IoT Based Fish Recomandation System A Machine Learning Apporch

Md.Shahriar Hossain Apu
Dept.of IoT and Robotics Engineering
Bangabandhu Sheikh Mujibur Rahman Digital University
Bangaldesh
1901036@iot.bdu.ac.bd

Bicrom Adhikari Utsa
Dept.of IoT and Robotics Engineering
Bangabandhu Sheikh Mujibur Rahman Digital University
Bangaldesh
1901034@iot.bdu.ac.bd

Md.Nur-E Ferdous
Dept.of IoT and Robotics Engineering
Bangabandhu Sheikh Mujibur Rahman Digital University
Bangaldesh
1901033@iot.bdu.ac.bd

Chowdhury Rafsan Mohammadullah
Dept.of IoT and Robotics Engineering
Bangabandhu Sheikh Mujibur Rahman Digital University
Bangaldesh
1801048@iot.bdu.ac.bd

***Abstract*—Fish farming has become a crucial component of the global food industry, with an increasing demand for sustainable and efficient practices. To address the challenges of maintaining water quality in aquaculture, a real-time Internet of Things (IoT) dataset has been developed to monitor and manage pond water conditions. This dataset aims to optimize fish farming by providing timely and accurate information to fish farmers, researchers, and aquaculture enthusiasts. The Real-time IoT Pond Water Dataset comprises various sensors and devices that continuously collect data on critical water parameters, including temperature, pH and turbidity. These parameters play a pivotal role in the health and growth of fish populations, making real-time monitoring essential.This dataset enables real-time analysis and visualization of pond water conditions, allowing farmers to make informed decisions regarding feeding, aeration, and water treatment. Researchers can utilize the data to gain insights into optimal conditions for different fish species, leading to improved growth rates and reduced mortality.**

***Index Terms*—Smart fish farming ,Aquatic biology ,IoT sensor, Machine learning**

I. Introduction

Bangladesh, a nation deeply rooted in agriculture, holds fisheries as a cornerstone of its economy and food security. With over 100 million people and a growing demand for protein, the aquaculture sector plays a pivotal role in nourishing the nation. However, the traditional methods of fish farming face several challenges that hinder productivity and sustainability.

One of the primary challenges lies in predicting fish growth and yield accurately. Current methods rely on manual observations and experience-based estimations, which often lead to inaccuracies and inefficiencies. This lack of precise prediction capabilities hampers farmers' ability to optimize feed management, stocking density, and harvesting time, ultimately impacting their profitability and environmental sustainability.

The advent of the Internet of Things (IoT) presents a transformative opportunity to revolutionize aquaculture in Bangladesh. By leveraging IoT technologies, fish farmers can gain real-time insights into critical water quality parameters, such as temperature, pH, dissolved oxygen, and turbidity. These parameters play a crucial role in influencing fish growth, health, and behavior.

An IoT-based fish prediction system can harness these real-time data streams to develop sophisticated predictive models. These models can analyze historical data, current environmental conditions, and fish species-specific characteristics to accurately forecast fish growth rates and yield. By empowering farmers with such predictive capabilities, IoT-based systems can revolutionize aquaculture practices in Bangladesh, leading to a range of significant benefits:

- Enhanced Productivity and Yield: Accurate fish growth predictions enable farmers to optimize feed management, ensuring that fish receive the right amount of nutrition at the right time. This optimized feeding strategy reduces feed wastage, enhances fish growth rates, and ultimately increases yield.
- Improved Resource Utilization: Predictive insights into fish growth rates allow farmers to plan stocking densities and harvesting schedules more effectively. This optimization prevents overcrowding and ensures that resources, such as water and feed, are utilized efficiently. Improved Resource Utilization: Predictive insights into fish growth rates allow farmers to plan stocking densities and harvesting schedules more effectively. This optimization prevents overcrowding and ensures that resources, such as water and feed, are utilized efficiently.
- Reduced Environmental Impact: By optimizing feed management and reducing waste, IoT-based fish prediction systems can minimize the environmental footprint of aquaculture. This reduction in pollution and resource consumption contributes to a more sustainable and environmentally friendly aquaculture industry.
- Empowered Farmers and Enhanced Livelihoods:

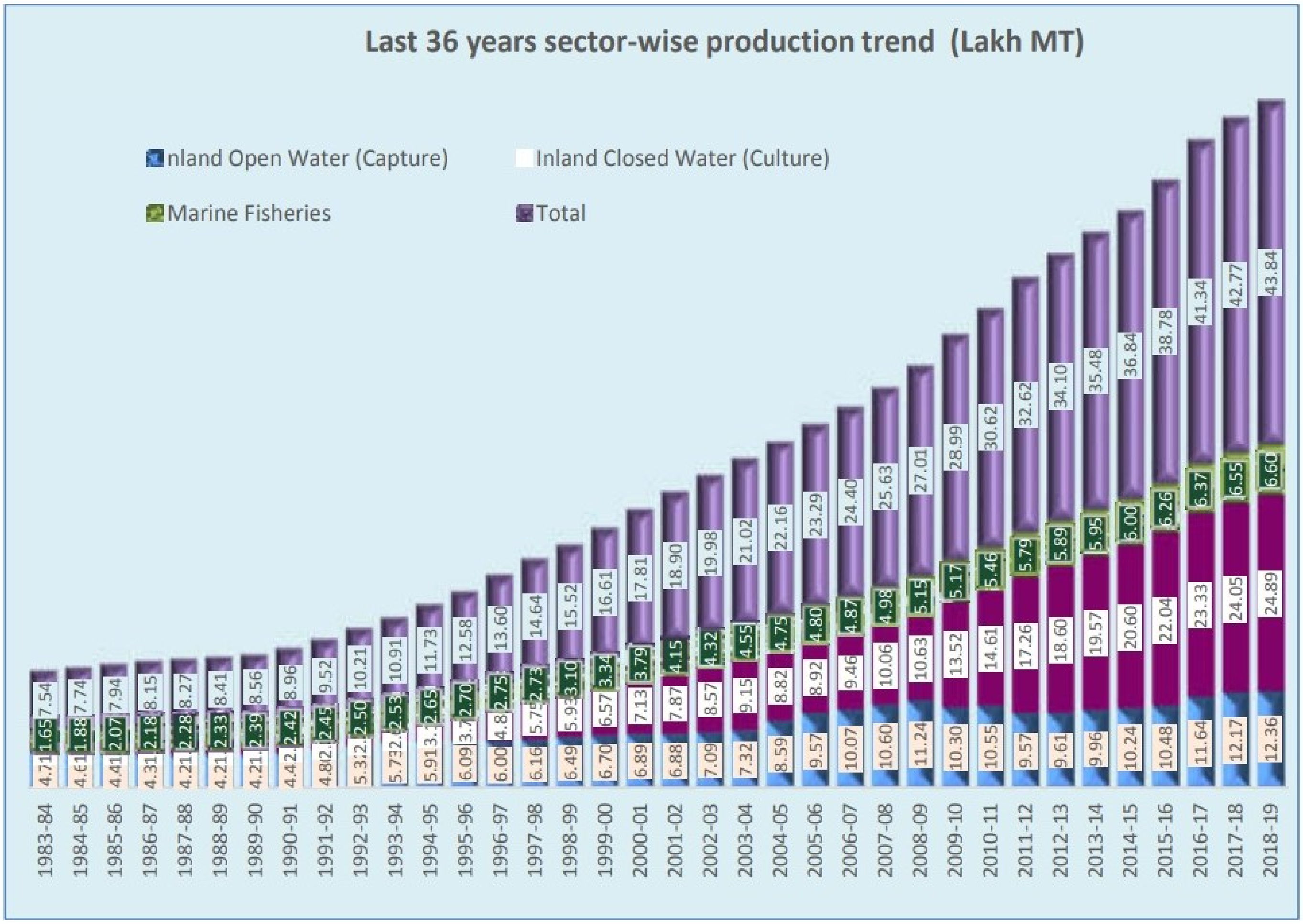

Fig. 1. Last 36 year sector wise production trend(Lakh MT)

Equipped with data-driven insights, farmers can make informed decisions, improve their aquaculture practices, and increase their profitability. This empowerment leads to enhanced livelihoods for fish farmers and their families.

- Data-driven Aquaculture Policymaking: IoT-generated data can inform policymaking decisions, enabling the government to identify areas for improvement, develop targeted interventions, and promote sustainable aquaculture practices nationwide.

If we see a of last 36 year of our country [3] then we can see the how this sector is growing up in figure 1 and figure 2 represents the statistics of pond aquaculture.

II. LITERATURE SURVEY

[12]This paper presents a real-time water quality dataset from five ponds in Bangladesh, collected using IoT technology. The dataset includes pH, temperature, turbidity, and fish data, with 11 fish categories. Only three ponds are found suitable for fish farming, and the Random Forest algorithm performs well in the analysis. The paper also provides information on the IoT system's hardware, offering a valuable resource for researchers and fish farmers for fish survival prediction.

This paper [13] discusses the use of the Internet of Things (IoT) for real-time monitoring and control in aquaculture. It introduces an IoT framework that employs sensors and Arduino microcontrollers to monitor various water-related parameters in fish farming ponds. The data collected is stored in an IoT cloud platform, and machine learning algorithms are applied to assess the suitability of ponds for fish cultivation. The study found that only three out of five ponds met the required standards, with Random Forest achieving the highest performance in classifying fish categories. The survey also includes prototype hardware details for the proposed IoT system.

[14]The paper survey discusses the challenge of fish species selection for aquaculture in specific aquatic environments and presents a machine learning approach using random forest to predict the ideal fish species. The model uses environmental characteristics like pH, temperature, and turbidity and is validated with an 11-fish dataset. Results show an 88.48% accuracy, 87.11% kappa statistic, and an 88.5% true positive rate.

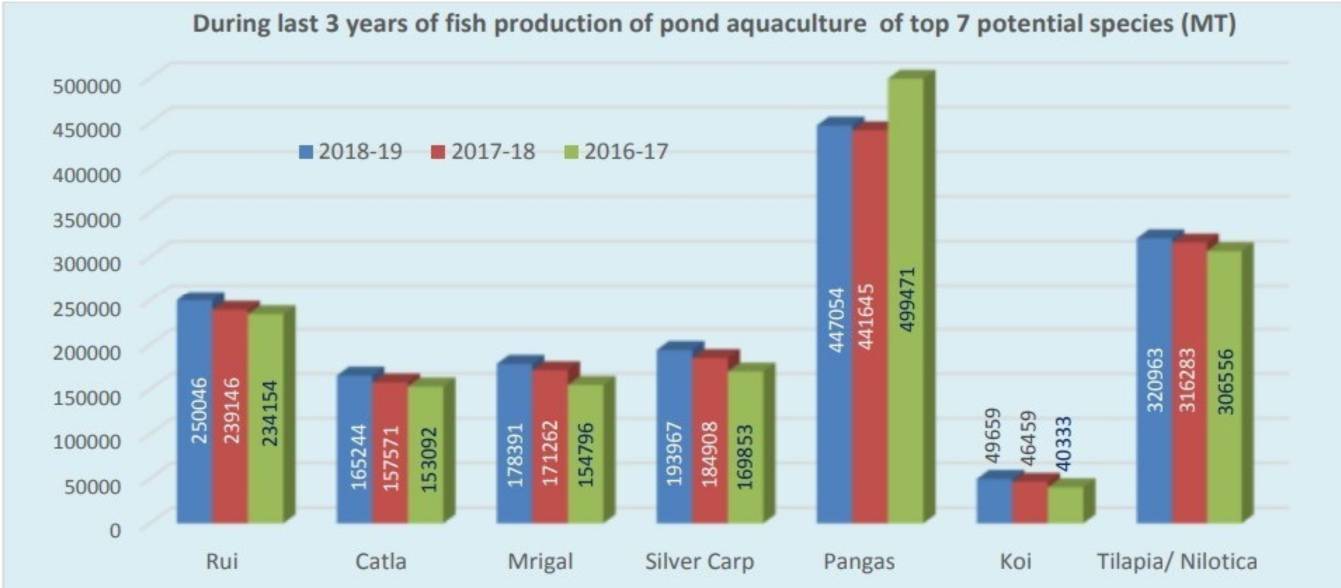

Fig. 2. : During last 3 years of production performance of top 7 potential species (MT))

Furthermore, the proposed model outperforms other existing models like J48, KNN, and CART in terms of accuracy, TP rate, and kappa statistics.

[10]This literature survey discusses the significance of pond water quality for prawn survival and aquaculture productivity. It emphasizes the need for an IoT-based system to assist cultivators in continuously monitoring water quality. The proposed system employs pH, salinity, and ultrasonic sensors, along with an ESP32 microcontroller, to collect and process data, which is then stored in a cloud firebase and displayed on an Android app. The results indicate that the system performs well, with minimal errors in sensor readings, and effectively notifies users of abnormal water parameter values.

[15]This paper describes an IoT system employing a K Nearest Neighbors Machine Learning Model to identify fish species by analyzing a dataset. Real-time sensor data is stored on a cloud server, and a dynamic website provides information about aquatic fish species. The website is linked to the cloud server for easy access, enabling users to make informed decisions about fish survival. The IoT system employs 5 sensors (mq7, pH, temperature, ultrasonic, and turbidity) connected to an Arduino Uno, with data stored in CSV format on a cloud server. This system allows remote monitoring for fish farming.

[11]This literature survey discusses the development of an embedded system for fish classification using deep learning and IoT technology. It outlines the creation of two datasets, one with original images and the other with unsharp masked photos. Several deep learning models, including InceptionV3, Xception, and DenseNet, were applied to these datasets, achieving a maximum accuracy of 97% with a hybrid CNN and Convolutional LSTM model. Additionally, an IoT-based solution was explored, revealing potential for real-time applications despite some limitations. Overall, the research demonstrates the effectiveness of integrating deep learning and IoT in automated fish classification systems.

III. METHODOLOGY

A. Dataset

[4]This data set contains real-time data gathered from a monitoring system that uses an Internet of Things framework to evaluate aquatic conditions. The system uses three sensors—pH, temperature, and turbidity—along with an Arduino controller to monitor the water quality in five different ponds on a continuous basis. The four columns in the data set—pH, temperature, turbidity, and fish—comprise 40,280 records. ”Fish” is the target variable in these columns, and the other three are the independent variables. Eleven different fish types with unique values—tilapia 8830 rui 6336 pangas 5314 silverCup 3906 katla 3786 sang 3776 prawn 3204 karpio 2112 prawn 1348 koi 964 magur 704—are present in the sample. The size of the dataset is one of the dataset’s remaining restrictions.

TABLE I
FEATURE OF THE DATA SET

Feature	Description
pH	Amount of pH present in the water
Temperature	Amount of the temperature in the water in celcius
Turbidity	Amount of turbidity in water
Fish	The target class

B. Hardware

For our system we used arduino,raspberry-pi,turbidity sensor,pH sensor,temaprature sensor.Sensor description are given below:

	ph	temperature	turbidity	fish
0	6.0	27.0	4.0	katla
1	7.6	28.0	5.9	sing
2	7.8	27.0	5.5	sing
3	6.5	31.0	5.5	katla
4	8.2	27.0	8.5	prawn

Fig. 3. Snapshot of Data Set

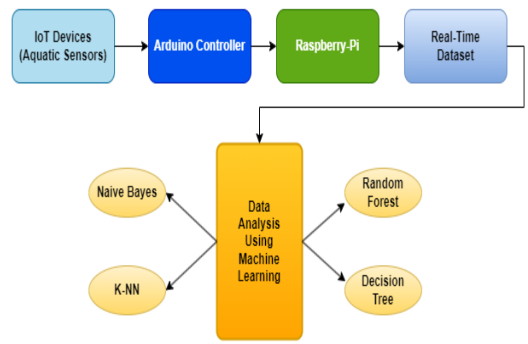

Fig. 4. Workflow of designed method

- **Arduino uno:**Arduino Uno is a popular microcontroller board. It is based on the Atmega328P microcontroller and has 14 digital input/output pins, 6 analog input pins, a 16 MHz crystal oscillator, a USB connection, a power jack, and an ICSP header.
- **Raspberry-Pi;** A Raspberry Pi is a small single-board computer (SBC) developed in the United Kingdom by the Raspberry Pi Foundation with the intention of promoting teaching of basic computer science in schools and colleges. It has since been adopted for many other purposes, including robotics, home automation, and media centers.
- **pH Sensor:** [1]pH sensor is one of the most important tools for measuring pH and is commonly used in water quality monitoring. This type of sensor is capable of measuring alkalinity and acidity in water and other solutions. When used properly, pH sensors can ensure the safety and quality of products and processes that occur in wastewater or manufacturing plants.
- **Turbidity Sensor:**A turbidity sensor [2]is an analytical sensor that measures turbidity. They are highly useful and effective instruments to identify the clarity and particle content in a solution, like water. Turbidity sensors are used to reduce waste, improve yields, and analyze water quality in a wide range of industries.
- **Temperature sensor:**For measuring water temperature and other fluids, waterproof temperature sensors are required. One of such temperature sensors [5] is DS18B20. This sensor can measure the temperature of the air, liquids like water, and ground. The sensor comes in two form-factors, of which one is a waterproof module. It can be used to sense temperature in applications like electric steam cookers, electric kettles, and temperature-controlled water storage.

C. System Implementation

The overall proposed method of getting the real-time aquatic values from the IoT framework is shown in Fig. 4 . First of all, we defined the aquatic sensors then the Arduino setup read the data through the sensors and send it into Raspberry-pi. The data set with the fish specifies is evaluated by machine learning models to predict the fish survival . But the Random forest has the highest accuracy that's why we use this model to train our system.The physical part of the IoT framework is displayed in Fig. 5 . Fig. 6 shows some of the experimental images of the proposed system.

IV. System Learning using ML Algorithm

We trained our using some well known Machine learning algorithm. These are-

A. DECISION TREE

[7]Decision tree is a supervised machine learning algorithm which can be used for both classification and regression. It is mostly preferred to solve classification problems in which data has to be classified into different categories. It has a tree like structure with internal node acting as features, branches as decision rule and leaf nodes as the outcome. It is basically a graphical representation of all the possible outcomes for a given problem based on the given conditions. It simply asks a question and based on the answers it further splits the trees. In order to build a tree, it uses CART algorithm.

B. RANDOM FOREST

[9]Random Forest is a classifier that contains number of decision tree on various subsets. It is a supervised machine learning algorithm used for both classification and regression. To improve the predicted accuracy of the dataset it takes the average of the decision trees. It is based on the concept of ensemble learning, which is basically a process of combining several classifiers to solve complex problems and to improve the performance of the model. To have the higher accuracy and prevent the problem of overfitting, greater number of trees are used.

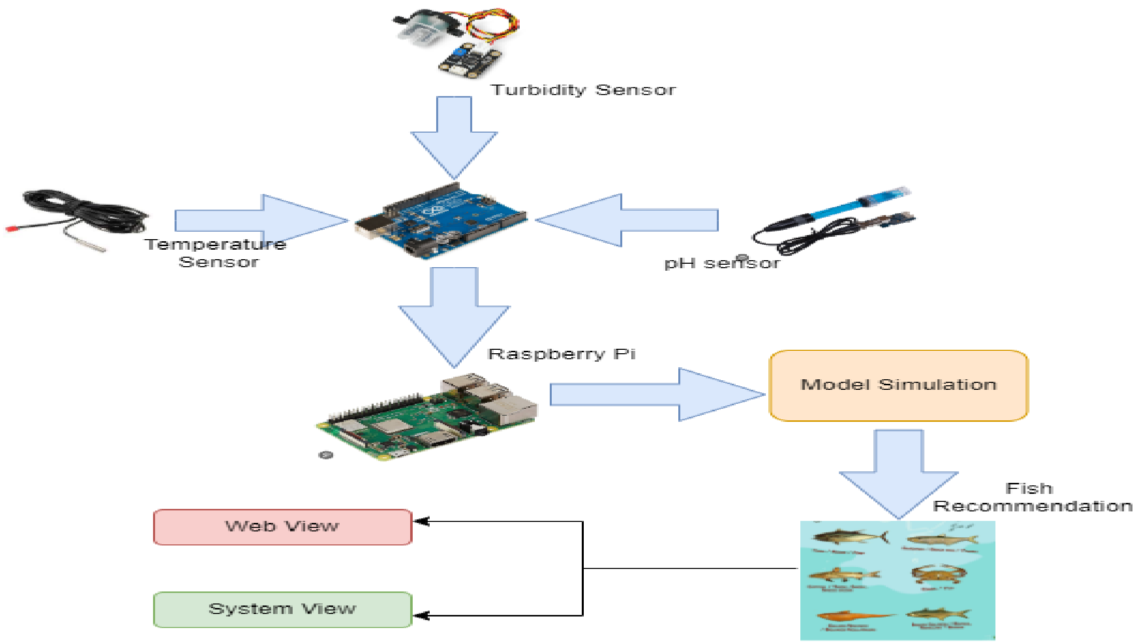

Fig. 5. System Architecture

C. KNN

[6]K-Nearest Neighbour is one of the simplest Machine Learning algorithms based on Supervised Learning technique. K-NN algorithm assumes the similarity between the new case/data and available cases and put the new case into the category that is most similar to the available categories. K-NN algorithm stores all the available data and classifies a new data point based on the similarity. This means when new data appears then it can be easily classified into a well suite category by using K- NN algorithm. K-NN algorithm can be used for Regression as well as for Classification but mostly it is used for the Classification problems. K-NN is a non-parametric algorithm, which means it does not make any assumption on underlying data. It is also called a lazy learner algorithm because it does not learn from the training set immediately instead it stores the dataset and at the time of classification, it performs an action on the dataset. KNN algorithm at the training phase just stores the dataset and when it gets new data, then it classifies that data into a category that is much similar to the new data.

D. Naïve Bayes Classifier

[8]Naïve Bayes algorithm is a supervised learning algorithm, which is based on Bayes theorem and used for solving classification problems. It is mainly used in text classification that includes a high-dimensional training dataset. Naïve Bayes Classifier is one of the simple and most effective Classification algorithms which helps in building the fast machine learning models that can make quick predictions. It is a probabilistic classifier, which means it predicts on the basis of the probability of an object. Some popular examples of Naïve Bayes Algorithm are spam filtration, Sentimental analysis, and classifying articles.

V. Result and Analysis

We Developed a User-Interface for the system and also for a web application to that user can easily use this sysytem.Besides this In this section, dataset used along with use of various Machine Learning algorithms for prediction is highlighted. The dataset is used is from Kaggle. This dataset consists of water quality metrics for 591 different water bodies. Key features used to tally the results are: pH value, Turbidity, Temperature. The results of the different machine Learning algorithms namely KNN, Decision Tree, Random Forest,Naive Bayes are discussed based on parameters like accuracy, precision, recall and F1-score. The main aspect taken into consideration is accuracy.

Accuracy=TP+TN/TP+FP+FN+TN

The results of the algorithms are mentioned below:

Acknowledgment

I would like to express my sincere gratitude to my teacher, Nurjahan Nipa, for her invaluable guidance and unwavering support throughout the course of this data science project. Her expertise, patience, and dedication were instrumental in helping me navigate the complexities of data analysis and machine learning.

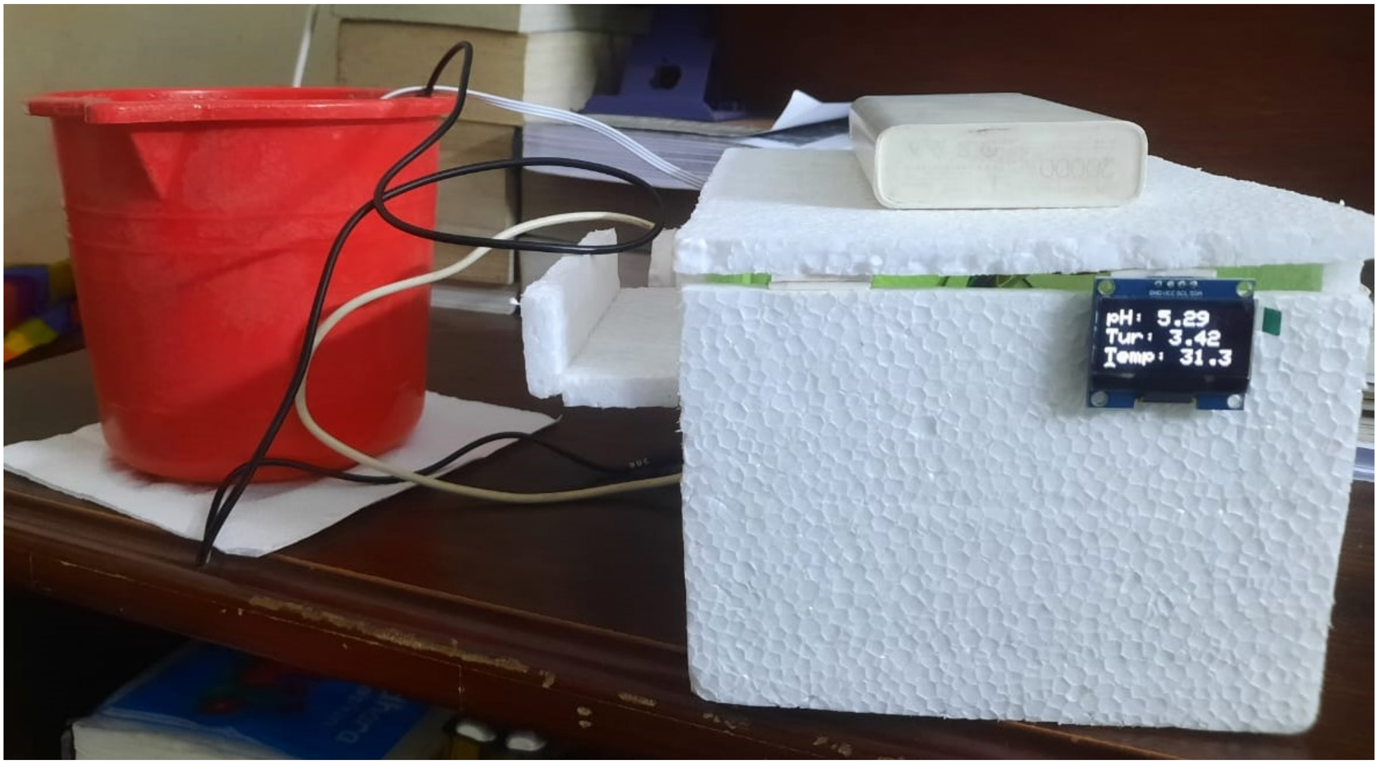

Fig. 6. Experimental Picture

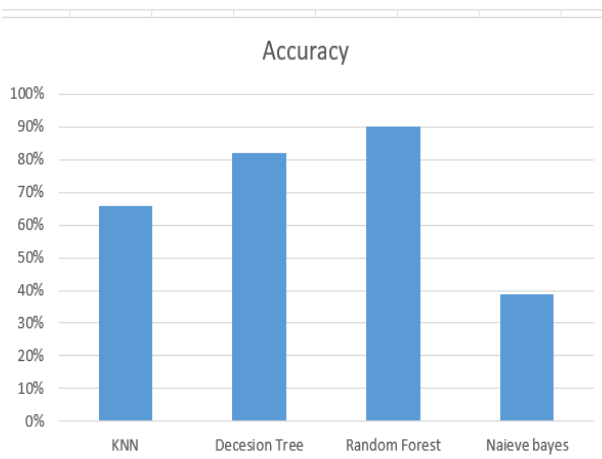

Fig. 7. Model Accuracy

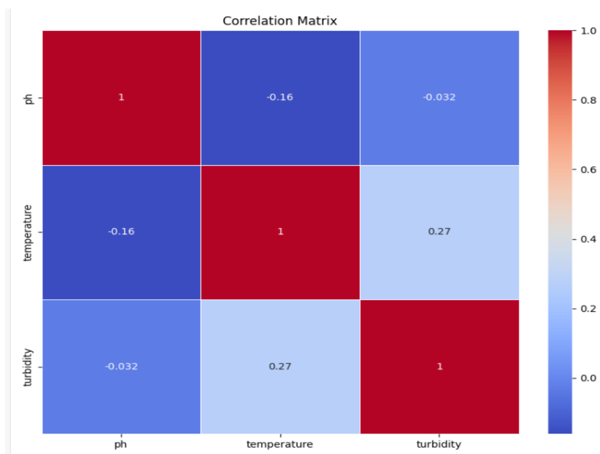

Fig. 8. Co-Relation Matrix

TABLE II
MODEL ACCURACY

Moldel	Accuracy	Precision	Recall	F1-Score
KNN	66%	71%	66%	67%
Naive Bayes	60%	42%	36%	39%
Decision Tree	82%	86%	82%	83%
Random Forest	90%	90%	90%	89%

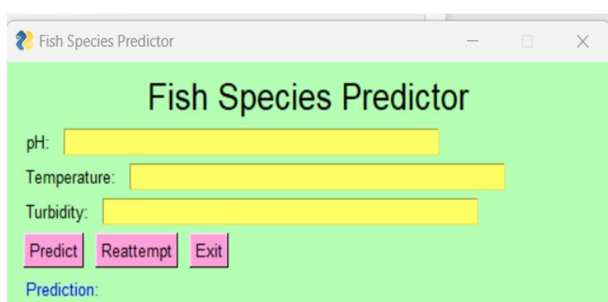

Fig. 9. System Dashboard

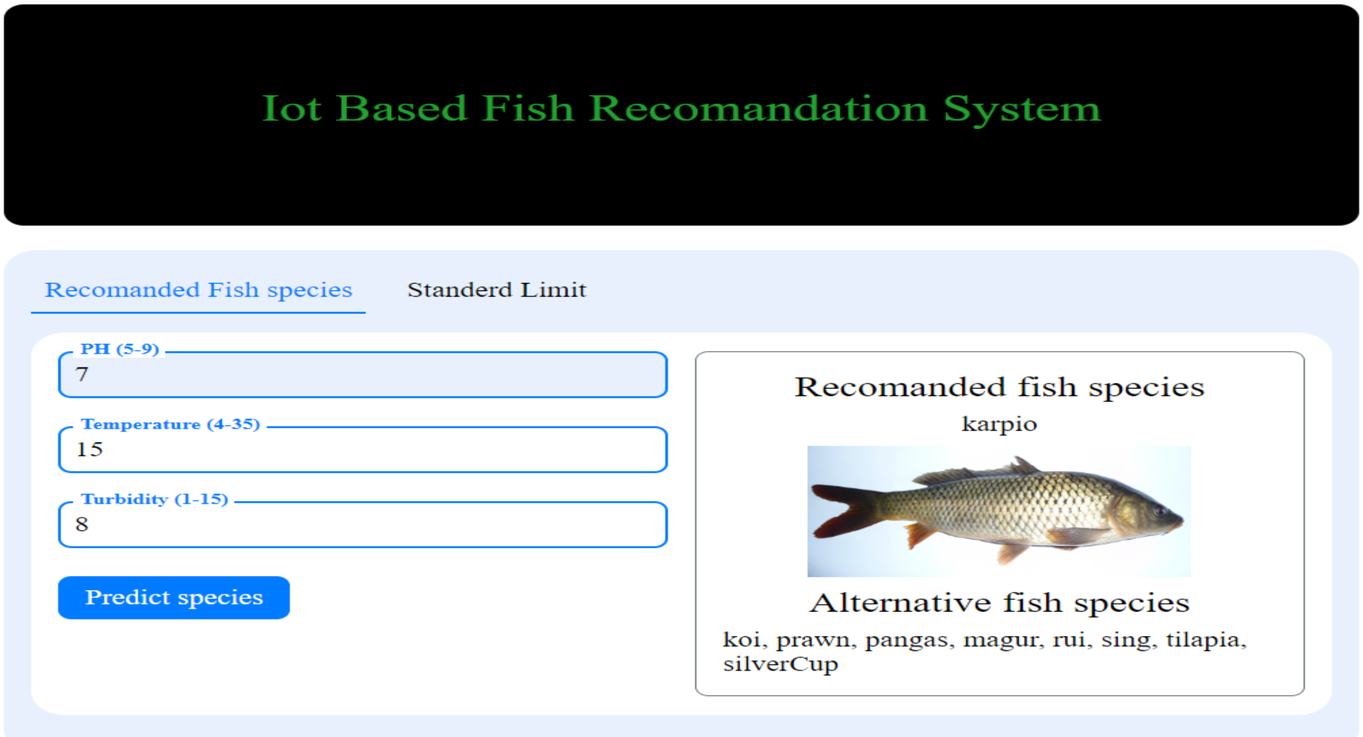

Fig. 10. Web Interface

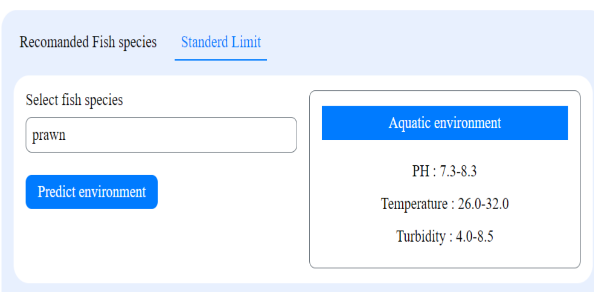

Fig. 11. Fish Environmental Parameter suggested in web

Nurjahan Nipa's insightful feedback and constructive criticism played a pivotal role in shaping the success of this project. Her commitment to fostering a deep understanding of data science concepts and her willingness to go the extra mile to assist her students were truly commendable.

I would also like to extend my appreciation to my fellow classmates who collaborated with me on this project, as well as my friends and family for their constant encouragement.

VI. CONCLUSION AND FUTURE WORK

IoT-based fish prediction systems offer a transformative solution to revolutionize aquaculture in Bangladesh. By leveraging real-time data streams collected from IoT sensors, these systems enable farmers to predict fish growth rates and yield with unprecedented accuracy. This predictive capability empowers farmers to optimize feed management, stocking densities, and harvesting schedules, leading to a range of benefits.

REFERENCES

[1] https://apureinstrument.com/blogs/about-ph-sensor/.
[2] https://atlas-scientific.com/blog/what-is-a-turbidity-sensor/: :text=a
[3] https://fisheries.portal.gov.bd/.
[4] https://www.kaggle.com/datasets/monirmukul/realtime-pond-water-dataset-for-fish-farming.
[5] https://www.kaggle.com/datasets/monirmukul/realtime-pond-water-dataset-for-fish-farminghttps://www.engineersgarage.com/arduino-based-water-temperature-monitor-using-ds18b20-1-wire-temperature-sensor/.
[6] https://www.javatpoint.com/k-nearest-neighbor-algorithm-for-machine-learning, 2023. Last accessed 16 September 2017.
[7] https://www.javatpoint.com/machine-learning-decision-tree-classification-algorithm, 2023.
[8] https://www.javatpoint.com/machine-learning-naive-bayes-classifier, 2023. Last accessed 16 September 2017.
[9] https://www.javatpoint.com/machine-learning-random-forest-algorithm, 2023.
[10] Ramzi Adriman, Maya Fitria, Afdhal Afdhal, and Alfy Yusyfa Fernanda. An iot-based system for water quality monitoring and notification system of aquaculture prawn pond. In *2022 IEEE International Conference on Communication, Networks and Satellite (COMNETSAT)*, pages 356–360. IEEE, 2022.
[11] Md. Asif Ahmed, Md. Shakil Hossain, Wahidur Rahman, Abdul Hasib Uddin, and Md. Tarequl Islam. An advanced bangladeshi local fish classification system based on the combination of deep learning and the internet of things (iot). *Journal of Agriculture and Food Research*, 14:100663, 2023.
[12] Md. Monirul Islam. Real-time iot pond water dataset for fish farming. *Data in Brief*, page 109761, 2023.
[13] Md Monirul Islam, Mohammod Abul Kashem, Salem A Alyami, and Mohammad Ali Moni. Monitoring water quality metrics of ponds with iot sensors and machine learning to predict fish species survival. *Microprocessors and Microsystems*, 102:104930, 2023.

[14] Md Monirul Islam, Mohammod Abul Kashem, and Jia Uddin. Fish survival prediction in an aquatic environment using random forest model. *Int J Artif Intell ISSN*, 2252(8938):8938, 2021.

[15] Md Monirul Islam, Jia Uddin, Mohammod Abul Kashem, Fazly Rabbi, and Md Waliul Hasnat. Design and implementation of an iot system for predicting aqua fisheries using arduino and knn. In *Intelligent Human Computer Interaction: 12th International Conference, IHCI 2020, Daegu, South Korea, November 24–26, 2020, Proceedings, Part II 12*, pages 108–118. Springer, 2021.