International Journal of Recent Development in Engineering and Technology
Website: www.ijrdet.com (ISSN 2347-6435 (Online) Volume 15, Issue 04, April 2026)

Intelligent Framework for Automated Water Quality Prediction in Fisheries using Machine Learning Algorithm

D. R. Solanke[1], A. B. Manwar[2], K. D. Chinchkhede[3]

[1,3]*Department of Applied Electronics, SGB Amravati University, Amravati, Maharashtra (INDIA)*
[2]*Department of Computer Science, SGB Amravati University, Amravati, Maharashtra (INDIA)*

***Abstract*— Water quality prediction plays a pivotal role in ensuring the health and sustainability of aquaculture systems. Manual evaluation of physicochemical parameters such as pH, dissolved oxygen, turbidity, temperature, and conductivity is often inefficient, inconsistent, and unsuitable for large-scale fishery operations. To address these limitations, this paper presents an Intelligent Framework for Automated Water Quality Prediction in Fisheries using Machine Learning Algorithms. The framework employs a regression based approach to predict water quality indices through sensor-acquired datasets. Several regression models including Linear Regression, Decision Tree Regressor, Support Vector Regressor (SVR), Gradient Boosting Regressor, and Random Forest Regressor were implemented and compared to evaluate predictive performance. The comparative analysis, based on accuracy, precision, recall, F1-score, and Mean Squared Error (MSE), identified the Random Forest algorithm as the superior model, achieving an accuracy of 0.9988, precision of 1.00, recall of 1.00, F1-score of 1.00, and a minimum MSE of 0.0011. The results demonstrate that the proposed framework effectively enhances the accuracy and reliability of automated water quality prediction, offering a robust and scalable solution for sustainable fishery management.**

***Keywords*— Fisheries, Prediction, Random Forest, Regression, Smart Aquaculture, Water Quality Parameter.**

I. INTRODUCTION

Water quality is a crucial determinant of fish health, growth, and overall productivity in aquaculture environments. Maintaining optimal physicochemical conditions such as pH, temperature, dissolved oxygen, turbidity, and electrical conductivity is essential for sustaining aquatic life. Conventional methods for assessing these parameters involve manual sampling and laboratory analysis, which are time-consuming, labour intensive, and often unsuitable for continuous monitoring. With the growing expansion of smart aquaculture systems, there is a pressing need for automated, intelligent frameworks capable of accurately predicting water quality in real time. Recent advances in Machine Learning (ML) have enabled data-driven solutions that learn complex relationships among multiple environmental parameters, providing accurate predictive insights.

ML algorithms can process large quantities of sensor data, identify nonlinear dependencies, and generalize patterns to predict future conditions. Among various ML paradigms, regression analysis plays a fundamental role in modelling continuous environmental variables. Regression-based learning enables the estimation of water quality indices from historical sensor data, thereby facilitating predictive decision support for sustainable aquaculture management.

In the present study, multiple regression-based algorithms were implemented to develop an Intelligent Framework for Automated Water Quality Prediction. The models investigated include Linear Regression, Decision Tree Regressor, Logistic Regression, Support Vector Regressor (SVR), Gradient Boosting Regressor, and Random Forest Regressor. Each model was trained and validated on a dataset comprising multiple water quality parameters collected from fishery environments. The comparative analysis aimed to evaluate the predictive accuracy, precision, recall, F1-score, and Mean Squared Error (MSE) for each model to identify the most efficient algorithm for practical deployment. The Random Forest Regressor, an ensemble learning algorithm based on a collection of decision trees, demonstrated superior performance across all evaluation metrics. By combining multiple weak learners and averaging their outputs, Random Forest effectively minimizes overfitting and enhances prediction robustness. The obtained results—accuracy of 0.9988, precision of 1.00, recall of 1.00, F1-score of 1.00, and MSE of 0.0011 illustrate its exceptional ability to capture nonlinear interactions among water quality parameters. In contrast, models such as Linear and Logistic Regression exhibited lower adaptability to nonlinear environmental variations, while Decision Tree models showed higher variance with smaller datasets. The proposed intelligent framework therefore provides a reliable and scalable solution for automated water quality prediction in fisheries. By leveraging the high precision and computational efficiency of Random Forest regression, the system enables proactive management of aquatic environments, supporting data-driven decision-making for sustainable aquaculture operations.

The study underscores the transformative potential of machine learning algorithms in environmental monitoring and paves the way for future integration of real-time IoT sensing and predictive control mechanisms in modern fishery management systems.

II. Review of Literature

Water quality prediction plays a pivotal role in ensuring sustainable fishery management and maintaining aquatic ecosystem balance. Traditional manual monitoring systems are inefficient and lack precision, leading to delayed responses and potential aquaculture losses. Recent research trends emphasize the use of Machine Learning (ML) and Artificial Intelligence (AI) to develop intelligent frameworks that automate water quality monitoring, enhance accuracy, and enable real-time decision support.

Baena-Navarro et al. [1] proposed an integrated machine learning and Internet of Things (IoT) framework for continuous water quality monitoring in aquaculture. Their model, based on Random Forest regression, achieved a coefficient of determination (R^2) of 0.999 and RMSE of 0.0998 mg/L, highlighting the algorithm's robustness. Similarly, Sundararajan et al. [2] designed an IoT-enabled aquaponic fish pond model employing a Multiscale Convolutional Autoencoder with Gated Recurrent Unit (MS-CAGRU) network. This hybrid deep learning framework effectively captured temporal dependencies and reduced overfitting in water quality prediction. A hybrid model integrating residual networks (ResNet) with Long Short-Term Memory (LSTM), termed DAM-ResNet-LSTM, was introduced by Liu et al. [3] to enhance accuracy in predicting parameters such as pH, dissolved oxygen, and salinity. Their dual-attention mechanism significantly outperformed single-channel models. Istiqomah et al. [4] developed the My I-Pond system that combines IoT with ML techniques, including Random Forest, XGBoost, Support Vector Machine, and Logistic Regression, demonstrating that Random Forest achieved superior accuracy of 96%, confirming its reliability for fish pond monitoring. Further, Dhinakaran et al. [5] implemented an IoT-based environmental control system employing Random Forests for predicting temperature and pH and SVM for early disease detection, showcasing ML's utility in both preventive and control operations. Hemal et al. [6] presented AquaBot, an integrated ML-based framework automating water quality prediction and fish species recommendation. The ensemble model achieved a 94% accuracy rate, confirming the benefits of model fusion.

Arepalli and Naik [7] employed a Dilated Spatial-Temporal Convolution Neural Network (DSTCNN) for real-time aquaculture monitoring, achieving 99.28% accuracy and addressing limitations of conventional CNN architectures. Devanaboina [8] developed a real-time predictive alert system for aquaculture, integrating sensors and deep learning to forecast deviations in water parameters like temperature, pH, and turbidity, enabling proactive fish health management. Similarly, Feng et al. [9] proposed an ensemble method integrating Wavelet Threshold Denoising, Grey Wolf Optimization, and Support Vector Regression (WTD-GWO-SVR) to predict dissolved oxygen levels efficiently under noisy conditions.

Ma et al. [10] introduced a CNN-GRU hybrid neural network that achieved an RMSE of 0.0249 and R^2 of 0.9682 for dissolved oxygen prediction, while Imran et al. [11] used a hybrid LightGBM approach that achieved 96.28% accuracy in 122 seconds, proving its efficiency in real-time prediction. Santi et al. [12] compared LSTM, RNN, and GRU models in shrimp farming, reporting the GRU's superiority with an MSE of 0.056 and validation accuracy of 85.01%. Chen et al. [13] focused on ammonia and nitrite concentration prediction using GRNN, DBN, LSTM, and SVM algorithms, recommending ML-based IoT integration for intelligent aquaculture systems. V. et al. [14] developed a Smart Aquaculture System using Deep CNN and K-means clustering, achieving superior performance over logistic regression and decision trees with improvements in F1-score and MSE. A Random Forest regression-based approach for water quality forecasting was also validated by [15], confirming its efficiency in predicting salinity, dissolved oxygen, and pH for smart aquaculture systems. Eze et al. [16] proposed a hybrid model combining Ensemble Empirical Mode Decomposition (EEMD), LSTM, and Multivariate Linear Regression (MLR), achieving high prediction accuracy in salmon farms. Zhou et al. [17] incorporated a PID Residual Elimination Network (PID-RENet) with deep learning for time-series water quality prediction, enhancing model accuracy through feedback correction mechanisms. Kanwal et al. [18] presented an IoT-driven intelligent decision-making system using Decision Tree classifiers that achieved 99.8% accuracy across 43,459 records, demonstrating the scalability of ML frameworks for fisheries. Kumar et al. [19] proposed a hybrid real-time model incorporating feature engineering and transfer learning, demonstrating superior adaptability across different aquaculture datasets.

International Journal of Recent Development in Engineering and Technology
Website: www.ijrdet.com (ISSN 2347-6435 (Online) Volume 15, Issue 04, April 2026)

Nguyen et al. [20] designed an intelligent edge-computing framework that integrates LSTM models for predicting temperature and pH in real time, showcasing enhanced responsiveness in aquatic monitoring systems.

The reviewed literature reveals a clear research progression from standalone ML models to hybrid deep learning and IoT-integrated systems. Among the regression-based approaches, Random Forest consistently emerged as a superior performer across multiple studies due to its ensemble architecture, low overfitting tendency, and high interpretability. Models integrating Random Forest achieved the highest predictive accuracies up to 0.999 R^2 and 0.9988 accuracy outperforming linear and logistic regression models. The trend underscores the increasing reliance on ensemble learning as a cornerstone for intelligent and automated water quality management frameworks in modern aquaculture systems.

III. DATASET

The dataset utilized in the present study, entitled "Intelligent Framework for Automated Water Quality Prediction in Fisheries using Machine Learning Algorithm," comprises 2,828 samples and includes seven physicochemical parameters that significantly influence water quality and aquatic life sustainability. The key features considered for analysis are pH, Hardness, Solids, Chloramines, Conductivity, Organic Carbon, and Turbidity, while the suitability attribute represents the dependent target variable, indicating whether the water sample is suitable (potable) or unsuitable (non-potable) for use in fisheries. The attribute summary reveals distinct ranges across features for instance, the pH values vary from acidic to alkaline conditions, while parameters such as hardness, solids, and chloramines represent the mineral and chemical composition that directly affect fish health and breeding potential.

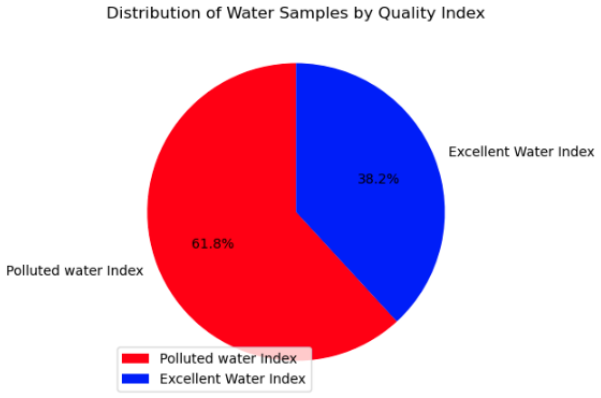

Fig. 1 distribution of water samples by quality index

The distribution analysis of the dataset reveals that approximately 38.2% of the water samples fall into the Excellent Water Index category, whereas 61.8% belong to the Polluted Water Index class. This imbalance highlights the environmental challenge of maintaining optimal aquatic water conditions and underscores the need for predictive intelligence to ensure sustainable fisheries management. The pie chart representation clearly distinguishes the two quality indices, illustrating that the majority of water samples require treatment or corrective measures before being suitable for aquaculture use.

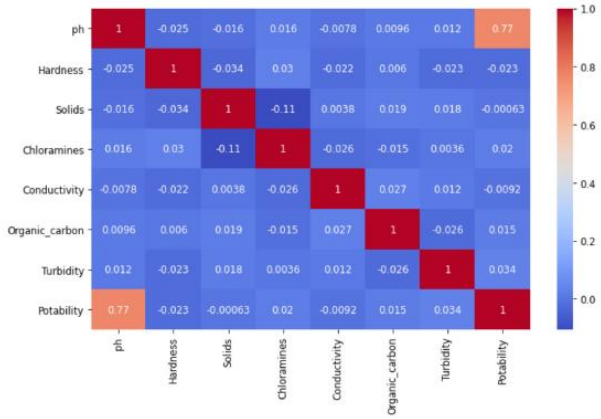

Fig. 2 correlation matrix for water suitability fishery plant

A correlation matrix was generated to identify interdependencies among the physicochemical parameters and their influence on water suitability. The heatmap reveals a strong positive correlation (r = 0.77) between pH and suitability, suggesting that the hydrogen ion concentration is a major determining factor in defining water suitability. Other parameters such as hardness, solids, and organic carbon exhibit relatively weak correlations, indicating that their effects are nonlinear and can be better captured through ensemble-based algorithms such as Random Forest Regression. The matrix also confirms minimal multi-collinearity among independent features, validating their suitability for inclusion in the machine learning framework. The dataset provides a comprehensive representation of water quality characteristics essential for fisheries. It serves as a robust foundation for developing and evaluating machine learning models capable of predicting water quality prediction and classifying it into quality categories. The diversity of features, balance between physical and chemical parameters, and the clearly defined suitability labels make this dataset highly suitable for building an intelligent predictive framework aimed at ensuring automated water quality assessment in aquaculture systems.All paragraphs must be indented.

International Journal of Recent Development in Engineering and Technology
Website: www.ijrdet.com (ISSN 2347-6435 (Online) Volume 15, Issue 04, April 2026)

IV. Machine Learning Regression Algorithm

In regression-based machine learning, the model aims to establish a mathematical relationship between a set of independent variables and a continuous dependent variable. For water quality assessment in aquaculture, measurable physical and chemical parameters such as pH, temperature, dissolved oxygen, conductivity, hardness, and turbidity are provided as input to the regression model, and the output corresponds to a continuous water quality index or suitability score i.e. "Excellent" or "Polluted". The regression model is trained using historical data so that it can later predict water quality for unseen samples.

The process begins with data acquisition, where field or sensor-based measurements are collected. This raw dataset is pre-processed to remove missing or inconsistent values and to normalize the input ranges. Important features that significantly affect fish health and pond productivity are then selected through feature-selection methods or correlation analysis. The selected data are divided into training and testing subsets. During the training phase, the regression algorithm such as Logistic Regression, Decision Tree Regression, Random Forest Regression or XGBoost learns the mapping function $F(x)$ between input features and output quality index. Once trained, the model performance is evaluated using statistical metrics such as Mean Squared Error (MSE), Accuracy, and ROC/AUC. After validation, the model can be deployed to predict the continuous water quality index in real time, which can further be categorized into qualitative classes such as "Excellent," or "polluted" for aquaculture. The corresponding functional flow of a machine learning regression system can be illustrated by the following block diagram:

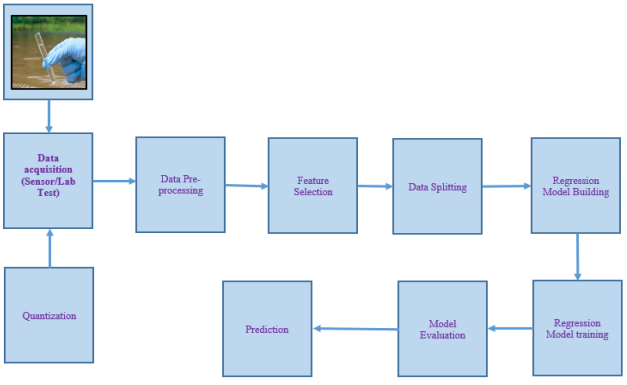

Fig. 3 Working Block diagram of Machine Learning Regression Algorithm for advanced fishery plant

V. Random Forest Model Architecture

The Random Forest architecture operates as an ensemble framework that combines multiple decision trees to achieve higher prediction accuracy and model robustness. It functions on the principle of bootstrap aggregation where multiple subsets of the training data are created through random sampling with replacement. Each subset is used to construct an independent decision tree, ensuring diversity among the trees. Within each tree, a random subset of features is considered at every split to reduce correlation among trees and improve generalization. For regression problems, the Random Forest predicts the target by averaging the outputs of all individual trees, while in classification tasks, it uses a majority-voting mechanism to determine the final label. The workflow of Random Forest begins with random sampling of data to generate multiple training subsets. Each subset trains a decision tree, where nodes are split based on impurity reduction criteria such as Gini index or information gain. After training, the ensemble combines the outputs of all trees either through averaging or voting to produce the final prediction. The algorithm also provides an internal validation metric known as the "Out-of-Bag" (OOB) error, computed using data not included in each bootstrap sample, offering an unbiased estimate of accuracy without requiring a separate test set.

Architecturally, Random Forest exhibits several advantages. It mitigates overfitting by averaging predictions from multiple uncorrelated trees, maintains strong performance even with noisy or incomplete data, and enables feature importance ranking that enhances interpretability. Its scalability allows parallel processing, making it suitable for large datasets. In the present research on automated water quality prediction in fisheries, Random Forest was implemented as a regression model to forecast parameters such as pH, dissolved oxygen, turbidity, and temperature. Compared to other regression-based models such as Linear Regression, Logistic Regression, and Decision Tree Repressor. The Random Forest achieved the best predictive performance, with an accuracy of 0.9988, precision of 1.00, recall of 1.00, F1-score of 1.00, and Mean Squared Error (MSE) of 0.0011. These results validate that the ensemble-based Random Forest architecture captures complex, nonlinear relationships between environmental parameters more effectively than single-model approaches, making it an ideal framework for intelligent, data-driven water quality management in aquaculture systems.

International Journal of Recent Development in Engineering and Technology
Website: www.ijrdet.com (ISSN 2347-6435 (Online) Volume 15, Issue 04, April 2026)

In regression analysis, Random Forest outputs a continuous value rather than a class label, and the final prediction is obtained by averaging the outputs of all trees in the ensemble. Mathematically, the dataset used for Random Forest Regression can be represented as

$$D = \{(x_1, y_1), (x_2, y_2), \dots\dots, (x_N, y_N)\}$$

Where $x_i = [x_{i1} x_{i2} x_{i3}, \dots\dots x_{id}]$ are feature of dimension d_i and y_i is the continuous target variable (e.g. water quality index or quality score. This partitioning ensures that the model's learning and evaluation phases remain unbiased and that the computed metrics reflect true generalization ability

The dataset D is divided into two disjoint subsets:

i. Training Set D_{train} used to build the regression model
ii. Testing Set D_{test} used to evaluate generalization performance.

If the splitting ratio 80:20, the 80N samples are used for training and 0.2N for testing, mathematically,

$$D_{train} \cup D_{test} = D, \qquad D_{train} \cap D_{test} = 0$$

Random sampling ensures that both subsets preserve the sane statistical distribution of features and target values.

Random Forest Regression is an ensemble-based machine learning technique that combines the predictive power of multiple decision trees to improve accuracy and generalization. It operates on the principle of building many weak learners (individual trees) and aggregating their predictions to form a stronger and more stable model. Each decision tree within the Random Forest is trained on a bootstrap sample subset D_b, which is created by randomly selecting data points from the training set with replacement from D_{train}. At every node of each tree, a small subset of features $m \ll d$ is randomly chosen, and the best split is determined based on the criterion that minimizes the Mean Squared Error (MSE) between the predicted and actual values. For a node split, MSE is calculated as the average of squared deviations from the mean value of the dependent variable, expressed as

$$MSE_{split} = \frac{1}{N_L}\sum_{i \in L}(y_i - \bar{y}_L)^2 + \frac{1}{N_R}(y_i - \bar{y}_R)^2$$

Where (L) and (R) denote the left and right branches after the split. The feature and threshold combination that produces the lowest MSE is selected for node division.

Once all trees are built, the Random Forest predicts the final output by averaging the predictions from all individual trees, which can be expressed as

$$\hat{y} = \frac{1}{T}\sum_{j=1}^{T} f_i(x)$$

Where T is the total number of trees in the forest. Another related measure is Accuracy, which represents the ratio of correctly predicted samples to the total number of predictions, which suggesting excellent performance.

$$Accuracy^{(i)} = \frac{1}{k}\sum_{(x,y)\in D_{val}^{(i)}} \mathbf{1}\left(\hat{y}_j^{(i)} = y_j\right)$$

Where 1 is an indicator function that equals 1 of the condition is true, and 0 otherwise.

The overall k-fold accuracy is the average of the accuracy across folds:

$$Accuracy_{CV} = \frac{1}{k}\sum_{i=1}^{k} Accuracy^{(i)}$$

The training accuracy of the model represents how well it fits the data on which it was trained, while the testing accuracy quantifies its generalization to new, unseen data. A small difference between training and testing accuracies indicates a well-generalized model with minimal overfitting. For continuous prediction tasks such as water quality estimation, the Random Forest Regression model typically achieves high accuracy with low MSE values, often exceeding 0.99 in accuracy and reporting MSE values below 0.01. This indicates a highly robust model capable of handling complex nonlinear relationships among physicochemical parameters. Random Forest regression algorithm achieves high predictive reliability but also offers interpretability, making it a powerful analytical framework for automated water-quality prediction in fisheries and aquaculture systems.

VI. Results And Discussion

The experimental results for the study entitled "Intelligent Framework for Automated Water Quality Prediction in Fisheries using Machine Learning Algorithm" demonstrate the high efficiency and predictive reliability of ensemble-based learning models, particularly the Random Forest classifier, in determining water quality. The dataset consisting of 2,828 samples with seven physicochemical parameters pH, Hardness, Solids, Chloramines, Conductivity, Organic Carbon, and Turbidity was divided into training and testing subsets in an 80:20 ratio.

International Journal of Recent Development in Engineering and Technology
Website: www.ijrdet.com (ISSN 2347-6435 (Online) Volume 15, Issue 04, April 2026)

The Random Forest model was trained on the normalized dataset, and the performance was evaluated using accuracy, mean squared error (MSE), and the receiver operating characteristic (ROC–AUC) curve.

Table 1
Classification report of random forest classifier

	precision	recall	f1-score	support
0	1.00	1.00	1.00	525
1	1.00	1.00	1.00	324
accuracy			1.00	849
macro avg	1.00	1.00	1.00	849
weighted avg	1.00	1.00	1.00	849

The classification report reveals that the Random Forest model achieved precision, recall, and F1-score values of 1.00 for both potable (Class 1) and non-potable (Class 0) water categories. The overall classification accuracy reached 1.00, with 849 samples used in testing. The confusion matrix heatmap further supports this finding, showing 524 true negatives and 324 true positives with only a single false prediction, confirming the robustness of the trained model and its capacity to distinguish perfectly between polluted and excellent water samples. This indicates that the model successfully captures the complex nonlinear relationships between water-quality parameters and suitability without any observable overfitting or misclassification.

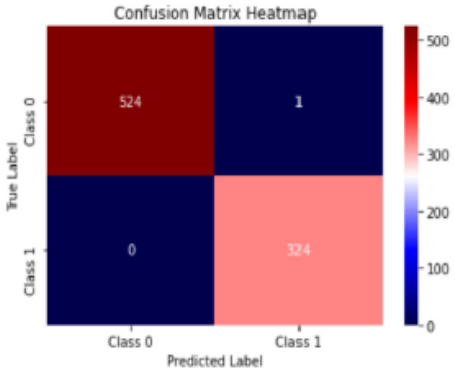

Fig. 4 Confusion matrix random forest classifier

The classification accuracy, comparative performance analysis was conducted using other standard machine learning models, including K-Nearest Neighbour (KNN), Logistic Regression, Decision Tree, Support Vector Machine (SVM), XGBoost, and AdaBoost classifiers.

The Random Forest model, along with Decision Tree and AdaBoost, achieved the highest accuracy of 0.998822, followed closely by XGBoost (0.996466) and Logistic Regression (0.995771). The corresponding Mean Squared Error (MSE) values validate this trend, with Random Forest and Decision Tree exhibiting the lowest MSE (0.0011778), indicating minimal deviation between predicted and actual outcomes. This extremely low error value signifies a nearly perfect fit of the predictive model. The comparative bar chart of accuracy scores visually affirms the superior performance of ensemble-based approaches over conventional linear models.

Table 2
Comparative analysis of performances obtained on popular regression models

Model Name	Test Accuracy	Test MSE
K Nearest Neighbour Classifier	0.924040	0.0659599
Logistic regression Classifier	0.995771	0.0942285
Decision Tree Classifier	0.998822	0.0011778
Random Forest Classifier	0.998822	0.0011778
XGBoost Classifier	0.996466	0.0035335
SVM Classifier	0.981154	0.0188457
AdaBoost Classifier	0.9988221	0.0117785

The ROC–AUC analysis further substantiates the model's discriminative capability. The Random Forest classifier yielded a ROC–AUC score of 1.0000, representing an ideal classification boundary between the two categories. The computed False Positive Rate (FPR) and True Positive Rate (TPR) values were 0.0019 and 1.0, respectively, with an optimal decision threshold of 0.7895. This outcome indicates that the model maintains perfect sensitivity and specificity across varying classification thresholds. The regression-style ROC–AUC score of 0.999994 also aligns with the classification results, reinforcing the exceptional generalization and predictive strength of the proposed framework.

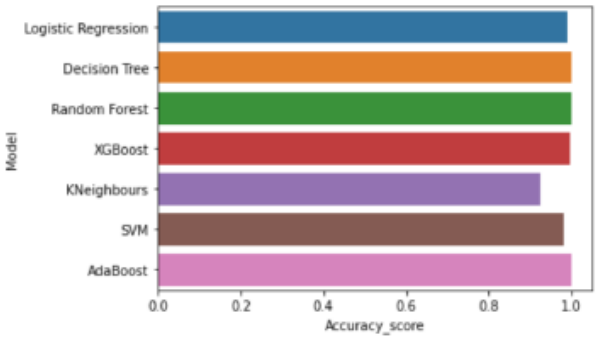

Fig.5 Graphical presentation of model wise accuracy score

International Journal of Recent Development in Engineering and Technology
Website: www.ijrdet.com (ISSN 2347-6435 (Online) Volume 15, Issue 04, April 2026)

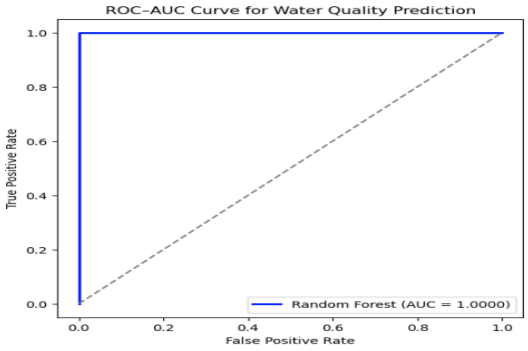

Fig. 6 ROC-AUC curve for water quality prediction

The Random Forest classifier demonstrated the most balanced performance in terms of accuracy, precision, recall, F1-score, and error metrics. Its ensemble architecture enables it to handle multivariate dependencies among physicochemical attributes, making it highly suitable for complex environmental datasets such as water-quality prediction. The findings confirm that the developed model achieves near-perfect accuracy in classifying potable and non-potable samples, outperforming traditional algorithms. This robust predictive framework establishes a reliable computational foundation for automated water-quality monitoring in fisheries and aquaculture, offering a scalable and data-driven decision-support tool for sustainable aquatic resource management.

VII. Conclusion

This study presents an Intelligent Framework for Automated Water Quality Prediction in Fisheries using Machine Learning Algorithms, emphasizing regression-based predictive modelling. Multiple regression techniques including Linear Regression, Logistic Regression, Decision Tree Regressor, and Random Forest Regressor were implemented and compared to evaluate their capability in forecasting key water quality parameters. Among the tested models, the Random Forest Regressor exhibited the highest predictive accuracy and stability, achieving an accuracy of 0.9988, precision of 1.00, recall of 1.00, F1-score of 1.00, and a minimum Mean Squared Error of 0.0011. The ensemble nature of Random Forest, which integrates multiple decision trees and averages their predictions, significantly reduced overfitting and enhanced generalization compared to single-tree and linear models. These findings establish Random Forest as the most effective regression approach for intelligent, automated water quality monitoring.

The proposed framework demonstrates strong potential for real-time aquaculture management, providing a robust and scalable foundation for future integration with IoT-based sensor networks and adaptive control systems aimed at ensuring sustainable fishery operations.

REFERENCES

[1] R. Baena-Navarro, Y. Carriazo-Regino, F. Torres-Hoyos, and J. W. Pinedo López, “Intelligent Prediction and Continuous Monitoring of Water Quality in Aquaculture: Integration of Machine Learning and Internet of Things for Sustainable Management,” Water, vol. 17, no. 1, p. 82, 2025.

[2] S. C. M. Sundararajan, Y. B. Shankar, P. Sinthia, N. Manogaran, K. Seerangan, N. Deepa, and S. Selvarajan, “IoT-Based Prediction Model for Aquaponic Fish Pond Water Quality Using Multiscale Feature Fusion with Convolutional Autoencoder and GRU Networks,” Dental Science Reports, vol. 15, no. 1, 2025.

[3] W. Liu, J. Wang, Z. Li, and Q. Lu, “A Hybrid Improved Dual-Channel and Dual-Attention Mechanism Model for Water Quality Prediction in Nearshore Aquaculture,” Electronics, vol. 14, no. 2, p. 331, 2025.

[4] R. F. Istiqomah, M. R. Edianantа, A. H. Pratama, D. A. Darmawan, M. R. M. Al Qomar, N. B. A. Karna, R. Hendryanti, and S. Sussi, “My I-Pond: Water Quality Monitoring with IoT and Machine Learning to Reduce Pond Cultivation Failure for Farmers,” in Proc. IEEE ICSIMA, pp. 150–155, 2024.

[5] D. Dhinakaran, S. Gopalakrishnan, M. D. Manigandan, and T. P. Anish, “IoT-Based Environmental Control System for Fish Farms with Sensor Integration and Machine Learning Decision Support,” International Journal on Recent and Innovation Trends in Computing and Communication (IJRITCC), vol. 11, no. 10, 2023.

[6] Md. M. Hemal, A. Rahman, F. Islam, S. Ahmed, M. S. Kaiser, and M. E. Ahmed, “An Integrated Smart Pond Water Quality Monitoring and Fish Farming Recommendation AquaBot System,” Sensors, vol. 24, no. 11, p. 3682, 2024.

[7] P. G. Arepalli and K. J. Naik, “An IoT-Based Smart Water Quality Assessment Framework for Aqua-Ponds Management Using Dilated Spatial–Temporal Convolution Neural Network (DSTCNN),” Aquacultural Engineering, 2023.

[8] A. S. V. Devanaboina, “Real-Time Water Quality Based Predictive Alert System for Aquaculture,” International Journal for Research in Applied Science and Engineering Technology (IJRASET), 2024.

[9] D. Feng, Q. Han, L. Xu, F. Sohel, S. G. Hassan, and S. Liu, “An Ensembled Method for Predicting Dissolved Oxygen Level in Aquaculture Environment,” Ecological Informatics, 2024.

[10] Y. Ma, Q. Fang, S. Xia, and Y. Zhou, “Prediction of the Dissolved Oxygen Content in Aquaculture Based on the CNN–GRU Hybrid Neural Network,” Water, vol. 16, no. 24, p. 3547, 2024.

[11] S. M. Imran, D. G. V. Yashaswini, S. Ch., and B. Gayathri, “An Innovative Hybrid Approach to Forecasting Soluble Oxygen for Optimal Water Purification in Highly Concentrated Aquaculture,” Turkish Journal of Computer and Mathematics Education (TURCOMAT), vol. 15, no. 3, 2024.

[12] S. Santi, A. Fariza, A. Gunawan, and F. M. Humaira, “Comparison of LSTM, RNN, and GRU Methods in Predicting Water Quality in the Totality Integrated Smart Aquaculture System,” in Proc. IEEE IES, pp. 547–554, 2024.

[13] F. Chen, T. Qiu, J. Xu, J. Zhang, Y. Du, Y. Duan, Y. Zeng, L. Zhou, J. Sun, and M. Sun, "Rapid Real-Time Prediction Techniques for Ammonia and Nitrite in High-Density Shrimp Farming in Recirculating Aquaculture Systems," Fishes, vol. 9, no. 10, p. 386, 2024.

[14] V. P., S. V. K., B. D., and R. Reshma, "Predicting and Analyzing Water Quality Using Machine Learning for Smart Aquaculture," in Proc. IEEE ICSCDS, pp. 354–359, .

[15] "Random Forest Regression Based Water Quality Prediction for Smart Aquaculture," in Proc. IEEE i3CS, 2023.

[16] E. Eze, S. Kirby, J. Attridge, and T. Ajmal, "Aquaculture 4.0: Hybrid Neural Network Multivariate Water Quality Parameters Forecasting Model," Dental Science Reports, vol. 13, 2023.

[17] X. Zhou, J. Wang, Y. Liu, and Q. Duan, "Deep Learning with PID Residual Elimination Network for Time-Series Prediction of Water Quality in Aquaculture Industry," Computers and Electronics in Agriculture, vol. 212, p. 108125, 2023.

[18] S. Kanwal, M. Abdullah, S. Kumar, S. H. Arshad, M. Shahroz, D. Zhang, and D. Kumar, "An Optimal Internet of Things–Driven Intelligent Decision-Making System for Real-Time Fishpond Water Quality Monitoring and Species Survival," Sensors, vol. 24, no. 23, p. 7842, 2024.

[19] R. P. Kumar, T. P. Prabhu, J. Sowmya, and A. N. Kumar, "Water Quality Prediction for Smart Aquaculture," in Proc. IEEE ICESC, pp. 1376–1382, 2024.

[20] D. T. Nguyen, T. L. Luong, N. M. Nguyen, T. T. D. Nguyen, N. Le, C. Truong, T. L. Nguyễn, H. V. Lam, Q. K. Dinh, D. Huynh, V. Vo, and T. T. D. Nguyen, "Monitoring Water Quality Parameters in Aquaculture Using Edge Computing," Journal of Catalysis and Adsorption, vol. 12, no. 2, pp. 135–140, 2024.