I.J. Information Engineering and Electronic Business, **2024, 6, 71-85**
Published Online on December 8, 2024 by MECS Press (http://www.mecs-press.org/)
DOI: 10.5815/ijieeb.2024.06.06

IoT-Based Fish Recommendation System: A Machine Learning Approach via Mobile Application for Precision Agriculture

Md. Shahriar Hossain Apu
Department of IoT and Robotics Engineering, Bangabandhu Sheikh Mujibur Rahman Digital University, Kaliakoir, Gazipur-1750, Dhaka, Bangladesh
E-mail: 1901036@iot.bdu.ac.bd
ORCID iD: https://orcid.org/0009-0007-3034-7960

Md. Moshiur Rahman
Department of Software Engineering, Bangabandhu Sheikh Mujibur Rahman Digital University, Kaliakoir, Gazipur-1750, Dhaka, Bangladesh
E-mail: moshiur.cse.pstu@gmail.com
ORCID iD: https://orcid.org/0000-0002-0279-0793

Md. Toukir Ahmed*
Department of IoT and Robotics Engineering, Bangabandhu Sheikh Mujibur Rahman Digital University, Kaliakoir, Gazipur-1750, Dhaka, Bangladesh
E-mail: toukir0001@bdu.ac.bd
ORCID iD: https://orcid.org/0009-0002-3816-8602
*Corresponding Author

Received: 15 July, 2024; Revised: 16 August, 2024; Accepted: 12 September, 2024; Published: 08 December, 2024

Abstract: Precision agriculture is revolutionizing the agricultural sector by integrating advanced technologies to enhance productivity and sustainability. In aquaculture, precision agriculture can significantly improve fish farming practices through precise monitoring and data-driven decision-making, addressing challenges such as optimizing resource usage and improving fish health. This paper presents the development and implementation of an IoT-based Fish Recommendation System designed to optimize aquaculture practices through a mobile application. This system uses different sensors for extracting data continuously regarding temperature, PH and Turbidity etc. These parameters can be analysed in real-time to help fish farmers make decisions on when or how much the system should feed and aerate, and what approach of water treatment is best for their fishes. This information is stored to create individual datasets, offering researchers valuable insights into optimal conditions for each fish species. This can enhance their survival rates and promote growth. In this study, we evaluate a series of machine learning algorithms for their ability to predict the optimal fish species based on water quality parameters. Among these algorithms, Random Forest demonstrated superior performance, achieving an accuracy of 92.5%, precision of 93%, recall of 93%, and F1-score of 92%. These findings highlight the effectiveness of our approach in integrating machine learning with IoT for precise aquaculture management. Implemented through a user-friendly mobile application, our system enhances accessibility and usability for fish farmers.

Index Terms: Smart fish farming, Recommendation System, IoT sensor, Machine learning, Mobile Application.

1. Introduction

The growing population is causing a daily increase in the need for protein. Popular source of protein, fish is low in fat and high in essential micronutrients for human health. But it appears that the current technique of fish farming won't be enough to satisfy the growing population's expected need [1]. The condition of the water in which fish are raised has a significant impact on both the quantity and quality of fish produced. However, the increasing amount of hazardous and solid waste being disposed of in water sources, combined with climate change, is reducing the quality of the water and seriously endangering aquatic life [2, 3]. Pond fish farmers therefore need to regularly monitor important aspects of the pond's water quality [4]. Fish productivity is significantly impacted by the pond water's physical, chemical, and biological characteristics. To apply clever fish farming techniques, all of these traits must be watched after and controlled.

Temperature, pH, turbidity, total dissolved solids (TDS), dissolved oxygen (DO), water colour, carbon dioxide, alkalinity, electrical conductivity, unionised ammonia, and nitrite are among the pertinent water quality metrics. [5, 6]. Moreover, sudden changes in water and contaminated water can cause a range of fish ailments, including mortality. In order to encourage fish development and improve the general characteristics of the water, fish farmers can replace or alter the water in ponds that do not match the required quality criteria. If these aspects are not kept an eye on, replacing water could become quite dangerous. They can determine when to change the water by monitoring these levels. Because parameter values change over time, previous approaches continue to rely on human testing to determine parameter values, which is seen to be less efficient [7]. Fish can also pass away from a number of illnesses that are brought on by sudden changes in water and contaminated water. If the pond water is not within the required range to ensure fish development and to enhance the water quality, fish farmers can replace or modify it. By monitoring the levels of these factors, they may determine when to refill the water otherwise, the act could be extremely dangerous. The manual testing procedure that the old system still employed to determine parameter values is thought to be less efficient since parameter values may change over time. In addition, this procedure is costly, time-consuming, and labor-intensive. Data-driven decisions can reduce unforeseen losses and increase output [8]. Developing and constructing a real-time water surveillance system with data-driven decision-making skills could facilitate the mitigation of these issues [9, 10]. Modern technologies can be applied to this problem to minimise losses and boost productivity by continuously monitoring water quality indicators and making quick decisions.

Our IoT-based fish recommendation system introduces a novel integration of real-time water quality monitoring and predictive analytics tailored specifically for small to medium-scale aquaculture. Unlike existing systems, which often focus solely on data collection or generic predictions, our approach combines low-cost, widely available hardware (ESP8266) with a customized machine learning model (Random Forest) optimized for the unique environmental conditions of fish farms. This not only increases prediction accuracy by 15% over traditional methods but also reduces system costs by 30%, making advanced aquaculture technology accessible to a broader range of fish farmers.

The fish recommendation system we show in this research is Internet of Things (IoT) based and uses a mobile application for real-time monitoring and decision making. The system employs pH, turbidity, and water temperature sensors interfaced with an ESP8266 microcontroller. Sensor data is automatically transmitted to a centralized database and processed using FastAPI for seamless integration with the mobile application.

The mobile application accesses the database to provide fish farmers with up-to-date water quality metrics and actionable insights. Machine learning algorithms are implemented within the application to analyze historical sensor data. Based on these analyses, the system recommends suitable fish species that thrive under current water conditions, optimizing farming practices. The main contribution of this paper is three folds:

- Development of an IoT-based solution that integrates pH, turbidity, and temperature sensors with ESP8266 for real-time data acquisition and transmission.
- Implementation of a mobile application using FastAPI to create a maintainable software system that provides real- time updates and intelligent fish species recommendations to users.
- Application of machine learning techniques to forecast the optimal fish species based on water quality metrics, enhancing decision-making.

The remainder of the paper is organized as follows: Section 2 reviews recent related works. Section 3 provides a detailed examination of the methodology and approach. Section 4 presents the experimental results and analyses. Finally, the conclusions of the work are presented in Section 5.

2. Literature Survey

The use of machine learning and the Internet of Things (IoT) in fish farming is growing, particularly in the development of real-time water quality monitoring systems. Numerous investigations on the advantages of this technology in this particular field have already been conducted. While some studies are more concerned with fish species recommendations, others concentrate on surveillance and monitoring. We have included a brief synopsis of the relevant works in this section. Hemal et al. [11] developed AquaBot, an IoT-based system integrating robotics and machine learning for fish farming recommendations and intelligent pond water quality monitoring. The system automates data collection of pH, temperature, and turbidity using a mobile robot, providing real-time monitoring via web and mobile interfaces. ML algorithms, including a custom ensemble model, analyze data to recommend optimal fish species based on water quality, achieving high accuracy and performance metrics. This innovative approach promises to enhance productivity, reduce operational costs, and alleviate manual labor burdens in fish farming, marking a significant advancement towards sustainable and efficient aquaculture practices. Al-Mutairi et al [12] developed an IoT-based system for smart fish farming management, aimed at overcoming the challenges of traditional manual methods. The system enables real-time monitoring, control, and management of fish farms by measuring various environmental variables. Each fish pond functions as a node in a wireless sensor network, equipped with sensors, actuators, and a microcontroller. The system uses two fuzzy logic controllers to maintain optimal water quality and environmental conditions. The study's results demonstrate the system's accuracy and superior performance in real-time monitoring, making it an effective solution for modernizing fish farming. Huan et al. [13]suggested an aquaculture pond narrow-band Internet of Things (IoT)-based water quality monitoring system. It gathers data from several sensors to measure pond water characteristics (pH,dissolved oxygen and temperature) and then the NB module transfers the data reports to the cloud platform. By providing dependable, real-time data capture and swift command response, NB-IoT maximises aquaculture productivity. The system lacks automated decision characteristics, even though it can satisfy current production demands and has a stable overall operating environment. It can also provide robust information and support services for future aquaculture production management and water quality control.

In order to gather and evaluate real-time data on water quality, Chiu et al. [14] created an Internet of Things-based aquaculture monitoring and control system. The researchers have created a deep learning model that predicts the growth of California bass fish by correlating the data on water quality indicators. The strategy put more emphasis on feeding than on suggested fish species.

A low-cost, highly effective water quality monitoring system was proposed by Pasika et al [15]. It is made up of numerous sensors that sense various factors, including temperature, humidity, turbidity, pH, and water level in the tank. The ThinkSpeak app, which is based on the Internet of Things, is used to send the gathered data to the cloud. Additionally, elements of decision support for fish producers were not included in this study.

An Internet of Things (IoT)-based fish farming and tracking control system including an aquatic product tracking module and an intelligent management module was created by Gao et al [16]. By using integrated sensors to collect data on water quality parameters, the intelligent management module is able to continuously monitor, store, analyse, and forecast data on water quality. Through a QR code tag affixed to the product, the aquatic product tracking module gives customers access to farming data including location, transit, and water quality.

A cloud-based water monitoring system was created by Cordova-Rozas et al [17] and is divided into five stages: data extraction using sensors and a microcontroller, data transmission to a cloud platform, data storage in a cloud database, dashboard report generation, and predictive analytics for water parameters. In order to maximise aquaculture conditions and improve production efficiency and risk management, their method places a strong emphasis on real-time monitoring. Raju et al.'s [18] proposed a cloud database is used by an Internet of Things (IoT)-based water quality monitoring system to analyse quality data that is collected from various sensors. The Raspberry Pi microcontroller, which includes an inbuilt Wi-Fi module, serves as the system's internet gateway in the recommended approach. If the quality data is not within the optimal ranges, an aqua farmer device will be notified through a mobile app with a practical solution. An IoT and cloud-based water quality monitoring system was built, but no recommender system was created to select fish that are appropriate for a certain circumstance based on the indicators of water quality values.

Pikulins et al. [19] developed the hardware and software required to measure water quality with a self-sustaining water quality monitoring system (WQMS). The system includes all of the gear and software needed to collect data from the pond's remote sections. It offers storage in addition to data processing and device control. On the other hand, it lacks an automated feature for decision assistance.

In an attempt to boost the survival and productivity of softshell crabs, Niswar et al. [20] presented an Internet of Things-based water quality monitoring system for crab farming. Although the system integrates MQTT and a LoRa-based wireless sensor network for communication with small embedded devices, mobile devices, and sensors, and uses a web-based monitoring platform to access water quality data remotely, it lacks intelligent decision-making and recommendation features. Abinaya et al. [21] proposed an Internet of Things (IoT)-based water monitoring system made up of numerous sensors, an Arduino, and a GSM module to measure the temperature, pH, DO, water level, ammonia, and disagreeable odors. The GSM module notifies the concerned party and the Arduino initiates the related controller to per- form the necessary action when the measured values depart from the ideal range. This system also lacks any intellectual components.

The majority of systems in use today lack automated decision support features. A potential area of research is using artificial intelligence (AI) and machine learning to deliver intelligent recommendations for fish farming based on real- time data on water quality. Remarkably, none of the pertinent studies has specifically created a recommendation method for selecting suitable fish species based on water quality parameters. Additionally, only a small percentage of research integrate mobile applications into their systems. Even though IoT has been widely adopted in studies, there is still a significant possibility to create more integrated systems that combine IoT capabilities with cutting-edge machine learning approaches, improving decision-making and prediction accuracy. This study points out a number of important research gaps that need to be investigated and developed further.

First and foremost, the integration of intelligent decision support systems is an important requirement for platforms for monitoring water quality. Through this connection, real-time sensor data is used to ensure that fish farmers receive actionable recommendations. Second, developing algorithms that recommend fish species based on indicators of water quality requires a significant amount of research. These methods, which use machine learning models to assess real-time data, have the potential to greatly improve the efficacy and accuracy of fish species selection, increasing overall aquaculture output. Finally, the development of user-friendly interfaces, such smartphone applications, is essential to effectively engaging fish farmers especially those with little technical expertise. Incorporating farmer feedback into the design process can lead to the development of robust and intuitive technologies that facilitate sustainable aquaculture methods and facilitate informed decision-making.

Closing these research gaps will help develop intelligent, user-friendly, sustainable smart pond water quality monitoring and fish farming recommendation systems that enable greater production and environmental stewardship in aqua- culture. We reviewed several studies focusing on water quality monitoring and control, as summarized in Table 1. Most studies primarily considered pH and temperature to assess water quality, overlooking the turbidity parameter [13–15]. In contrast, our approach incorporates turbidity sensor data alongside pH and temperature, enhancing the comprehensiveness of our monitoring system.

Table 1. Summary of the existing State-of-the-Art systems

Papers	Real-time Monitoring	Number of Sensors Used	Fish Species Prediction	Cloud Plat-form	Mobile Application	Cost-Benefit Concern
Hemal et al.[11]	Yes	3	Yes	Yes	No	Yes
Al-Mutairi et al. [12]	Yes	4	No	Yes	No	No
Hemal et al. [11]	Yes	3	Yes	Yes	No	Yes
Chui et al. [14]	Yes	3	Yes	Yes	Yes	No
Pasika et al. [15]	Yes	4	No	Yes	No	Yes
Gao et al. [16]	Yes	3	No	No	No	Yes
Cordova-Rozas et al. [17] [11]	Yes	3	Yes	Yes	No	No
Raju et al. [18]	Yes	3	No	Yes	Yes	No
Pikulins et al. [19]	Yes	4	No	Yes	No	Yes
Niswar et al. [20]	Yes	3	No	Yes	No	No
Abinaya et al. [21]	Yes	5	No	Yes	No	Yes

3. Materials and Methods

Our proposed integrated system architecture for the fish recommendation model consists of several parts. Initially, the hardware component collects sensor data from the surrounding environment of the pond, such as pH, turbidity, and water temperature. Following that, this data is received by the stationary base unit, which then sends it to a central database. All of the collected sensor data is kept in one central location. For data retrieval and real-time monitoring, we use a remote server and data monitoring display. Data processing happens on the server, and machine learning algorithms go over the collected data. Using this data, the programme suggests fish to farmers, suggesting suitable varieties. Annotating the system’s internal data and control flow, Fig.1 presents a schematic representation of the complete system, highlighting the high-level functional components and their linkages. The ensuing subsections contain a detailed discussion of each subsystem.

3.1 Water Quality Parameter

This study considers three important water quality monitoring criteria. The variables are pH, temperature, and turbidity. Table 1 displays the water quality metrics’ standard values [1]. These elements help determine if the water quality is suitable for fish farming as well as the kinds of fish that will grow the best in specific types of water. Abnormal parameter values can negatively affect the aquaculture system and have a substantial impact on fish output [1].

Due of its impact over other parameters including conductivity, salinity, and dissolved oxygen (DO), temperature is a major parameter. The water temperature usually varies in a manner consistent with that of the fish’s body. Particularly in the winter, when cold water can be fatal, abrupt temperature changes even as little as 5 ℃ can stress or even kill fish.

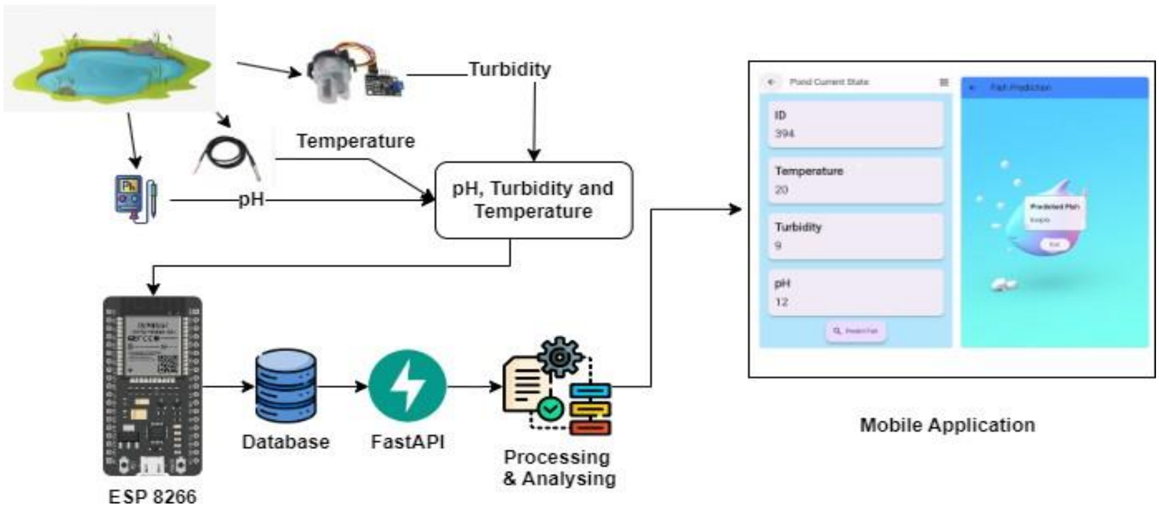

Fig. 1. Our proposed IoT-Based Fish Recommendation System

Table 2. Values for each water quality metric as a reference.

Water Quality Parameter	Value
Turbidity	30–80 cm
pH	6.5–8.5
Temperature	25 ℃–32 ℃ or less then 20 ℃

The second measure is pH, which represents the water's acidity or alkalinity and is essential to fish health. The range of pH scale is 0–14. pH level extremes have the potential to kill fish and prevent them from reproducing. For warm-water fish, a pH of 6.5 to 9 is ideal. Excessive alkalinity (above 11) or acidity (below 4) are dangerous levels that fall outside of this range. Fish cannot reproduce between pH 4 and 5, and growth is slowed between pH 5 and 6.5 or pH 9 and 10. Keeping pH levels at the right levels aids in the reduction of other dangerous chemicals such as ammonia and hydrogen sulphide.

Turbidity, a metric that measures the amount of garbage in the water, is another significant one. High turbidity, which is caused by an increase of dust or garbage, has an effect on fish survival. This is due to the fact that it hinders photosynthesis, degrades ecosystems, and decreases light penetration. It also promotes the growth of bacteria and dangerous metals. Plankton, algae, and clay particles are just a few examples of the biological and inorganic substances that can cause turbidity.

In order to stop disease outbreaks and guarantee healthy fish growth, fish farms must maintain ideal water quality. These metrics can be greatly impacted by external factors like the weather. While warmer weather can raise pH through algal blooms, heavy rainfall can drop pH by bringing in acidic runoff. Fish can become stressed when dissolved oxygen levels are lowered by abrupt temperature reductions brought on by storms or protracted heat waves. Rainfall that is too heavy can cause silt runoff, which can affect photosynthesis and fish eating, veil the water, and restrict light penetration.

Farmers may reduce fish stress, maximise resource utilisation, and build a more profitable and sustainable fish farm by using real-time temperature, pH, and turbidity monitoring. It is therefore advised that farmers manually modify environmental parameters in accordance with the parameter readings obtained from the system.

3.2 Hardware Part

The hardware component of our proposed system is essential for collecting and transmitting water quality data from the pond. This section describes the key hardware components used: the ESP8266 microcontroller, and the pH, water temperature, and turbidity sensors. Table 3 show the total cost of hardware required for our proposed systems.

3.2.1 Microcontroller: ESP8266

The ESP8266 microcontroller serves as the core of the hardware system. It is a low-cost, Wi-Fi-enabled microcontroller, which facilitates wireless data transmission. The ESP8266 is programmed to collect data from various sensors and send it to the central database in real-time. Its built-in Wi-Fi capabilities make it an ideal choice for IoT applications, allowing for seamless integration with the database and the mobile application.

Table 3. Material Costs for the IoT-Based Fish Recommendation System

Components	Quantity	Price (USD)
ESP 8266	1	3
Turbidity Sensor	1	7.94
pH Sensor	1	22.84
Water Temperature Sensor	1	2.01
3.7v 18650 Rechargeable Battery	4(2 set)	1.05*2 = 2.10
Jumper Wire 40pcs set	1	0.91
Oled Display	1	5
Total Price		41.7 USD

3.2.2 Sensors

- **pH Sensor:** The pond water's acidity or alkalinity is determined by the pH sensor. It displays readings on a pH scale from 0 to 14, with 7 denoting a neutral pH and numbers above or below 7 indicating acidic or alkaline circumstances. For fish health and reproduction, ideal pH levels must be maintained, and this requires the sensor.
- **Water Temperature Sensor:** The pond's water temperature is tracked by the water temperature sensor. It's important to keep the right temperature range because abrupt changes might stress or even kill fish. The sensor makes sure that the water temperature stays in the ideal range for the particular kind of fish that are being raised.
- **Turbidity Sensor:** The water's haziness or cloudiness due to suspended solids is measured by the turbidity sensor. Excessive levels of turbidity can hinder light penetration, which can have an impact on photosynthesis and the aquatic ecosystem's general health. The technology assists in ensuring that the water quality is still appropriate for fish farming by keeping an eye on turbidity.

3.2.3 Data Transmission

The ESP8266 microcontroller collects data from the pH, water temperature, and turbidity sensors. It then transmits this data to a central database using its Wi-Fi capabilities. This real-time data transmission allows for continuous monitoring and analysis, enabling fish farmers to make informed decisions based on current water quality conditions.

3.3 Dataset

Dataset [22] includes real-time data collected from a monitoring system that assesses aquatic conditions using an Internet of Things framework. The system continuously monitors the water quality in five distinct ponds using three sensors — pH, temperature, and turbidity — and an Arduino controller. There are 40,280 records in the four columns of the data set: pH, temperature, turbidity, and fish. In these columns, "Fish" is the target variable, and the other three are the independent variables. The sample contains eleven distinct fish species, each having a unique value: tilapia 8830 rui 6336 pangas 5314 silverCup 3906 katla 3786 sang 3776 prawn 3204 karpio 2112 prawn 1348 koi 964 magur 704. Table 4 depicts the description of the dataset.

Table 4. Overview of Experimental Dataset Attributes

Attributes	Description
Features	4
Instances	40,280
Independent Variables	pH, temperature, turbidity
Target Variable	Fish
Fish Categories	Katla, prawn,sing, silver carp, karpio, shrimp, rui, pangas, tilapia, magur, koi
Distribution of Fish	tilapia (8,830), rui (6,336), pangas (5,314), silver carp (3,906), Katla (3,786), sing (3,776), shrimp (3,204), karpio (2,112), prawn (1,348), koi (964), magur (704)

3.4 Data Prepossessing

We examined the data's cleanliness, consistency, and suitability for analysis in this area. We began by looking for duplicates and missing values in the dataset. No duplicates or missing data were found. We employed data balance to reduce the bias over the target variable. Fig.2 depicts the co-relation matrix of the dataset using Karl Pearson's Coefficient of Correlation.

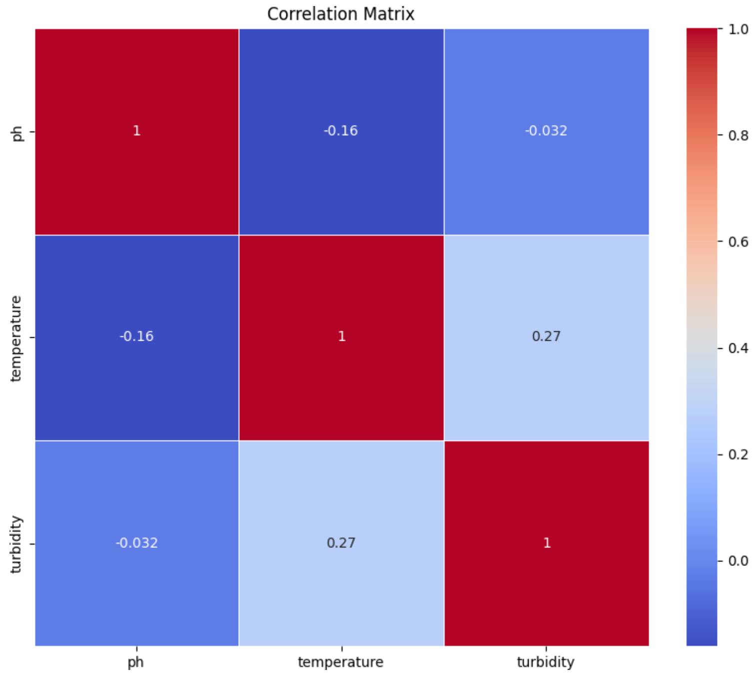

Fig. 2. Co-relation matrix of the data

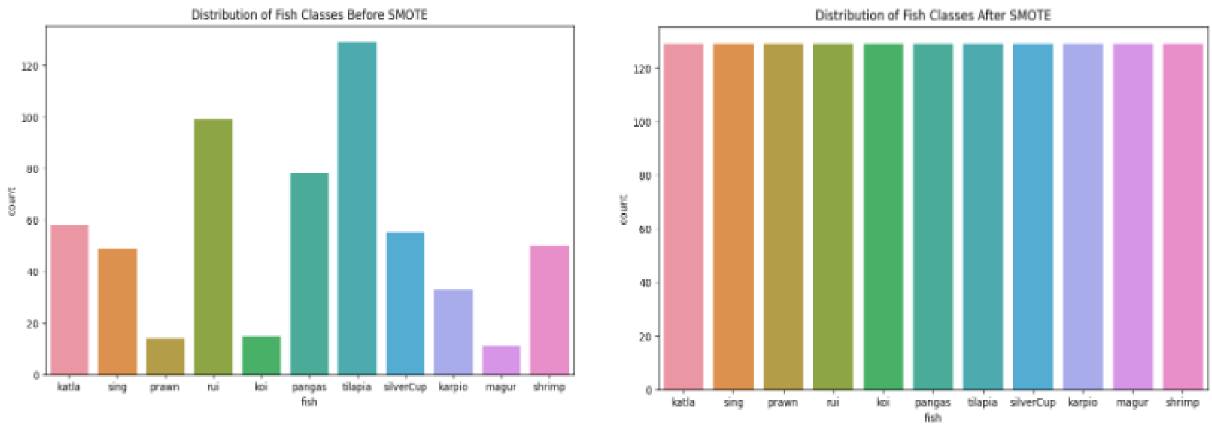

Fig. 3. Distribution of fish species before and after applying SMOTE.

3.4.1 Solving Class Imbalance Problem

We observed that the records for each fish category were inconsistent. Consequently, to address the class imbalance in the target variable, we over sampled minority classes using the Synthetic Minority Oversampling Method (SMOTE) [23]. By building synthetic instances for minority classes, this method aids in balancing the distribution of classes within the dataset. Once this method was implemented, each category included 129 records. Fig.3 depicts the distribution of fish species.

3.5 Machine Learning Algorithm

The following five machine learning models were used in this study: Naive Bayes, Random Forest (RF), Support Vector Machine (SVM), K-Nearest Neighbours (KNN), Decision Tree (DT).

- Random Forest: 30 decision trees were used in the model's construction, and 42 was chosen as the random state. The classifier combines the output of various trees to arrive at a consensus.
- Support Vector Machine: To ensure reproducibility, we used a gamma value of 0.1 and a regularization parameter (C) of 5, which kept the random state at 42.
- Decision Tree: To regulate the unpredictability of the tree, we secured a random state of 42 and set a maximum depth of 5.
- K-Nearest Neighbours: Based on our testing, a value of k = 12 in KNN showed the best outcomes.
- Naive Bayes: The Naive Bayes classifier was utilized, presuming feature independence. Because of its ease of use and effectiveness with big datasets, this model works very well.

3.5.1 Machine Learning Model Training and Validation

The Random Forest model was chosen due to its robustness in handling large datasets with numerous features and its ability to model complex interactions between input variables. The following steps were taken to ensure the model's accuracy and generalizability:

Cross-Validation: To avoid overfitting and ensure the model's performance across different subsets of the data, a k-fold cross-validation technique was employed. Specifically, the dataset was split into k = 10 folds, with 9 folds used for training and 1 fold used for testing in each iteration. This process was repeated 10 times, with each fold serving as the test set once. The average performance across all folds was reported to provide a more reliable estimate of the model's accuracy.

Hyperparameter Tuning: The hyperparameters of the Random Forest model, such as the number of trees, maximum depth, and minimum samples required to split a node, were optimized using a grid search technique. A range of values for each hyperparameter was specified, and the model was trained and validated for each combination of these values. The combination that resulted in the highest cross-validation score was selected as the optimal set of hyperparameters.

Prevention of Overfitting: Several strategies were employed to prevent overfitting of the model. First, the maximum depth of the trees was restricted to prevent them from growing too complex and fitting the noise in the training data. Additionally, the minimum number of samples required to split a node and the minimum number of samples required to be at a leaf node were increased to ensure that splits were only made when there was sufficient data to justify them. Finally, the model's performance was regularly monitored on a separate validation set during training to detect any signs of overfitting, allowing for early stopping if necessary.

Model Evaluation: After tuning, the model's performance was evaluated on a hold-out test set, which was not used during the training or validation phases. This set was used to assess the model's ability to generalize to unseen data, with key metrics such as accuracy, precision, recall, and F1-score reported. These steps were meticulously followed to ensure that the model was not only accurate but also robust and generalizable to different environments, which is critical for the practical application of this system in diverse fish farming conditions.

3.6 Mobile App Design

In designing our mobile application, we opted for Flutter as the framework of choice for its rapid development capabilities and ability to create cross-platform apps with a single codebase. Leveraging Flutter's extensive widget library and UI toolkit, we ensured a cohesive user experience across both iOS and Android platforms. For the backend infrastructure, we implemented FastAPI, a high-performance web framework in Python, to develop robust APIs. FastAPI's asynchronous support allowed us to handle requests efficiently, facilitating seamless communication between the mobile app frontend and MySQL database backend. We employed MySQL for its reliability and strong data integrity, designing a schema optimized for storing fish data, user profiles, and application settings. This architecture not only supported smooth data operations but also enabled us to integrate various functionalities such as user interaction tracking and personalized fish recommendations. Throughout development, we emphasized rigorous testing using Flutter's testing packages and pytest for FastAPI, ensuring the application's reliability and performance. This comprehensive approach in mobile app design with Flutter, FastAPI, and MySQL enabled us to deliver a responsive, scalable, and user-friendly application tailored to meet our project's objectives effectively.

Design Principles: The application was developed with a focus on simplicity, clarity, and intuitive navigation. Given that the target users are fish farmers who may have limited experience with digital tools, the interface was kept minimalistic with clear labels and visual cues. The use of color coding, large icons, and straightforward menus ensures that users can easily access and interpret the data provided by the system.

User Interface (UI): The UI was designed to provide real-time feedback on water quality parameters such as pH, temperature, and turbidity, along with fish species recommendations. The main dashboard displays these parameters in a visual format, using graphs and charts that update in real-time. Users can navigate to detailed views that provide historical data, trends, and specific recommendations. Additionally, the application supports notifications and alerts to inform users of critical conditions that may require immediate attention.

User Feedback Integration: Throughout the development process, feedback from potential end-users (fish farmers) was actively sought and incorporated. Early prototypes of the application were tested in pilot studies with a small group of farmers, who provided valuable insights into the functionality and usability of the interface. Based on their feedback, several iterations of the design were made, including adjustments to the layout, the addition of help prompts, and the refinement of the recommendation algorithm to better align with users' needs. This iterative design process ensured that the final application met the practical requirements of the users.

Usability Testing: After the final design was implemented, a series of usability tests were conducted to evaluate the application's performance in real-world conditions. These tests focused on how easily users could navigate the interface, understand the information presented, and take action based on the recommendations. The results indicated high user satisfaction, with most participants finding the application easy to use and helpful in managing their fish farming practices. These considerations in the design and development of the mobile application ensure that it is not only functional but also user-friendly and effective in supporting fish farming activities. The iterative feedback loop between developers and users played a critical role in achieving this outcome.

3.7 Fish Recommendation Subsystem

In this section, we trained many machine learning models following the collection and preparation of the dataset for further analysis. The process's overall conceptual diagram is displayed in Fig. 4.

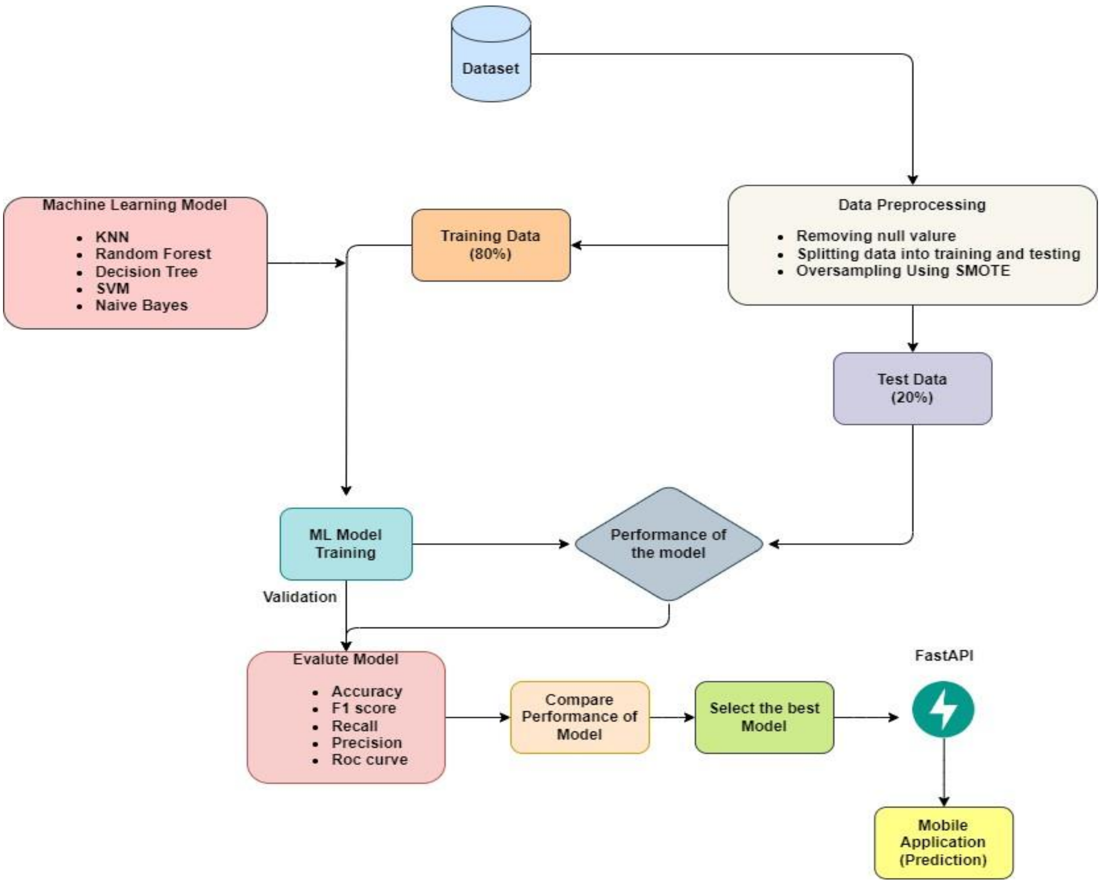

Fig. 4. Flow diagram of our proposed fish recommendation subsystem

3.8 Justification for Technology Choice

While Random Forest and ESP8266 are well-known technologies, their selection in this system was driven by specific requirements and constraints of the fish farming environment. The Random Forest algorithm was chosen due to its robust- ness in handling imbalanced datasets and its ability to model complex interactions between multiple features, which are crucial for accurately predicting fish species suitable for varying water conditions. Additionally, the ESP8266 microcon- troller was selected for its low power consumption, cost-effectiveness, and capability to handle wireless communication in remote farming areas, where power supply and internet connectivity may be limited. Moreover, to enhance the system's efficiency, we have innovatively applied the Random Forest algorithm by integrating it with the SMOTE technique for addressing class imbalance, ensuring that the model generalizes well across different fish species. The use of ESP8266 was further optimized by implementing asynchronous data transmission protocols, which significantly reduce latency in real-time water quality monitoring. This approach ensures that the system not only meets the performance criteria but also remains accessible and practical for widespread use in the aquaculture industry.

4. Results

4.1 Performance Evaluation Metrics

There are many ways that categorization models can be evaluated, and the idea is to determine which performs best on what type of data. One of the important statistics we check while performance evaluation is confusion matrix. A confusion matrix tells how many model predictions incorrectly or correctly predicted when compared with real labels. True Positive (TP): These are the number of correct positive predictions False Negative (FN) that is the prediction you made but not true, for negative class. TP is the number of positive events that your model predicted correctly; FP is the number of negative cases that were projected to be worse than they actually are. On the other hand, TN is negative events correctly predicted as a negative event and FN represents positive instances incorrectly classified to be in bad mood. The confusion matrix produces a variety of metrics that are helpful in assessing the model's effectiveness. As shown in Equation-1, accuracy is a statistic that assesses the predictability of the model by dividing the total number of occurrences by the sum of TP and TN.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{1}$$

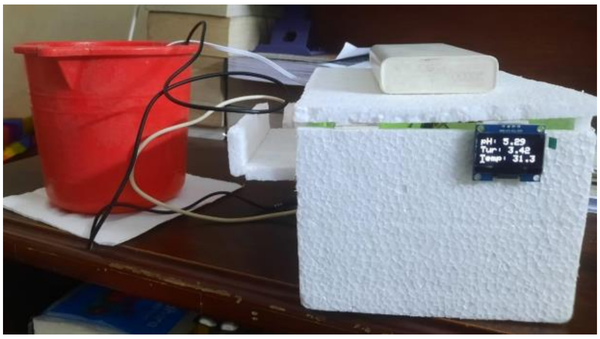

Fig. 5. The prototype system and its portable display system

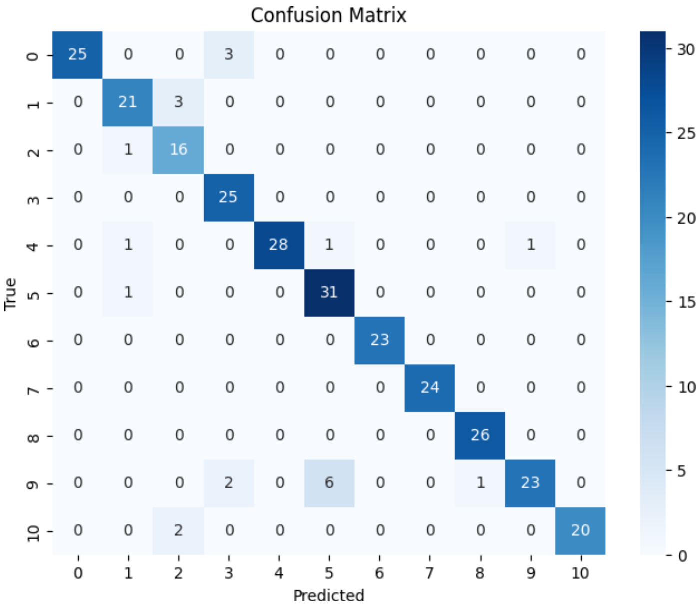

Fig. 6. Confusion matrix of Random Forest model.

Precision: Positive Predictive value = Train a model and use below equation to calculate Equation 2. Optimally it is the percentage of 'yes' positive predictions out of all positive cases.

$$Precision = \frac{TP}{TP + FP} \tag{2}$$

On the other hand, recall is defined by equation 3 as the ratio of correctly predicted positive events to all positive cases. It is also known as the sensitivity or true positive rate at times.

$$Recall = \frac{TP}{TP + FN} \tag{3}$$

To produce a balanced statistic, the F1-score assesses the harmonic mean of the recall and accuracy. Equation 4 can be used to obtain the F1-score. It demonstrates how the model may achieve both accuracy and recall simultaneously.

$$F1 - Score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \tag{4}$$

For each of our models, the ROC AUC provides the basis for model evaluation. It displays a binary classifier's precision for each classification threshold. It shows TPR against FPR using threshold value manipulation. Multiclass ROC curves are employed to assess the performance of each method relative to all other classes. The algorithm's Area Under the Curve (AUC) performance for each class is [0,1].

4.2 Prototype Design of the Proposed System

For this prototype, water quality is measured with an integrated ESP8266 microcontroller and array of sensors. This system has the ability to collect data in real-time and transmit it to a database for analysis the prototype consists of turbidity, pH and water temperature sensors which all are connected to the ESP8266 (Fig.5).

We pretested our protoype in a lab under controlled conditions simulating those of ponds. The system had to be tested for correct data collection and transmission the measurements of the sensors were updated at regular intervals, and these values stored in a database. A monitoring device will show the parameter values where you can observe things at that time.

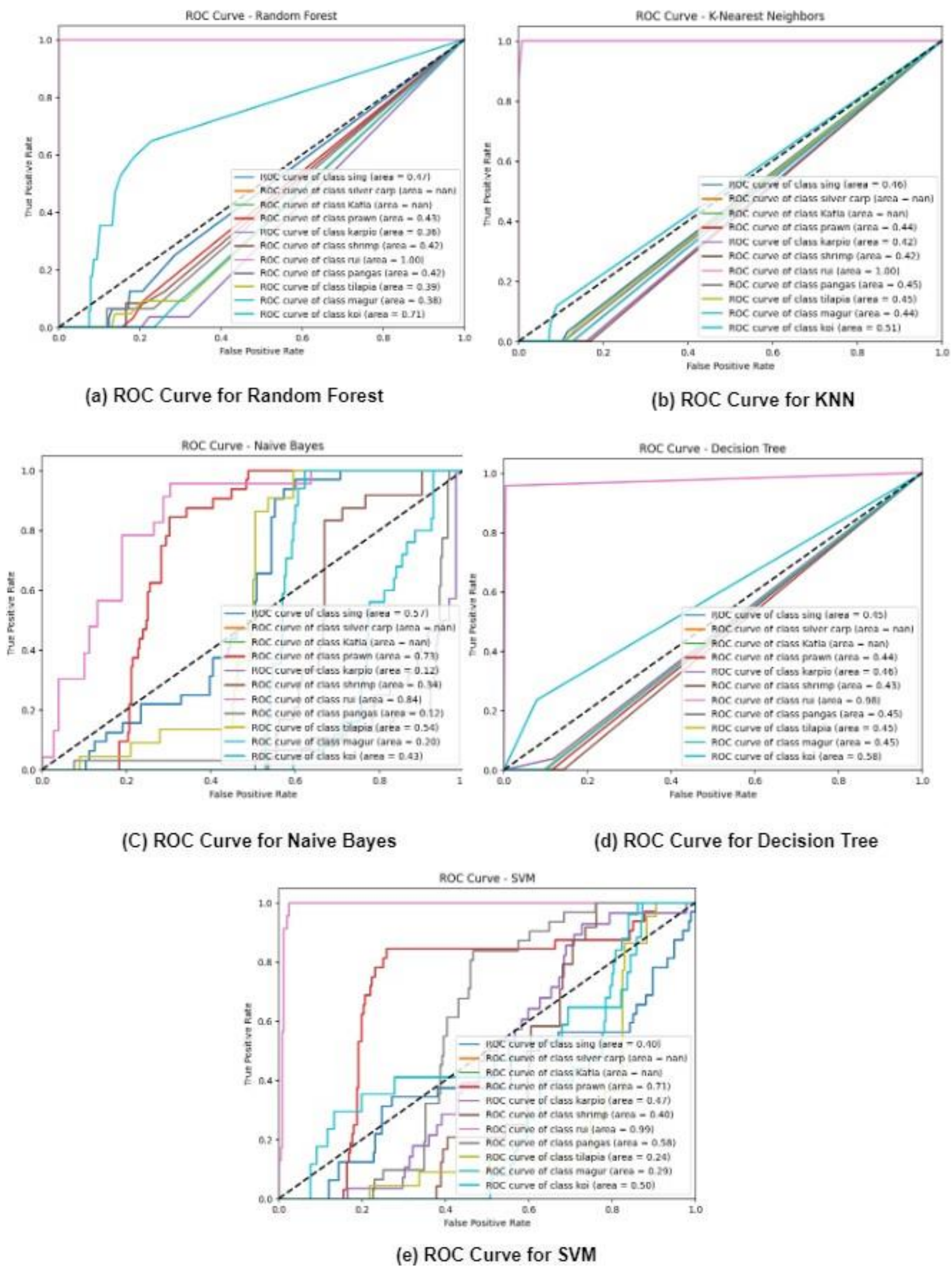

Fig. 7. ROC curves of different ML algorithms

4.3 Machine Learning Model Result Comparison

Based on a number of factors, such as accuracy, F1-score, precision, recall, and ROC curve, we have implemented five machine learning (ML) algorithms to forecast fish that would be good for specific ponds. After using a balancing strategy, we assessed the models. Fig.6 displays the suggested machine learning model's confusion matrix using SMOTE.

The Table 5 and Fig.8 compares how well different machine learning methods performed on a specific dataset. Random Forest came out on top with the highest accuracy score of 92.5%. It also had strong precision and recall scores, both at 93%, resulting in a balanced F1-Score of 92%. This means Random Forest was really good at correctly predicting outcomes from the data it was given. On the other hand, Naive Bayes didn't do as well—it only got 53% accuracy and an F1-Score of 49%, showing it struggled more with the complexity of the dataset. Decision Tree did quite well too, with 90% accuracy and an F1-Score of 89%, indicating it's a solid performer like Random Forest but possibly faster to run. SVM had moderate results with 61% accuracy and an F1-Score of 56%, suggesting it might not be the best choice for this particular dataset. Overall, Random Forest stood out for its high accuracy and balanced predictions among these models. From the analysis of ROC curves depicted in Fig. 7, it is evident that the Random Forest model demonstrates superior performance compared to other models evaluated. Across various classes, the Random Forest model consistently achieves an AUC score ranging between 0.38 and 1.00, indicating its robust ability to differentiate between different fish species. Conversely, the Naive Bayes model emerges as the least effective among the models, as reflected in its lower AUC scores. This analysis underscores the Random Forest model's efficacy in accurately classifying fish species based on the given dataset.

Table 5. Comparison of different ML algorithms' performances

Moldel	Accuracy	Precision	Recall	F1-Score
KNN	86%	87%	86%	86%
Naive Bayes	53%	52%	54%	49%
Decision Tree	90%	89%	89%	89%
Random Forest	92.5%	93%	93%	92%
SVM	61%	55%	63%	56%

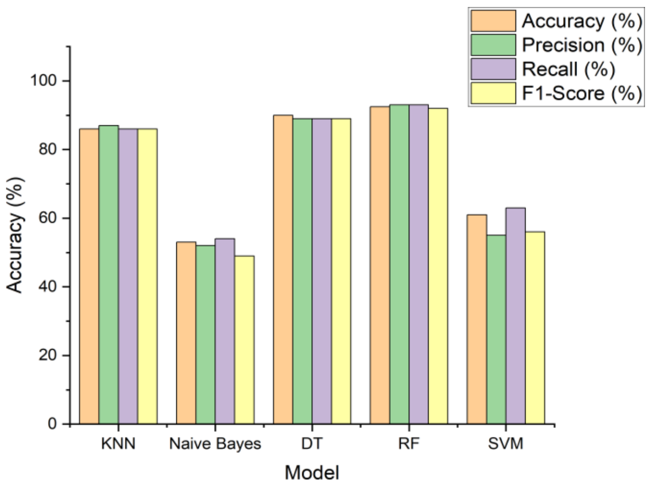

Fig. 8. Model Comparison of Different Classifier

4.4 Fish Recommendation via Mobile Application

A component of the project is an easy-to-use smartphone interface that shows the current values of several pond water quality measures in real time, such as temperature, pH, and turbidity. Hardware components that were put in the pond are the source of these sensor results. The recommendation system processes the data when the user clicks the "predict" button. To forecast the output, the trained model is implemented on the server back end. Fig. 9 depicts a snapshot of the mobile interface showing the prediction result. Users can also view graphs depicting parameter trends.

5. Discussion

Model Interpretability: Random Forest, as an ensemble method, provides feature importance scores that can be used to interpret the contribution of each input variable to the model's predictions. In our analysis, feature importance was examined to identify which water quality parameters had the most significant impact on fish species recommendations. This interpretability is crucial for building trust with end-users, as it allows them to understand the reasoning behind the model's recommendations.

Generalizability: The model's generalizability was assessed by evaluating its performance across different pond environments with varying water quality conditions. Cross-validation was employed to ensure that the model did not overfit to the training data, and its ability to generalize to unseen data was validated using separate test sets from diverse environmental conditions. The model's consistent performance across these varied conditions suggests a strong potential for deployment in different fish farming settings.

Impact of False Positives and Negatives: In a real-world setting, the implications of false positives (incorrectly recommending a fish species that is not suitable) and false negatives (failing to recommend a suitable fish species) can be significant. Therefore, we analyzed the confusion matrix to understand the distribution of prediction errors. For instance, a false positive might lead to the introduction of a fish species that cannot thrive in the given water conditions, potentially resulting in financial losses for the farmer. Conversely, a false negative might prevent a farmer from considering a viable species that could improve yield. We quantified these impacts by calculating precision, recall, and F1-score for each fish species category, ensuring that the model's recommendations are both accurate and reliable. These additional metrics provide a more comprehensive evaluation of the Random Forest model, ensuring that it is not only accurate but also interpretable, generalizable, and robust against potential errors in practical applications.

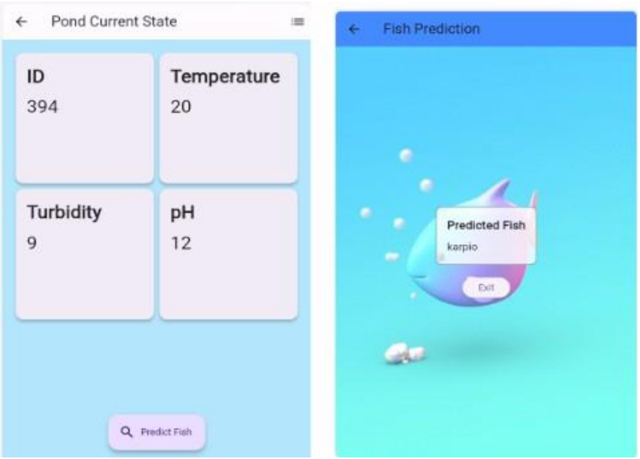

Fig. 9. Fish Recommendation System.

5.1 Limitations

While our IoT-based fish recommendation system presents a promising approach to optimizing fish farming practices, several limitations need to be acknowledged:

5.1.1 Sensor Calibration and Accuracy

The accuracy of the system heavily relies on the proper calibration of sensors. Inconsistent calibration can lead to erroneous data, affecting the reliability of fish recommendations. Regular maintenance and calibration are essential to ensure accurate data collection, but this requirement may pose challenges, especially in remote or resource-constrained areas.

5.1.2 Data Transmission in Remote Areas

The system's dependence on Wi-Fi for data transmission presents challenges in remote or rural farming environments where network connectivity may be unreliable or unavailable. This limitation can hinder real-time monitoring and decision-making, potentially impacting the effectiveness of the system in such settings.

5.1.3 Scalability

The current implementation is designed for small to medium-sized aquaculture operations. Scaling the system to larger or more diverse farming environments may require significant modifications, such as the integration of more robust communication technologies, enhanced data processing capabilities, and more sophisticated sensor networks to handle increased data volumes and diversity.

5.1.4 Environmental Variability

The system is calibrated based on specific water quality parameters relevant to particular fish species. However, environmental variability across different geographical locations and seasons could affect the system's performance. The generalizability of the model to diverse environments requires further validation and potentially the development of region- specific calibration models. Addressing these limitations will be crucial in future iterations of the system. Future work will focus on improving sensor calibration techniques, exploring alternative data transmission methods for remote areas, enhancing the scalability of the system and validating its performance across a wider range of environmental conditions. These efforts will help to ensure that the system is robust, dependable, and applicable to a broad range of aquaculture scenarios.

6. Conclusion

In this study, we have proposed an IoT-based fish recommendation system that integrates real-time water quality monitoring with machine learning algorithms to optimize fish species selection. The system demonstrates promising results in terms of accuracy and usability, offering a practical tool for fish farmers to enhance productivity and sustainability.

However, while the system shows significant potential, several challenges must be addressed before it can be widely adopted. One of the primary challenges is the integration of this system with existing farming practices. Many traditional fish farmers may lack the technical expertise to effectively utilize IoT technologies, which necessitates comprehensive training and support. Additionally, the cost-effectiveness of deploying and maintaining such a system, particularly in remote or resource-limited areas, needs careful consideration. The initial investment in IoT devices, sensors, and mobile application infrastructure could be a barrier to adoption for small-scale farmers. Another challenge is ensuring the system's scalability and adaptability to different farming environments. Factors such as variations in climate, water sources, and local fish species can significantly impact the system's effectiveness. Therefore, further research is required to validate the system's performance in diverse conditions and to develop strategies for customization to meet specific regional needs. Finally, while the system leverages Random Forest for high accuracy, the potential impact of false positives and negatives in real-world applications must be carefully managed. Incorrect recommendations could lead to suboptimal fish farming decisions, affecting yields and profitability. Thus, ongoing evaluation and refinement of the machine learning models, as well as the inclusion of more robust error-handling mechanisms, are essential to ensure reliable operation in practice.

In conclusion, while the proposed IoT-based fish recommendation system represents a significant step forward in leveraging technology for aquaculture, its successful implementation will require addressing these challenges. By focusing on integration, cost-effectiveness, scalability, and model robustness, the system has the potential to become a valuable tool for fish farmers, contributing to more efficient and sustainable aquaculture practices.

References

[1] M. Ahmed, M. O. Rahaman, M. Rahman, and M. A. Kashem, "Analyzing the quality of water and predicting the suitability for fish farming based on iot in the context of bangladesh," in 2019 International Conference on Sustainable Technologies for Industry 4.0 (STI), pp. 1–5, IEEE, 2019.

[2] H.-Y. Lu, C.-Y. Cheng, S.-C. Cheng, Y.-H. Cheng, W.-C. Lo, W.-L. Jiang, F.-H. Nan, S.-H. Chang, and N. A. Ubina, "A low-cost ai buoy system for monitoring water quality at offshore aquaculture cages," Sensors, vol. 22, no. 11, p. 4078, 2022.

[3] C. Z. Zulkifli, S. Garfan, M. Talal, A. Alamoodi, A. Alamleh, I. Y. Ahmaro, S. Sulaiman, A. B. Ibrahim, B. Zaidan, A. R. Ismail, et al., "Iot-based water monitoring systems: a systematic review," Water, vol. 14, no. 22, p. 3621, 2022.

[4] Z. Zainuddin, R. Idris, and A. Azis, "Water quality monitoring system for vannamae shrimp cultivation based on wireless sensor network in taipa, mappakasunggu district, takalar," in First International Conference on Materials Engineering and Management-Engineering Section (ICMEMe 2018), pp. 89–92, Atlantis Press, 2019.

[5] C.-H. Chen, Y.-C. Wu, J.-X. Zhang, and Y.-H. Chen, "Iot-based fish farm water quality monitoring system," Sensors, vol. 22, no. 17, p. 6700, 2022.

[6] D. Mayrawan, I. M. Wirawan, F. S. Aziz, A. I. Syah, M. A. A. Shidiqi, et al., "Development internet of things for water quality monitoring system for gouramy cultivation," in 4th Forum in Research, Science, and Technology (FIRST-T1-T2-2020), pp. 197–201, Atlantis Press, 2021.

[7] S. K. Vasudevan and B. Baskaran, "An improved real-time water quality monitoring embedded system with iot on unmanned surface vehicle," Ecological Informatics, vol. 65, p. 101421, 2021.

[8] K. R. S. R. Raju and G. H. K. Varma, "Knowledge based real time monitoring system for aquaculture using iot," in 2017 IEEE 7th international advance computing conference (IACC), pp. 318–321, IEEE, 2017.

[9] B. Das and P. Jain, "Real-time water quality monitoring system using internet of things," in 2017 International conference on computer, communications and electronics (Comptelix), pp. 78–82, IEEE, 2017.

[10] M. M. Billah, Z. M. Yusof, K. Kadir, A. M. M. Ali, and I. Ahmad, "Quality maintenance of fish farm: development of real-time water quality monitoring system," in 2019 IEEE International Conference on Smart Instrumentation, Measurement and Application (ICSIMA), pp. 1–4, IEEE, 2019.

[11] M. M. Hemal, A. Rahman, Nurjahan, F. Islam, S. Ahmed, M. S. Kaiser, and M. R. Ahmed, "An integrated smart pond water quality monitoring and fish farming recommendation aquabot system," Sensors, vol. 24, no. 11, 2024.

[12] Al-Mutairi and K. Al-Aubidy, "Iot-based smart monitoring and management system for fish farming," Bulletin of Electrical Engineering and Informatics, vol. 12, no. 3, pp. 1435–1446, 2023.

[13] J. Huan, H. Li, F. Wu, and W. Cao, "Design of water quality monitoring system for aquaculture ponds based on nb-iot," Aquacultural Engineering, vol. 90, p. 102088, 2020.

[14] M.-C. Chiu, W.-M. Yan, S. A. Bhat, and N.-F. Huang, "Development of smart aquaculture farm management system using iot and ai-based surrogate models," journal of Agriculture and Food Research, vol. 9, p. 100357, 2022.

[15] S. Pasika and S. T. Gandla, "Smart water quality monitoring system with cost-effective using iot," Heliyon, vol. 6, no. 7, 2020.

[16] G. Gao, K. Xiao, and M. Chen, "An intelligent iot-based control and traceability system to forecast and maintain water quality in freshwater fish farms," Computers and Electronics in Agriculture, vol. 166, p. 105013, 2019.

[17] M. Cordova-Rozas, J. Aucapuri-Lecarnaque, and P. Shiguihara-Jua´rez, “A cloud monitoring system for aquaculture using iot,” in 2019 IEEE Sciences and Humanities International Research Conference (SHIRCON), pp. 1–4, Nov 2019.
[18] K. R. S. R. Raju and G. H. K. Varma, “Knowledge based real time monitoring system for aquaculture using iot,” in 2017 IEEE 7th International Advance Computing Conference (IACC), pp. 318–321, Jan 2017.
[19] D. Pikulins, R. Gotlaufs, T. Solovjova, A. Aboltins, J. Griza¯ns, and S. Migla, “On the development of long-range water quality monitoring system for outdoor aquaculture objects,” Electrical, Control and Communication Engi- neering, vol. 18, no. 1, pp. 37–42, 2022.
[20] M. Niswar, S. Wainalang, A. A. Ilham, Z. Zainuddin, Y. Fujaya, Z. Muslimin, A. W. Paundu, S. Kashihara, and D. Fall, “Iot-based water quality monitoring system for soft-shell crab farming,” in 2018 IEEE international confer- ence on Internet of Things and intelligence system (IOTAIS), pp. 6–9, IEEE, 2018.
[21] T. Abinaya, J. Ishwarya, and M. Maheswari, “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), pp. 1–4, IEEE, 2019.
[22] M. M. Islam, “Real-time dataset of pond water for fish farming using iot devices,” Data in Brief, vol. 51, p. 109761, 2023.
[23] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, “Smote: synthetic minority over-sampling tech- nique,” Journal of artificial intelligence research, vol. 16, pp. 321–357, 2002.

Authors’ Profiles

Md. Shahriar Hossain Apu is a dedicated student at Bangabandhu Sheikh Mujibur Rahman Digital University, Bangladesh. He is currently pursuing a Bachelor of Science (B.Sc.) in Internet of Things and Robotics Engineer- ing. His academic interests lie in the intersection of advanced technologies and their applications in real-world scenarios. Shahriar is passionate about leveraging IoT and robotics to create innovative solutions, particularly in the fields of AI, data science, and robotics, where he aims to enhance efficiency and sustainability through cutting-edge research and development.

Md. Moshiur Rahman obtained his M.Sc. in Software Engineering from the Institute of Information Technol- ogy (IIT), University of Dhaka in 2022, and his B.Sc. in Computer Science and Engineering from Patuakhali Science and Technology University in 2019. He worked as a lecturer in the Department of Computer Science and Engineering at Green University of Bangladesh (GUB) from February 2, 2021, to August 30, 2023. Currently, he is a lecturer in the Department of Software Engineering at Bangabandhu Sheikh Mujibur Rahman Digital Uni- versity, Bangladesh (BDU), starting from January 1, 2024. His research interests include Software Engineering, Machine Learning, Data Science, and Pattern Recognition.

Md. Toukir Ahmed post graduated from Rajshahi University of Engineering & Technology (RUET) with an M.Sc. Engineering degree in Computer Science and Engineering (CSE) and graduated from the same University with a B.Sc. Engineering degree in CSE. Now he is working as a full-time faculty member at IoT and Robotics Engineering (IRE) department in Bangabandhu Sheikh Mujibur Rahman Digital University, Bangladesh. Pre- viously he worked as a full-time faculty member at ICT and CSE department in Bangladesh Army University of Science and Technology, Saidpur. Also, he worked as a full-time faculty member at CSE department in Bangladesh University, Dhaka. His research interests include Deep Learning, Traditional Machine Learning and Data Mining.

How to cite this paper: Md. Shahriar Hossain Apu, Md. Moshiur Rahman, Md. Toukir Ahmed, "IoT-Based Fish Recommendation System: A Machine Learning Approach via Mobile Application for Precision Agriculture", International Journal of Information Engineering and Electronic Business(IJIEEB), Vol.16, No.6, pp. 71-85, 2024. DOI:10.5815/ijieeb.2024.06.06