset integrity. To overcome this, we use the median method to adjust records with outliers. In circumstances when more than one number is absent, we compute the mean of the two nearest values. By using these data cleaning approaches, we ensure that our dataset is reliable and accurate for further analysis and modelling.

3.3 Feature Selection

At this step, we employ a correlation matrix and a heatmap to identify essential traits. We priorities variables with the greatest effect on the prediction variable.

3.4 Imbalanced data handling.

In our suggested approach, we divide the dataset into nine unique categories. Initially, we see a non-uniform distribution of these classes. For a classification issue with 7,338 occurrences (rows), the class ratios (given as percentages) are as follows: NS:P:C: B:BW: BP: BPW: BC: BCCP = 6:4:6:7:17:13:25:10:13 the distribution of classes is given in Table1.

The dataset has Class imbalance hampers multiclass categorization. To overcome this problem, we recommend employing a number of performance metrics like as accuracy, recall, F1-Score, Kappa, and ROC curves. Skewed datasets should not be evaluated just on the basis of accuracy.

To address the imbalance, consider the following approaches:

Over-sampling is the process of randomly adding occurrences (rows) within the minority cohort into the training dataset, using replacement.

Under-sampling is a manner of removing occurrences (rows) of a class with too many copies. Optimization Challenges:

Selecting examples from the minority class, selecting sample features, sample kinds, sampling rates, and developing a technique for distributing minority class samples are all part of a complicated optimization issue.

By carefully resolving class imbalance, we improve the resilience and reliability of our classification system.

Table 1 Distribution of classes in percentage.

Categorical label for dependent variable	Labelled dependent variable	Distribution %
NS	1	5
P	2	5
C	3	7
B	4	6
BW	5	16
BP	6	14
BPW	7	26
BC	8	11
BCP	9	12

1. Nitrogen Species: Nitrogen species are vital in aquaponics systems. Understanding and regulating these characteristics is critical to sustaining water quality. The key nitrogen species are: Ammonia (NH3-N) concentrations should be less than 0.2 mg/L, whereas NO2-N levels should be below 0.1 mg/L. Proper nitrogen species control promotes the health and well-being of fish, plants, and microbes in the aquaponic system.
2. Phosphorus (P) is an essential component for aquaponics systems, affecting plant

growth and overall System wellness. Important information about phosphorus in aquaponics. Phosphorus is required for energy transmission, cell division, and photosynthesis in plants. Plant Access: Maintaining an ideal phosphorus level ensures that plants have enough access to this nutrient. pH Influence: Phosphorus solubility is influenced by pH. Most plants can efficiently use phosphorus between 5.5-7.5. Regular monitoring is important to prevent deficits or excesses.

3. Fish breathing emits CO_2 into the water, which affects pH.CO_2 spontaneously transforms into carbonic acid (H_2CO_3) when in contact with water. As CO_2 levels rise, the pH of water lowers. Higher fish stocking density causes increased CO_2 emission, reducing pH levels (12).

4. Carbonate hardness (sometimes called alkalinity) assesses water's buffering ability. Carbonate hardness affects pH because it acts as a buffer, lowering the pH. Well water is often hard because it contains natural minerals and carbonates24.Maintaining proper CO_2 levels in aquaponics is crucial for plant development and system health. Carbon dioxide availability influences photosynthesis, nutrient intake, and ecological equilibrium. Monitoring CO_2 levels promotes a healthy habitat for fish and plants.

5. Boron has a critical role in plant physiological activities, including growth. Cell Wall Formation Boron helps to strengthen and stabilize cell walls. Sugar Transport It promotes the flow of sugars throughout plants. Reproduction Boron is necessary for floral growth and seed formation. Maintaining adequate boron levels in aquaponics promotes healthy plant growth and reproduction. Boron deficiency can cause toxicity. Inadequate boron causes symptoms such as deformed leaves, poor blooming, and limited fruit set. Excess boron can be hazardous to plants, resulting in leaf burn and reduced development. Boron levels are regularly monitored to help avoid imbalances. Water Testing and Management Start-up systems should be evaluated on a daily basis to ensure that settings are adjusted swiftly. Once stable, weekly testing is sufficient. Boron influences DO levels, which affect fish and beneficial bacteria12. Boron affects pH stability, nutritional availability, and microbial activity.

To enhance the development of new instances for underrepresented classes, SMOTE disregards neighboring samples and instead draws from examples from other classes, resulting in increased class overlap and noise. Furthermore, SMOTE struggles with data that is highly dimensional, which might result in redundant observations for minority groups. As a consequence, adopting SMOTE usually results in datasets with many duplicates, which reduces model accuracy—a key downside of the SMOTE technique. In response to these constraints, Section 3.4.1 offers H-SMOTE, a novel modified approach for synthetic minority oversampling designed to address imbalanced class concerns. H-SMOTE begins with outlier detection and mitigation, using the Remove Outlier function to find and remove anomalies via the score, or Z approach. This equation computes the score for Z by adding Equations 3 and 4, which establish a mean as well as standard deviation, correspondingly. Equations 5 and 6 identify upper and lower bounds, while Eq. 7 locates outliers, which are subsequently eliminated by Eq. 8, resulting in a devoid-of-noise collection. The collected data is next segmented into dependent and independent factors using Eq. 9, yielding a Sx and Sy dataset that is sent to the H-SMOTE function. The dominant class's length is estimated using Eq. 10, which is indicated by CMJ, while the minority class's list is established using Eq. 11 and given to CMN. Eq. 12 computes the aggregate amount of minority classes (M), whereas Eq. 13 calculates the minority category's ratio to the dominant class. A ratio less than 0.8 suggests a class inequality. The aggregate amount of G synthetic samples generated for minority groups are calculated using Eq. 14. Equations 15 and 16 are then utilized to produce G samples and include them in the dataset. This method creates a fresh dataset that is devoid of outlier values, making it appropriate for preparing specimens for minority classes.

3.4.1 H-SMOTE Algorithm

Define Remove Outliers (Dataset):
Problem description: Use the Z-score approach to eliminate outliers from the dataset.
Input: A dataset with attribute ranges and bounds.
Output: A dataset devoid of noise and outliers.
S → Dataset for columns in dataset columns: z = (x-μ)/σ # μ denotes mean and σ represents standard deviation. (1)

$\mu = \frac{(\sum_{i=1}^{n} Col[i]}{length(col)}$ (2)

$\sigma = \sqrt{\frac{\sum_{i=1}^{n}(col[i]-\mu)^2}{length(col)}}$ (3)

Finding the Boundary Values.

GBx = μ + z * σ (4)

KBx = μ - z * σ (5)
# Identifying Outliers

OV = dataset[(S(col[i]) > GBx) OR(S(col[i] < KBx)] (6)

Removing outliers and creating a new dataset. ND

NR = S[(S(col[i]) < GBx) AND (S(col[i] > KBx) (7)

Sx, Sy → Split (NR) (8)

Return Sx and Sy.
Algorithm H-SMOTE (Dataset, Sy)

Problem description: Perform oversampling for minority classes without overlapping.
Input: A dataset including dependent feature Sx and independent feature Sy.
Output: Data for minority classes created synthetically with no class overlap.
Sx, Sy = Remove Outlier (dataset, Sy); # Detect and mitigate outliers.
CMT = Max (Count (Sy)) (9).
SET CMP = CK - CMT (10).
M < -length (CMP) (11).

J=0
WHILE J<M:DO

Ratio = count (CMP[J])/count (MJ) (12)
If the ratio is less than 0.8, then
G = [count (Mv) - count (CMP[J])] *β # Where $0 < \beta \leq 1$ (13)
For z: 1 to G, use Si = Xi + (QZi - Wi) * τ to generate a new synthetic sample (14).
Where τ is random (0,1).
End For

Dataset → Si # Newly produced sample added to dataset (15)

L = L+1 ENDWHILE

The approach is repeated until the dataset includes unique samples from the attribute's authorized range.

3.5 Aquaponics Farming Prediction

In the context of aquaponics farming prediction, we compare six classifier algorithms using performance criteria such as accuracy, execution time, precision, and recall. These classifiers are trained and assessed using a validation dataset. The proposed classifier incorporates CatBoost, gradient boost, XGBoost, a decision tree, Random Forest, with KNN. Each model has distinct characteristics for water quality measures. The proposed system employs the selected classifier to choose the optimal prediction model for incoming input data and then makes predictions.

4 Result and Discussion

The study evaluates the effectiveness of an Internet of Things (I prediction system designed to identify water quality for fish from cold waters, warm-water fish, plants, and microbes. The Python-based solution utilizes a data set given by the West Bengal Pollution Control Board, which comprises 7338 freshwater records from rivers and lakes throughout 27 Indian states. The dataset includes essential parameters. pH, oxygen concentration, temperature, ammonia, nitrite, and nitrate are among the variables examined. The data is broken down into three categories: 70% for training, 15% for validation, and 15% for testing. Several classification algorithms are utilized, and their performance is assessed using several measures, including accuracy, precision, recall, F1-measure, Kappa, ROC curve, Matthew's correlation coefficient (MCC), and execution time. For the optimum prediction, a voting method is employed, with performance indicators including accuracy, Kappa, the F1 measurement, and time to execution. The evaluation is divided into three stages: first, without utilizing any data augmentation techniques (H-SMOTE), secondly with the H-SMOTE methodology, and finally, execution time with and without H-SMOTE.

4.1 Model Achievement Without the H-SMOTE

The suggested system's performance is evaluated without using the H-SMOTE approach. Classifiers such as CatBoost, Gradient Boosting, XGBoost, Decision Tree, Random Forest, and KNN are evaluated using performance metrics such as accuracy, AUC curve, recall, precision, F1 score, Kappa, and MCC. With the exception of KNN, all classifiers achieve a prediction accuracy of more than 93%.

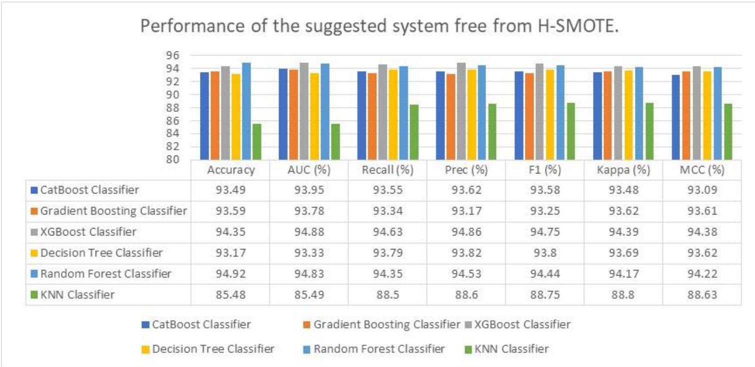

Fig. 2 Performance of the suggested system free from H-SMOTE.

4.2 Model Achievement with the H-SMOTE

The proposed technique's functionality is assessed through the H-SMOTE method. Classifiers such as CatBoost, Gradient Boosting, XGBoost, Decision Tree, Random Forest, and KNN are evaluated using indicators of performance such as accuracy, AUC curve, recall, precision, F1 score, Kappa, and MCC. Except for the KNN classifier, the suggested system obtains a prediction accuracy of more than 96%. Particularly, the XGBoost and Random Forest classifiers acquire accuracy rates of 96% and 99.50%, respectively.

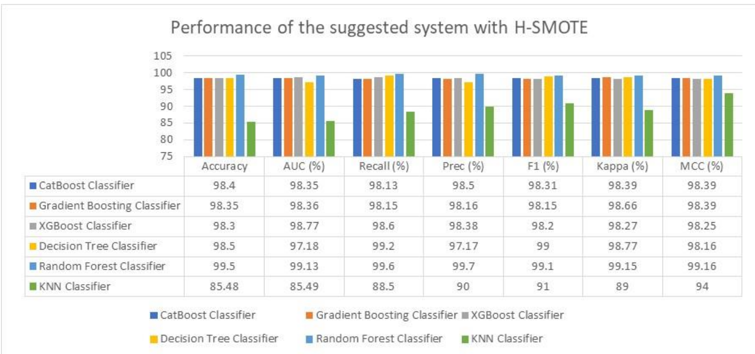

Fig. 3 Performance of the suggested system with H-SMOTE.

5 Conclusion

India ranks 3 for marine output and 2 in aquaculture. The seafood sector engages over 154 lakh persons and contributes to one percent of India's gross domestic product. The National Fisheries Development Board (NFDB) of India estimates that fishing-related businesses generated Rs 334.41 billion in export revenue. Aquaponic systems may help fish breeders make more money. To maintain an optimal environment for fish, plants, and bacteria throughout an ecological framework, the Internet of Things (IoT) system analyses water quality data and offers a method for balancing factors including pH, temperature, DO, ammonia, nitrate, and nitrite. The proposed method for aquaponic culture of a variety of species, including cold-water fish, warm-water fish, bacteria, and plants for aqueous production, yields a most exact prediction imaginable. It is based on a voting system that uses a variety classifier. To attain the highest possible prediction result of 99.50%, the proposed approach includes an effective system for data cleaning, feature selection, and handling of imbalanced data using H-SMOTE, as well as innovative classification and optimal prediction methods based on the voting principal technique. In the future, the suggested technique is expected to apply a return-on-investment approach to improve aquaponics farming and better prediction systems for other aquatic animal farms.

References

[1] Dewali, S., Sharma, N., Melkani, D., Arya, M., Kathayat, N., Panda, A. K., & Bisht, S. S. (2023). "Aquaculture: Contributions to Global Food Security," In Emerging Solutions in Sustainable Food and Nutrition Security (pp. 123–139). Springer.

[2] High Level Panel of Experts on Food Security and Nutrition. (2014). "Sustainable fisheries and aquaculture for food security and nutrition," Food and Agriculture Organization of the United Nations.

[3] B. Nemade and D. Shah, "An IoT-Based Efficient Water Quality Prediction System for Aquaponics Farming," in Computational Intelligence: Select Proceedings of InCITe 2022, Singapore: Springer Nature Singapore, 2023, pp. 311-323. [Online]. Available: https://doi.org/10.1007/978-981-19-7346-8_27

[4] Jones, R., Dewey, B., & Seaver, B. (2022). "Aquaculture Critical for Feeding the World in a Changing Climate," *Perspectives*. The Nature Conservancy.

[5] Food and Agriculture Organization (FAO). (2022). "Record fisheries and aquaculture production makes critical contribution to global food security,"

[6] B. Nemade and D. Shah, "IoT based Water Parameter Testing in Linear Topology," in 2020 10th Int. Conf. Cloud Comput., Data Sci. Eng. (Confluence), Noida, India, 2020, pp. 546-551. doi: 10.1109/Confluence47617.2020.9058224.

[7] Zambrano, A. F., Giraldo, L. F., Quimbayo, J., Medina, B., & Castillo, E. (2021). "Machine learning for manually-measured water quality prediction in fish farming," *PLoS ONE*, 16(8), e0256380.

[8] Capetillo-Contreras, O., Pérez-Reynoso, F. D., Zamora-Antuñano, M. A., Álvarez-Alvarado, J. M., & Rodríguez-Reséndiz, J. (2024). "Artificial Intelligence-Based Aquaculture System for Optimizing the Quality of Water;" A Systematic Analysis. Journal of Marine Science and Engineering, 12(1), 161.

[9] B. Nemade, V. Bharadi, S. S. Alegavi, and B. Marakarkandy, "A Comprehensive Review: SMOTE-Based Oversampling Methods for Imbalanced Classification Techniques, Evaluation, and Result Comparisons," Int. J. Intell. Syst. Appl. Eng., vol. 11, no. 9s, pp. 790-803, 2023.

[10] Li, D., & Du, L. (2022). "Recent advances of deep learning algorithms for aquacultural machine vision systems with emphasis on fish," *Artificial Intelligence Review*, 55, 4077–4116.

[11] V. Shirsath, V. Kaul, R. S. Kumar, and B. Nemade, "Intelligent traffic management for vehicular networks using machine learning," ICTACT J. Commun. Technol., vol. 14, no. 3, 2023.

[12] G. F. M. Baganz et al., "The aquaponic principle — It is all about coupling," Wiley, 2021. DOI: 10.1111/raq.12596

[13] B. Nemade, S. S. Alegavi, N. B. Badhe, and A. Desai, "Enhancing information security in multimedia streams through logic learning machine assisted moth-flame optimization," ICTACT J. Commun. Technol., vol. 14, no. 3, 2023.

[14] M. Lowe, R. Qin, and X. Mao, "A Review on Machine Learning, Artificial Intelligence, and Smart Technology in Water Treatment and Monitoring," Water, vol. 14, no. 9, p. 1384, 2022. DOI: 10.3390/w14091384

[15] M. Elmi Abdi et al.," Correlation Study on Water Quality and Indoor Environment Parameters of Aquaponic Systems Using Statistical and Machine Learning Techniques," SSRN. [Online]. Available: http://dx.doi.org/10.2139/ssrn.4716148

[16] Zambrano, A. F., Giraldo, L. F., Quimbayo, J., Medina, B., & Castillo, E. (2021). "Predicting Aquaculture Water Quality Using Machine Learning Techniques," 2023 International Conference on Sustainable Computing and Data Communication Systems (ICSCDS), 23-25 March 2023. DOI: 10.1109/ICSCDS.2023.10104677.

[17] B. Nemade, S. Doshi, P. Desai, A. Prajapati, K. Patel, and K. Maharana, "Amphibious Trash Collector System," Rivista Italiana di Filosofia Analitica Junior, vol. 14, no. 2, pp. 1360-1371, 2023. ISSN: 2037-4445

[18] S.Sen et al., "Smart Prediction of Water Quality System for Aquaculture using Machine Learning Algorithms," ResearchGate, Oct 2023.[Online]. Available: https://www.researchgate.net/publication/369426212_Smart_Prediction_of_Water_Quality_System_for_Aquaculture_using_Machine_Learning_Algorithms

[19] R. Mishra, B. Nemade, K. Shah, and P. Jangid, "Improved Inductive Learning Approach-5 (IILA-5) in Distributed System," Int. J. Intell. Syst. Appl. Eng., vol. 11, no. 10s, pp. 942-953, 2023.

[20] Zambrano, A. F., Giraldo, L. F., Quimbayo, J., Medina, B., & Castillo, E. (2021). "An Artificial Neural Network Model

for Water Quality and Water Consumption Prediction," *PLoS ONE*, 16(8), e0256380.

[21] B. C. Surve, B. Nemade, and V. Kaul, "Nano-electronic devices with in machine learning capabilities," ICTACT J. Microelectron., vol. 9, no. 3, pp. 1601-1606, 2023. doi: 10.21917/ijme.2023.0277

[22] Capetillo-Contreras, O., Pérez-Reynoso, F. D., Zamora-Antuñano, M. A., Álvarez-Alvarado, J. M., & Rodríguez-Reséndiz, J. (2024). "Application of Machine Learning in River Water Quality Management," A Comprehensive Review. Journal of Marine Science and Engineering, 12(1), 161.

[23] S. S. Alegavi, B. Nemade, V. Bharadi, S. Gupta, V. Singh, and A. Belge, "Revolutionizing Healthcare through Health Monitoring Applications with Wearable Biomedical Devices," Int. J. Recent Innov. Trends Comput. Commun., vol. 11, no. 9s, pp. 752–766, 2023. [Online]. Available: https://doi.org/10.17762/ijritcc.v11i9s.7890

[24] Li, D., & Du, L. (2022). "Short-term Water Quality Variable Prediction Using a Hybrid CNN–LSTM Model," Stochastic Environmental Research and Risk Assessment, 34(2), 415–433.

[25] Niknam, A. R. R., Sabaghzadeh, M., Barzkar, A., & Shishebori, D. (2024). "Comparing ARIMA and Various Deep Learning Models for Long-Term Water Quality Prediction,"

[26] Zambrano, A. F., Giraldo, L. F., Quimbayo, J., Medina, B., & Castillo, E. (2021). "An Optimized Hybrid ARIMA-LSTM Model for Time Series Forecasting of Agricultural Production in India," *PLoS ONE*, 16(8), e0256380.

[27] Taslim, M., & Rahman, M. S. (2024). "Comparative analysis of ARIMA and LSTM for predicting fluctuating time series data," Bulletin of Electrical Engineering and Informatics, 13(3), 107–119.

[28] Putro, S. B., Cahyo, B., Mustika, I. W., & Wahyunggoro, O. (2019). "Improved Time Series Prediction Using LSTM Neural Network for Smart Agriculture Application," In *2019 5th* International Conference on Science and Technology *(ICST)* (pp. 1–6).

[29] Wang, X., Li, Y., Qiao, Q., Tavares, A., & Liang, Y. (2023). "Water Quality Prediction Using ARIMA- SSA-LSTM Combination Model," Entropy, 25(8), 1186. DOI: 10.3390/e25081186.

[30] K. P. R. Abdul Haq and V. P. Harigovindan, "Water Quality Prediction for Smart Aquaculture Using Hybrid Deep Learning Models," IEEE Access, vol. 10, pp. 12345-12356, June 2022, doi: 10.1109/ACCESS.2022.3180482.

[31] X. Wang, Y. Li, Q. Qiao, A. Tavares, and Y. Liang, "Water Quality Prediction Based on Machine Learning and Comprehensive Weighting Methods," Entropy, vol. 25, no. 8, pp. 1186, Aug. 2023. doi: 10.3390/e25081186.

[32] S. -M Petrea et al., "An Analytical Framework on Utilizing Various Integrated Multi-Trophic Scenarios for Basil Production," MDPI, vol. 12, no. 3, p. 540, Jan. 2023. [Online]. Available: https://www.mdpi.com/2223-7747/12/3/540

[33] V. K. P, S. K, B. M. D, and R. Reshma, "Predicting and Analyzing Water Quality using Machine Learning for Smart Aquaculture," in Proc. 2023 International Conference on Sustainable Computing and Data Communication Systems (ICSCDS), Erode, India, 2023, pp. 354-359. doi: 10.1109/ICSCDS56580.2023.10104677.

[34] K. P. R. Abdul Haq and V. P. Harigovindan, "Water Quality Prediction for Smart Aquaculture Using Hybrid Deep Learning Models," IEEE Access, vol. 10, pp. 60078-60098, 2022, doi: 10.1109/ACCESS.2022.3180482.

[35] R. I. Singh and U. K. Lilhore, "A Survey of Machine Learning Models for Water Quality Prediction," in Proc. 2023 8th International Conference on Communication and Electronics Systems (ICCES), Coimbatore, India, 2023, pp. 1069-1074. doi: 10.1109/ICCES57224.2023.10192728.

[36] H. Qin, Y. Tian, H. Wang, Q. Li, X. Cong, and J. Sui, "Research Progress of Aquaculture Environmental Factor Prediction Model Based on Machine Learning," in Proc. 2023 9th International Conference on Control Science and Systems Engineering (ICCSSE), Shenzhen, China, 2023, pp. 263-269. doi: 10.1109/ICCSSE59359.2023.10244955.

[37] V. K. P, S. K, B. M. D, and R. Reshma, "Predicting and Analyzing Water Quality using Machine Learning for Smart Aquaculture," in *Proc. 2023* International Conference on Sustainable Computing and Data Communication Systems *(ICSCDS)*, Erode, India, 2023, pp. 354-359. doi: 10.1109/ICSCDS56580.2023.10104677.

[38] K. P. R. Abdul Haq and V. P. Harigovindan, "Water Quality Prediction for Smart Aquaculture Using Hybrid Deep Learning Models," *IEEE Access*, vol. 10, pp. 60078-60098, 2022, doi: 10.1109/ACCESS.2022.3180482.

[39] R. I. Singh and U. K. Lilhore, "A Survey of Machine Learning Models for Water Quality Prediction," in *Proc. 2023 8th* International Conference on Communication and Electronics Systems (ICCES), Coimbatore, India, 2023, pp. 1069-1074. doi: 10.1109/ICCES57224.2023.10192728.

[40] H. Qin, Y. Tian, H. Wang, Q. Li, X. Cong, and J. Sui, "Research Progress of Aquaculture Environmental Factor Prediction Model Based on Machine Learning," in *Proc. 2023 9th* International Conference on Control Science and Systems Engineering (ICCSSE), Shenzhen, China, 2023, pp. 263-269. doi: 10.1109/ICCSSE59359.2023.10244955.

[41] R. Jongjaraunsuk, W. Taparhudee, and P. Suwannasing, "Comparison of Water Quality Prediction for Red Tilapia

Aquaculture in an Outdoor Recirculation System Using Deep Learning and a Hybrid Model," Water, vol. 16, p. 907, 2024. DOI: 10.3390/w16060907

[42] Q.K Abd Rahman, T. S. M. Osman, and S. Setumin, "Artificial Neural Network: Physico-Chemical and Macronutrients Parameters in an Aquaponic System," 2020. [Online]. Available: https://ir.uitm.edu.my/id/eprint/81473/1/81473.pdf

[43] R. Rena et al., "A Method for Predicting Dissolved Oxygen in Aquaculture Water in an Aquaponics System," Computers and Electronics in Agriculture, vol. 152, pp. 153-161, 2018. DOI: 10.1016/j.compag.2018.06.013

[44] J. Cardenas-Cartagena, M. Elnourani, and B. Beferull-Lozano, "Forecasting Aquaponic Systems Behaviour with Recurrent Neural Networks Models," WISENET Center, University of Agder, Norway, Mar. 2022. DOI: 10.7557/18.6236

[45] S. Lauguico et al., "Adaptive Neuro-Fuzzy Inference System on Aquaphotomics Development for Aquaponic Water Nutrient Assessments and Analyses," Presented at the IEEE International Conference on Industrial Technology and Electrical Engineering (ICITEE), Oct. 2020. DOI: 10.1109/ICITEE49829.2020.9271736

[46] Y. Hendriana et al., "Water Quality Monitoring for Smart Farming Using Machine

[47] Learning Approach," Dec. 2023. [Online]. Available: http://repository.unu-jogja.ac.id/id/eprint/760

[48] B. Nemade and D. Shah, "An IoT-Based Efficient Water Quality Prediction System for Aquaponics Farming," in Proceedings of the International Conference on Internet of Things (ICIOT), 2023. DOI: 10.1007/978-981-19-7346-8_27

[49] C.L Kok et al., "Smart Aquaponics: An Automated Water Quality Management System for Sustainable Urban Agriculture," Electronics, vol. 13, p. 820, 2024. DOI:10.3390/electronics13050820 A. F. Zambrano et al., "Machine Learning for Manually-Measured Water Quality Prediction in Fish Farming," PLoS ONE, vol. 16, no. 8, p. e0256380, 2021. DOI: 10.1371/journal.pone. 0256380

[50] S. Sen et al., "Smart Prediction of Water Quality System for Aquaculture Using Machine Learning Algorithms," Mar. 2023. DOI: 10.36227/techrxiv.22300435

[51] W. Khan et al., "Water Quality Requirements and Management Strategies for Fish Farming: A Case Study of Ponds around Gurgaon Canal, Nuh, Palwal," International Journal of Trend in Scientific Research and Development, vol. 2, no.1, 2018A. C. Godoy et al., "Water Quality in a Reservoir used for Fish Farming in Cages in Winter and Summer Periods," Water Air Soil Pollut, vol. 229, p. 63, 2018. DOI: 10.1007/s11270-017-3669-x

[52] R. Y. Tallar and J.-P. Suen, "Aquaculture Water Quality Index: a low-cost index to accelerate aquaculture development in Indonesia," Aquaculture International, 2015. DOI: 10.1007/s10499-015-9926-3

[53] T. Y. Kyaw and A. K. Ng, "Smart Aquaponics System for Urban Farming," Energy Procedia, vol.141, pp. 385-391, Jul. 2017. DOI: 10.1016/j.egypro.2017.12.694

[54] Y. Chen et al., "A Review of the Artificial Neural Network Models for Water Quality Prediction," Applied Sciences, vol. 10, no. 17, Aug. 2020. DOI: 10.3390/app10175776 A. Ahmed et al., "Machine learning methods for better water quality prediction," Journal of Hydrology, 2019. DOI: 10.1016/j.jhydrol.2019.124084

[55] H. Haghiabi et al., "Water quality prediction using machine learning methods," Water Quality Research Journal of Canada

[56] T. H. H. Aldhyani, M. Maashi, M. AL-Yaari, and Alkahtani, "Water Quality Prediction Using Artificial Intelligence Algorithms," vol.2020, Article ID 6659314, 12 pages, Dec. 2020. DOI: 10. 1155/2020/

[57] 6659314. A. Najah, A. EL-Shafie, and O. Karim, "Application of artificial neural networks for water quality prediction," Neural Computing and Applications, vol. 21, no. 6, pp. 1495-1503, Apr. 2012. DOI: 10.1007/s00521-012-0940-3

[58] R. Avila, B. Horn, E. Moriarty, R. Hodson, and E. Moltchanova, "Evaluating statistical model performance in water quality prediction," Journal of Environmental Management, vol. 206, pp. 1197- 1206, Feb. 2018. DOI: 10.1016/j.jenvman.2017.11.049.

[59] EncinasC, RuizE, CortezJ,EspinozaA(2017) "Design and implementation of a distributed IoT system for the monitoring of water quality in aquaculture," Wirel Telecommun Symp. https:// doi.org/10.1109/WTS.2017.7943540

[60] Yep B, Zheng Y (2019) "Aquaponic trends and challenges—a review," J Clean Prod 228:1586 1599. https://doi.org/10.1016/j.jclepro.2019.04.290

[61] Francisco HR, Corrêia AF, Feiden A (2019) "Classification of areas suitable for fish farming using geotechnology and multi-criteria analysis," ISPRS Int J Geo-Inf 8(9). https://doi.org/10. 3390/ijgi8090394

[62] Wirza R, Nazir S (2021) "Urban aquaponics farming and cities—a systematic literature review," Rev Environ Health 36(1):47–61. https://doi.org/10.1515/reveh-2020-0064

[63] Villarroel M et al (2016) "Survey of aquaponics in Europe," Water (Switzerland) 8(10):3–9.

https://doi.org/10.3390/w8100468

[64] Yogev U, Barnes A, Gross A (2016) "Nutrients and energy balance analysis for a conceptual model of three loops off grid, aquaponics," Water (Switzerland)8(12). https://doi.org/10.3390/ w8120589

[65] Gunning D, Maguire J, Burnel G (2016) "The development of sustainable saltwater-based food production systems: a review of established and novel concepts," Water (Switzerland) 8(12). https://doi.org/10.3390/w8120598

[66] Duque G, Gamboa-García DE, Molina A, Cogua P (2020) "Effect of water quality variation on fish assemblages in an anthropogenically impacted tropical estuary, Colombian Pacific," Environ Sci Pollut Res 27(20):25740–25753. https://doi.org/10.1007/s11356-020-08971-2

[67] Junge R, König B, Villarroel M, Komives T, Jijakli MH (2017) "Strategic points in aquaponics," Water (Switzerland) 9(3):1–9. https://doi.org/10.3390/w9030182

[68] Yildiz HY, Robaina L, Pirhonen J, Mente E, Domínguez D, Parisi G (2017) "Fish welfare in aquaponic systems: its relation to water quality with an emphasis on feed and faeces—a review," Water (Switzerland) 9(1):1–17. https://doi.org/10.3390/w9010013

[69] Chen JH, Sung WT, Lin GY (2016) "Automated monitoring system for the fish farm aquaculture environment" In: Proceedings—2015 IEEE international conference on systems, man and cybernetics SMC 2015, pp 1161–1166. https://doi.org/10.1109/SMC.2015.208

[70] SurnarSR, SharmaOP, SainiVP (2015) "Aquaponics: innovative farming.Int," J Fish Aquat Stud2(4):261– 263

[71] Abinaya T,Ishwarya J,Maheswari M(2019) "A novel methodology for monitoring and controlling of water quality in aquaculture using internet of things (IoT)," In: 2019 International conference on computer communication and informatics, ICCCI 2019, pp 1–4. https://doi.org/10.1109/ICCCI.2019.8821988

[72] "A study on fish culture system in Kotalipara Upazila," Gopalganj 2(3):59–70 (2013)

[73] Bhatnagar A, Devi P (2013) "Water quality guidelines for the management of pond fish culture," Int J Environ Sci 3(6):1980–2009. https://doi.org/10.6088/ijes.2013030600019

[74] Khan W, Vahab A, Masood A, Hasan N (2017) "Water quality requirements and management strategies for fish farming a case study of ponds around Gurgaon canal NUH Palwal," Int J Trend Sci Res Dev 2(1):388–393. https://doi.org/10.31142/ijtsrd5914

[75] Ahmed M, Rahaman O, Rahman M, Kashem MA (2020) 2020 2nd international conference on "sustainable technologies for Industry 4.0," STI 2020, pp 1–5

[76] GodoyACetal(2018) "Water quality in a reservoir used for fish farming in cages in winter and summer periods," Water Air Soil Pollut 229(3). https://doi.org/10.1007/s11270-017-3669-x

[77] Tallar RY, Suen JP (2016) "Aquaculture water quality index: a low-cost index to accelerate aquaculture development in Indonesia," Aquac Int 24(1):295–312. https://doi.org/10.1007/s10499-015-9926-3

[78] Kyaw TY, Ng AK (2017) "Smart aquaponics system for urban farming," Energy Proc 143:342-347. https://doi.org/10.1016/j.egypro.2017.12.694

[79] Nemade and D. Shah, "An efficient IoT based prediction system for classification of water using novel adaptive incremental learning framework," in Journal of King Saud University-Computer and Information Sciences, vol. 34, no. 8, pp. 5121-5131, 2022. [Online]. Available: https://doi.org/10.1016/j.jksuci.2022.01.009

[80] K. Maharana, S. Mondal, and B. Nemade, "A review: Data pre-processing and data augmentation techniques," in Global Transitions Proceedings, vol. 3, no. 1, pp. 91-99, 2022. [Online]. Available: https://doi.org/10.1016/j.gltp.2022.04.020

[81] Nemade and D. Shah, "IoT based Water Parameter Testing in Linear Topology," in 2020 10th Int. Conf. Cloud Comput., Data Sci. Eng. (Confluence), Noida, India, 2020, pp. 546-551. doi: 10.1109/Confluence47617.2020.9058224.

[82] B. Nemade and D. Shah, "An IoT-Based Efficient Water Quality Prediction System for Aquaponics Farming," in Computational Intelligence: Select Proceedings of InCITe 2022, Singapore: Springer Nature Singapore, 2023, pp. 311-323. [Online]. Available: https://doi.org/10.1007/978-981-19-7346- 8_27

[83] K. B. T. S. and M. Challa, "Prediction of Water Quality System for Aquaculture using Machine Learning," Turkish Online Journal of Qualitative Inquiry (TOJQI), vol. 12, no. 6, pp. 321-328, Jun. 2021.