IoT-based Phytoplankton and Zooplankton Growth Monitoring System for Smart Aquaculture

by

Iqbal Hoshen Saif (201014040)
Nahian Raisa (201014076)
Md. Arif Fuad Akash (201014083)

Capstone Project (CSE499) report submitted in partial fulfillment of the requirements for the degree of

Bachelor of Science in Computer Science and Engineering

Under the supervision of

Muhammad Golam Kibria, PhD, SMIEEE

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

UNIVERSITY OF LIBERAL ARTS BANGLADESH

SUMMER 2024

DECLARATION

Project Title	IoT-based Phytoplankton and Zooplankton Growth Monitoring System for Smart Aquaculture
Authors	*Iqbal Hoshen Saif*, *Nahian Raisa*, and *Md. Arif Fuad Akash*
Student IDs	201014040, 201014076, 201014083
Supervisor	Muhammad Golam Kibria, PhD, SMIEEE

We declare that this capstone project entitled *IoT-based Phytoplankton and Zooplankton Growth Monitoring System for Smart Aquaculture* is the result of our own work except as cited in the references. The capstone project has not been accepted for any degree and is not concurrently submitted in candidature of any other degree.

Iqbal Hoshen Saif
201014040

Department of Computer Science and Engineering
University of Liberal Arts Bangladesh

Nahian Raisa
201014076

Department of Computer Science and Engineering
University of Liberal Arts Bangladesh

Md. Arif Fuad Akash
201014083

Department of Computer Science and Engineering
University of Liberal Arts Bangladesh

Date: November 3, 2024

Department of Computer Science and Engineering
University of Liberal Arts Bangladesh
Mohammadpur, Dhaka - 1207

CERTIFICATE

This is to certify that the capstone project entitled **IoT-based Phytoplankton and Zooplankton Growth Monitoring System for Smart Aquaculture**, submitted by **Iqbal Hoshen Saif** (Student ID: 201014040), **Nahian Raisa** (Student ID: 201014076), and **Md. Arif Fuad Akash** (Student ID: 201014083) who are undergraduate students of the **Department of Computer Science and Engineering**, has been examined. Upon recommendation by the examination committee, we hereby accord our approval of it as the presented work and submitted report fulfill the requirements for its acceptance in partial fulfillment for the degree of *Bachelor of Science in Computer Science and Engineering*.

Muhammad Golam Kibria, PhD, SMIEEE, Professor
Dept. of CSE, University of Liberal Arts Bangladesh

Mohammed Ashikur Rahman, PhD, Assistant Professor
Dept. of CSE, University of Liberal Arts Bangladesh

Rubaiya Hafiz, Senior Lecturer
Dept. of CSE, University of Liberal Arts Bangladesh

Suravi Akhter, Lecturer
Dept. of CSE, University of Liberal Arts Bangladesh

Place: Dhaka
Date: November 3, 2024

ACKNOWLEDGEMENTS

We would like to express our deep and sincere gratitude to our research supervisor, *Muhammad Golam Kibria, PhD, SMIEEE*, for giving us the opportunity to conduct research and providing invaluable guidance throughout this work. His dynamism, vision, sincerity and motivation have deeply inspired us. He has taught us the methodology to carry out the work and to present the works as clearly as possible. It was a great privilege and honor to work and study under his guidance.

We are greatly indebted to our honorable teachers of the Department of Computer Science and Engineering at the University of Liberal Arts Bangladesh who taught us during the course of our study. Without any doubt, their teaching and guidance have completely transformed us to the persons that we are today.

We are extremely thankful to our parents for their unconditional love, endless prayers and caring, and immense sacrifices for educating and preparing us for our future. We would like to say thanks to our friends and relatives for their kind support and care.

Finally, we would like to thank all the people who have supported us to complete the project work directly or indirectly.

Iqbal Hoshen Saif, Nahian Raisa, and Md. Arif Fuad Akash
University of Liberal Arts Bangladesh
Date: November 3, 2024

Dedicated to my beloved siblings
the driving force behind my resilience and ambition

– Iqbal Hoshen Saif

Dedicated to my supportive family
the source of my strength and inspiration.

– Nahian Raisa

Dedicated to my beloved supervisor Professor Muhammad Golam Kibria
a good soul.

– Md. Arif Fuad Akash

ABSTRACT

Plankton is a fundamental part of the aquatic food chain in aquaculture, serving as the primary food source for fish. Maintaining optimal plankton levels in fish farm ponds is essential for sustaining a healthy aquatic ecosystem. However, traditional monitoring methods, such as using a Secchi disk or manual sampling, often prove inadequate and may lead to misinterpretation by farmers. This can result in either insufficient plankton, which hampers fish growth, or excessive plankton, which can cause harmful algal blooms that deplete oxygen levels in the pond, leading to mass fish mortality. To address these challenges, this study proposes an IoT-based smart aquaculture system designed to monitor plankton growth in real-time using key water quality indicators. The system employs an IoT-embedded setup consisting of sensors and an Arduino micro-controller to measure pH, temperature, total dissolved solids (TDS), and turbidity from different points in the pond. Data was collected from January to mid-June from 13 fish ponds located in Dhaka, Magura, and Noakhali, Bangladesh. In addition, plankton samples were collected and analyzed in the laboratory to ensure the accuracy of the measurements. The data was then pre-processed and used to train and test various machine learning models for predicting plankton abundance and dissolved oxygen (DO), with DO serving as a critical feature in plankton prediction. The machine learning models explored include Support Vector Machine (SVM), Random Forest (RF), and K-Nearest Neighbors (KNN), Convolutional Neural Networks (CNN), Recurrent Neural Networks (RNNs), Gradient Boosting, and XGBoost. Among these, the RNN model delivered the highest accuracy for plankton abundance prediction at 93%, while XGBoost followed with 90% for Phytoplankton and 93% Zooplankton accuracy. The best-performing models were subsequently integrated into a mobile application that provides fish farmers with real-time updates on plankton levels. This system enables early detection of abnormal plankton growth, allowing farmers to take preventive measures, ensuring healthy fish production and increased profitability.

Keywords: Phytoplankton, Zooplankton, IoT, Aquaculture, Life Under Water.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Background Study

The adoption of IoT technologies in aquaculture has brought significant advancements in water quality monitoring and management, offering real-time insights that enhance fish production and sustainability. Maintaining a balanced aquatic environment in fish farming is critical to fish health, growth, and productivity, particularly in large-scale operations like those in Bangladesh, where aquaculture contributes significantly to national fish production.

However, traditional practices in monitoring plankton growth, water quality, and environmental conditions have posed several challenges. These methods, which rely on laboratory tests or expensive technologies, often fall short due to cost, time, and accessibility issues for small and medium-scale fish farmers.

IoT-based solutions provide a more cost-effective and scalable alternative by integrating sensors to continuously monitor critical water quality parameters such as pH, temperature, turbidity, and total dissolved solids (TDS). These parameters are directly linked to plankton growth, which plays an essential role in aquaculture ecosystems by serving as a primary food source for fish. Maintaining the right balance in plankton growth is vital, as both excessive and insufficient growth can result in harmful consequences like toxic algal blooms or oxygen depletion, causing massive fish deaths.

Recent research has emphasized the importance of monitoring plankton abundance and dissolved oxygen (DO) levels to prevent such risks. Although DO sensors are expensive, IoT-based systems offer innovative approaches to estimating DO using indirect parameters like temperature and turbidity, reducing costs while ensuring accurate predictions. Furthermore, advancements in machine learning (ML) algorithms like Random Forest, Decision Trees,Convolutional Neural Networks (CNN), Gradient Boosting and Support Vector Machines (SVM) have allowed for predictive analytics, making it possible to forecast plankton growth and water quality trends.

These technological developments highlight the potential of IoT and ML to revolutionize aquaculture by improving resource management, sustainability, and profitability. Integrating real-time monitoring with predictive analytics through mobile applications can provide fish farmers with timely insights, empowering them to make informed decisions that ensure a healthy aquatic environment. As research in this field grows, IoT-based monitoring systems stand poised to offer comprehensive solutions that are both accessible and efficient, meeting the needs of the modern aquaculture industry.

1.2 Problem statement

Phytoplankton and Zooplankton are vital in maintaining the health and productivity of aquatic ecosystems, serving as the foundation of the food chain. So balancing plankton abundance through constant monitoring is essential for ensuring healthy and stable aquatic life. Unfortunately, traditional methods of checking plankton abundances such as lab testing on nitrogen, and phosphorus levels, measuring the chlorophyll-a, and monitoring high-resolution zooplankton growth images are still inadequate and challenging in terms of real-time monitoring and cost. Measuring nitrogen and phosphorus levels through lab testing is time-consuming and impractical for constant monitoring. Measuring chlorophyll-a using sensor offers an automated solution but remains prohibitively expensive for small-scale fish farm owners. Furthermore, real-time monitoring of high-resolution zooplankton growth images becomes unfeasible in turbid water conditions due to the high cost and time demands. These issues create the need for a more affordable and time-effective smart solution for phytoplankton and zooplankton growth monitoring.

1.3 Objectives and Significance of the study

1.3.1 Aims

This research aims to design and propose an IoT-based Phytoplankton and Zooplankton growth monitoring system for ensuring a smart aquatic ecosystem. The real-time monitoring system will be capable of monitoring the abundance and growth of plankton depending on a few vital parameters of water such as water pH, temperature, turbidity, total dissolved solids(TDS), and dissolved oxygen(DO). The implementation of IoT sensors will ensure system efficiency at a low cost.

1.3.2 Objective

- Design a cost and time-effective IoT system for Phytoplankton and Zooplankton growth monitoring.
- Introduce a real-time plankton growth monitoring system to develop a smart aquatic environment.
- Mitigation of the financial loss faced by fish farm owners due to the large-scale fish deaths every year.

1.3.3 Motivation

In 2020, News18 (2020) reported that an acute shortage of oxygen resulting from plankton imbalance killed over 600 metric tonnes of fish worth Rs 10 crore in Bangladesh.

On 7 January 2022, Filstrup et al. (2018) published that advanced imaging techniques such as Video Plankton Recorders (VPRs) and other high-tech imaging systems are expensive and require significant processing power for real-time monitoring, which makes them difficult to use for continuous, cost-effective plankton growth tracking.

Additionally, On July 27, 2020, the Hablützel et al. (2021) published that phosphorus and nitrogen levels directly affect chlorophyll-a, a proxy for phytoplankton biomass, but applying this knowledge to real-time, large-scale monitoring is difficult and resource-intensive.

Figure 1.1: Imbalanced plankton growth threatening aquatic ecosystem

1.3.4 Significance

The significance of this research relies on its probable ability to transform the way aquatic ecosystems are monitored and managed. By designing a cost-effective and time-efficient system, the fish farm owners can monitor Phytoplankton and Zooplankton growth in real-time, ensuring a stable and healthy aquatic ecosystem. This innovation is crucial in reducing the financial losses that fish farm owners face due to large-scale fish deaths, often caused by imbalances in plankton abundance. Through constant monitoring, early detection of harmful conditions is possible, allowing for timely interventions that can prevent mass fish deaths and improve overall productivity in aquaculture. Consequently, this promotes a more sustainable and economically feasible approach to fish farming.

1.4 Scope and Limitation

1.4.1 Design specifications and success criteria

The design specifications and success criteria for the IoT-based Phytoplankton and Zooplankton Growth Monitoring System for Smart Aquaculture are outlined as follows.

1. **Measurement Precision:**

   Specification: Sensors should have a precision within ±0.1 for pH, ±0.2°C for temperature, ±0.5 mg/L for DO, ±1 NTU for turbidity, and ±5 ppm for TDS.

Success Criteria: The accuracy should meet or exceed 95% consistency in sensor readings when compared to calibrated instruments over a 24-hour period.

2. **Real-Time Monitoring and Data Transmission:**

   Specification: Real-time data collection and transmission should occur at a rate of one reading per minute for each parameter.

   Success Criteria: The system should be able to consistently transmit data every 60 seconds with no more than 1% of data loss over a 24-hour period.

3. **Plankton Growth Prediction Accuracy:**

   Specification: Prediction models should achieve at least 85% accuracy in predicting plankton growth.

   Success Criteria: A validated accuracy of 85% or higher across all algorithms during testing phases with test datasets.

4. **Power Efficiency:**

   Specification: The system should consume less than 10 W of power on average during operation.

   Success Criteria: The system should operate within the specified power limit for 90% of operational time.

5. **Environmental Robustness:**

   Specification: The sensors and IoT devices should function effectively in temperatures ranging from 5°C to 35°C and depths up to 3 meters.

   Success Criteria: The system should maintain 100% operability under specified conditions during testing.

6. **Scalability:**

   Specification: The system should support scalability to monitor at least 10 different aquaculture ponds with up to 5 sensors per pond.

   Success Criteria: The system should maintain real-time monitoring across all sites without performance degradation.

1.4.2 Scopes of the research

1. **Real-Time Data Collection and Analysis:**

   The system will continuously collect data on four key water parameters(pH, temperature, turbidity, and total dissolved solids). Data will be collected in real-time using IoT-enabled sensors deployed in aquaculture ponds. Machine learning models such as SVM, LR, DT, and RF will analyze this data to predict phytoplankton and zooplankton growth. The system will use real-time data to improve the accuracy of predictions.

2. **Alerts on the mobile application or web application:**

   The system will send real-time alerts (via mobile app or web interface) to the fish farm quality controllers if certain thresholds are breached, such as low total dissolved solids or abnormal turbidity that may indicate a rapid plankton bloom or harmful condition.

3. **Adaptive Machine Learning Models:**

   The system will continuously learn and adapt from the data it collects. With time, machine learning models will be retrained with new datasets to improve prediction accuracy.

4. **Environmental Sustainability Focus:**

   This system will contribute to sustainable aquaculture practices by optimizing resource usage such as water and feed, reducing the need for manual testing, and preventing potential environmental issues such as overpopulation of plankton or poor water quality that can lead to fish kills.

5. **Scalability and Integration:**

   The system will be designed to scale easily, allowing for the future integration of more ponds and sensors. The system architecture will be flexible, allowing for the introduction of new technologies such as more nitrogen sensors or machine learning models for predicting additional environmental variables.

1.4.3 Limitations of the research

1. **Data Reliability and Sensor Accuracy:**

   The accuracy of predictions and real-time insights heavily depends on the reliability and precision of the sensors. Sensors exposed to harsh environmental conditions like extreme temperatures, and heavy biofouling may degrade over time, leading to inaccurate readings. Sensor calibration and maintenance will be crucial. The system's effectiveness will diminish if sensors become uncalibrated or malfunction. Hence, regular maintenance will be required, which adds to operational costs.

2. **Model Limitations and Generalization:**

   While machine learning models can predict plankton growth with a high degree of accuracy, they may struggle to generalize beyond the specific conditions for which they were trained. For example, models trained on data from one pond may not perform well when applied to a different pond with unique environmental conditions. The system's predictive capabilities may be limited during extreme weather events or abrupt environmental changes such as storms, or pollution incidents that were not present in the training data.

3. **Network and Connectivity Issues:**

   The system's ability to transmit data in real-time depends on network connectivity. Remote aquaculture locations might have poor connectivity, leading to data delays or interruptions in monitoring and predictions. Backup data storage options may be needed to ensure that data is not lost during connectivity issues. Real-time data processing may be hampered by slow or unreliable internet connections, especially in rural or remote areas where aquaculture fish farms are often located.

4. **Power Dependency:**

   The IoT devices, including sensors, controllers, and communication modules, are dependent on a reliable power supply. If the system is deployed in remote locations without stable power sources, solar or battery-powered options may be necessary. However, these options may add costs or limit the operational time between charges. In the case of extensive system deployment across multiple ponds, increased power consumption may become a concern, particularly for the continuous operation of sensors and data transmission modules.

5. **Environmental Variability:**

   The system's prediction models assume a stable or predictable relationship between water parameters and plankton growth. However, environmental factors such as weather, seasonal changes, or the introduction of new species into the pond may disrupt these relationships, leading to inaccurate predictions. Abrupt ecological changes, such as pollution, intrusive species, or chemical spills, might not be effectively accounted for by the machine learning models, requiring manual intervention and reevaluation of the system's operation.

The scope of the IoT-based smart aquaculture system is determined, with the potential to significantly change sustainable aquaculture practices. However, its success will be determined by the technology's precision, the adaptability of machine learning models, and the availability of reliable power and connectivity infrastructure. Addressing these limits through good planning, frequent maintenance, and the implementation of backup systems will be essential for building a reliable and efficient solution.

1.5 Project Planning

The project planning strategy for the IoT-based phytoplankton and zooplankton growth monitoring system for smart aquaculture is structured into several well-defined phases, ensuring the systematic execution and achievement of project milestones.

The project began with a research phase, starting on September 15, 2023, focusing on the in-depth exploration of the topic. This research phase is set to conclude on October 31, 2023, laying the foundation for the subsequent stages of development. Figure 1.2 shows a timeline of the project.

Simultaneously, a literature review was conducted over a longer period, extending from September 15, 2023, to April 30, 2024. This phase allows for a comprehensive review of existing research, helping identify the gaps in current knowledge and guiding the direction of the project. Following this, the requirement analysis phase ran from January 20 to May 15, 2024, during which the system's specifications were clearly defined to inform the design process.

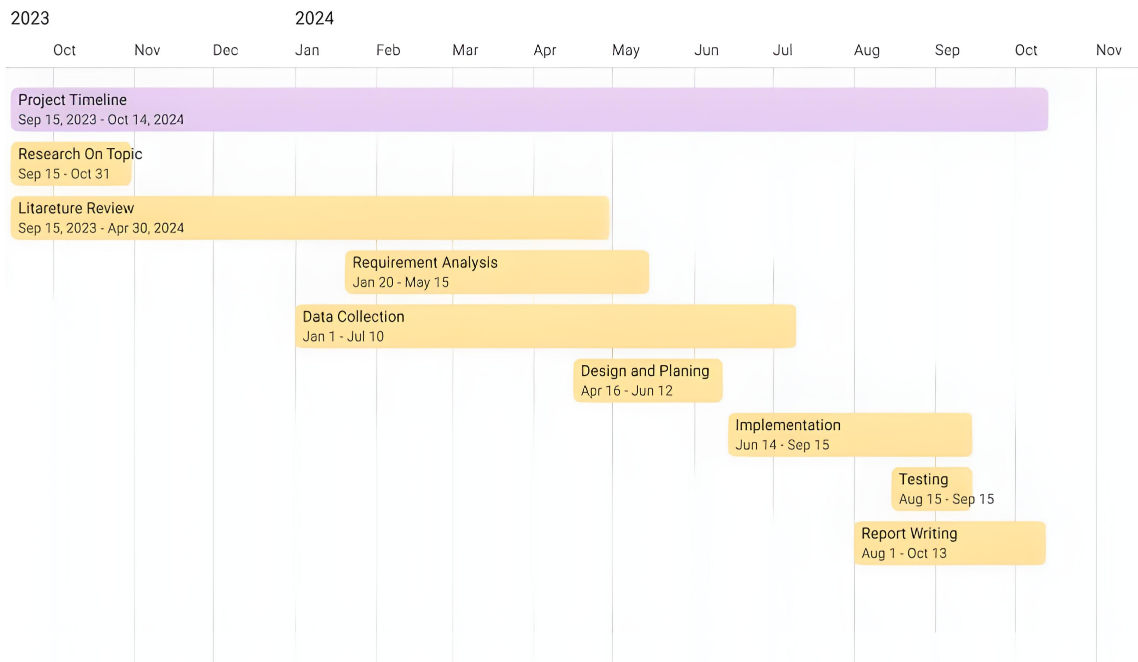


Figure 1.2: Project Timeline

The data collection phase, running from January 1 to July 10, 2024, focused on gathering essential real-world data on phytoplankton and zooplankton growth in aquaculture environments. During this period, IoT-based sensors monitor water quality and plankton metrics across various sites, ensuring diverse and reliable datasets. This data is critical for training and validating the predictive models, setting the stage for the system's accuracy and effectiveness in later phases of the project.

The design and planning phase took place from April 16 to June 12, 2024, focusing on creating an architecture for the system that integrates IoT technology and machine learning algorithms for real-time data collection and analysis. Once the design is complete, the project is expected to move into the implementation phase from June 14 to September 15, 2024, during which the system will be built and deployed.

Rigorous testing on the system will be done from August 15 to September 15, 2024, to ensure the system's reliability, accuracy, and cost-effectiveness.

Finally, the project will conclude with the report writing phase, from August 1 to October 13, 2024, synthesizing all findings and outcomes into a comprehensive document.

1.6 Report Layout

- **Chapter 1: Introduction**

  This chapter introduces the core idea behind the project, providing an overview of the importance of plankton in aquaculture. It outlines the need for real-time monitoring of plankton growth and explains the inadequacies of traditional methods. The chapter also states the problem, objectives, and significance of the study, framing the research as a solution to issues faced by fish farmers in maintaining balanced aquatic ecosystems.

- **Chapter 2: Literature Review**

  In this chapter, a comprehensive review of related literature is presented, including studies on IoT applications in aquaculture and the role of machine learning in water quality prediction. It examines the importance of plankton in fisheries, and various characteristics of plankton in aquaculture, and reviews the different machine learning models that can predict water quality and plankton abundance. This sets the foundation for the technical approaches adopted in the project.

- **Chapter 3: Design and Development**

  This chapter focuses on the project's design and development process. It covers the work breakdown structure, project management activities, and financial considerations. The selection of sensors, microcontrollers, and communication modules is discussed, alongside the system architecture. This chapter also explains how IoT devices are integrated and how the data from sensors are collected and processed.

- **Chapter 4: Result Analysis**

  This chapter presents the analysis of data collected through the IoT-based system, including plankton sampling, water quality, and environmental correlations. It evaluates the performance of various machine learning models like Support Vector Machines (SVM), Random Forest (RF), and others in predicting plankton abundance. Comparative analyses of different models' accuracy and effectiveness are also discussed, alongside the cross-validation and performance metrics for both phytoplankton and zooplankton.

- **Chapter 5: Conclusions and Future Work**

  The final chapter summarizes the research findings and discusses the impact of the project on aquaculture, including its societal, environmental, and ethical implications. It highlights the project's contributions toward sustainable aquaculture practices and presents recommendations for future research, particularly focusing on areas where the system can be improved or expanded to meet additional challenges in aquaculture.

Each chapter builds on the previous one to provide a comprehensive understanding of how the IoT-based monitoring system was developed and its impact on aquaculture management.

Chapter 2

Literature Review

2.1 Introduction

Aquaculture, a rapidly growing sector, plays a crucial role in global food production, yet maintaining water quality and plankton balance in fish farms remains a challenge. Plankton, serving as the primary food source for fish, is heavily influenced by various environmental parameters such as temperature, pH, Total Dissolved Solids (TDS), and Dissolved Oxygen (DO). Monitoring these factors traditionally involves manual methods that are prone to inaccuracies. Recent advancements in technology have introduced the Internet of Things (IoT)-based systems for real-time water quality monitoring, automating data collection, and ensuring optimal conditions. For instance, Anani et al. (2022) developed an IoT-based monitoring system to measure these parameters and ensure the sustainability of fish production, while Kumar et al. (2023)created a system specifically designed to maintain an eco-friendly environment for particular fish species. These innovations, alongside machine learning techniques for predicting plankton abundance, are significantly improving farm management practices, enhancing productivity, and ensuring sustainable fish farming operations.

2.2 IoT-based Aquaculture Monitoring

Recent advancements in aquaculture have explored various IoT-based fish farm monitoring systems that utilize water quality sensors to track key parameters like pH, Total Dissolved Solids (TDS), temperature, turbidity, and Dissolved Oxygen (DO), ensuring optimal fish farming conditions. These systems provide fish farmers with real-time data, helping them maintain water quality through mobile applications. As a result, IoT-based systems automate critical processes, minimize human error, and detect unfavorable conditions faster, ensuring a healthy environment for fish.

Several studies have demonstrated how these systems enhance fish farming operations. Saha et al. (2018), for example, designed a water quality monitoring system using Raspberry Pi, Arduino, and various sensors alongside a smartphone camera and an Android application. This setup allowed for real-time monitoring and data collection, with the system being able to detect and alert farmers about the water conditions, ensuring early intervention to maintain optimal water quality in aquaculture environments.

Similarly, Manoj et al. (2022) introduced an IoT-based Water Quality Monitoring System (WQMS) for fish ponds, which employed underwater sensors to measure water parameters such as DO, temperature, and turbidity. Their system aimed to provide farmers with an effective means to monitor and maintain the water conditions necessary for healthy fish growth. It emphasizes the integration of IoT and sensor technologies to streamline the management of fish ponds, thereby increasing productivity and reducing the risks of fish mortality caused by poor water quality.

Islam et al. (2020) developed a system that utilized wireless sensor networks (WSNs) to monitor water contamination levels. The system continuously compared collected data with predefined threshold values and made decisions about the water quality. If the quality dropped below acceptable levels, the system would immediately send notifications to the farmer's mobile app. This decision-making process not only automated water quality control but also offered an easy-to-use interface for fish farmers, ensuring quick responses to potential hazards.

In another contribution, Rashid et al. (2021) proposed an IoT-based solution designed to increase both the efficiency and productivity of aquaculture. Their system employed sensors to gather data, which was then analyzed using a machine learning model. The system incorporated Artificial Intelligence (AI) to generate decisions and provide feedback to the user, sending notifications via a mobile app. The integration of AI allowed the system to predict water quality trends and suggest preventive measures, making it an intelligent solution that significantly improved water quality monitoring processes.

Additionally, Nayan et al. (2020) designed a system to predict and detect water quality changes early, preventing potential losses. They used a Gradient Boosting Model (GBM) to forecast water quality by analyzing parameters collected from rivers across Bangladesh. The system employed automatic water parameter measuring tools, which gathered real-time data and facilitated predictive analysis to enhance water quality monitoring and early intervention.

These contributions illustrate the growing trend toward using IoT and machine learning technologies in aquaculture. Each system has successfully automated key aspects of water quality monitoring, reducing manual efforts while ensuring real-time control of water conditions. By providing accurate data and decision-making capabilities, these systems enable farmers to maintain a healthier and more sustainable aquatic environment.

2.3 Importance of Plankton in Fisheries

In aquaculture, plankton, particularly phytoplankton and zooplankton, play an essential role in ensuring the health and productivity of fish farms. Plankton forms the base of the aquatic food chain, providing nutrients and energy for fish at various stages of growth, especially during early development. Das & Sharma (2022) observed that fish cultures, particularly in lakes, depend heavily on the abundance of phytoplankton and zooplankton.

Fertilization of water bodies to encourage phytoplankton growth subsequently leads to an increase in zooplankton, which serves as natural feed for fish. The study highlighted a strong correlation between plankton abundance and fish production, suggesting that higher zooplankton and phytoplankton levels can improve fish yields and diversity in aquaculture systems.

Similarly, Lomartire et al. (2021) emphasized the significance of zooplankton as a preferred food source for fish, particularly during early life stages. The study found that long-term monitoring of zooplankton populations helps in understanding environmental conditions and predator-prey dynamics. Zooplankton not only serves as a direct food source for many fish species but also acts as a bio-indicator for water quality, influencing fish recruitment.

Several other studies have corroborated the impact of plankton on fish health. Abdel-Wahed et al. (2018) and El-Feky et al. (2019) found that plankton-rich water environments promote fish health by providing direct nourishment to planktivore fish species. Additionally, the development of phytoplankton and zooplankton in these waters supports the growth of fouling communities, which further benefits fisheries.

Araujo et al. (2022) also stressed the importance of cultivating plankton in aquaculture systems. By managing plankton populations, farmers can ensure a consistent and nutrient-rich food source for fish, which is vital for maintaining the overall productivity and sustainability of fish farms. Brander (2017) explored the relationship between plankton productivity and fishery production, noting that fish production tends to be highest in areas with significant plankton growth, such as nutrient-rich up-welling zones and shelf seas.

In conclusion, the balance and availability of phytoplankton and zooplankton are vital to the health and productivity of aquaculture systems. Their role as a primary food source and bio-indicator makes them essential components for successful fish farming, and monitoring their levels can greatly enhance fish yield and environmental health.

2.4 Characteristics of Plankton in Aquaculture

To further understand plankton population dynamics, several studies have explored the relationship between zooplankton and various environmental parameters, highlighting the importance of continuous monitoring in aquaculture settings. One notable investigation is by Odulate et al. (2017), which studied the water quality of the lower Ogun River with plankton abundance and diversity. This study showed highly significant correlations ($P \leq 0.05$)between plankton abundance, diversity, and physico-chemical parameters monitored during the study period.

The findings emphasize that the abundance and diversity of plankton are directly influenced by factors such as pH, DO, and transparency. Continuous monitoring of these water quality parameters is vital for maintaining the delicate balance of the ecosystem, ensuring that plankton populations are sustained and fish farms remain productive.

In a similar vein, Musa et al. (2022) analyzed the relationship between phytoplankton availability and the growth rate of Vannamei shrimp in response to water quality changes. The study employed both Principal Component Analysis (PCA) and Canonical Correspondence Analysis (CCA) to assess the impact of environmental parameters on phytoplankton abundance. Results demonstrated that pH, nitrate, and total organic matter (TOM) significantly influenced phytoplankton populations.

These findings indicate that water parameters can either promote or hinder the growth of phytoplankton, which serves as a critical food source for shrimp. By incorporating secondary water quality data into the PCA, the study provided a comprehensive understanding of how these factors interact, further emphasizing the need for precise control of water quality in aquaculture environments.

Khan et al. (2023) conducted a study in a eutrophic fish pond in Bangladesh to analyze the seasonal dynamics of zooplankton with environmental factors. Pearson's correlation coefficient analysis showed a positive correlation between total zooplankton abundance and pH, while negative correlations were observed with transparency, dissolved oxygen, phosphates, nitrates, and temperature. The Canonical Correspondence Analysis (CCA) further illustrated how environmental parameters like pH, transparency, and temperature significantly influenced the distribution and abundance of zooplankton groups. This highlights the importance of understanding seasonal and environmental fluctuations, as such knowledge can help fish farmers adapt their practices to maintain optimal zooplankton levels, thereby sustaining the health of aquatic ecosystems and improving fish production.

Smith & Piedrahita (1988) investigated the relationship between phytoplankton density and dissolved oxygen levels at dawn, proposing two mathematical models to describe this relationship. An analytical model predicted the upper limit of oxygen production by phytoplankton, while an empirical model, based on data from channel catfish ponds, helped quantify the interaction between these factors. This research is particularly relevant to fish farming, as it provides insight into how phytoplankton density affects oxygen levels, which are critical for maintaining a healthy aquatic environment. The study's findings support the idea that effective monitoring of DO and phytoplankton densities is essential for optimizing fish pond conditions.

Abdel-Wahed et al. (2018) explored the relationship between water quality, plankton abundance, and fish growth performance. Their findings indicated that water quality parameters like TDS, pH, and temperature have a direct impact on the growth rates of plankton, which in turn influence fish development. By monitoring these parameters, fish farmers can optimize the conditions required for both plankton and fish to thrive. This study aligns with the broader body of research that stresses the importance of maintaining water quality to ensure the sustainability of aquaculture operations.

Ding et al. (2018) investigated the contribution of internal phosphorus (P) loading to nitrogen (N) limitation in aquaculture systems. Their study found that internal phosphorus loading played a significant role in causing nitrogen limitation during the pre-bloom period, affecting the growth of phytoplankton. This finding is particularly relevant for fish farms where nutrient management is crucial for maintaining the balance between plankton populations and water quality.

Rui & Qiu-jin (2013) demonstrated that during the logarithmic phase of phytoplankton growth, increases in total dissolved solids (TDS) were positively correlated with increases in phytoplankton growth rate, biomass, and chlorophyll-a concentration. These results emphasize that maintaining appropriate TDS levels is essential for promoting phytoplankton growth, which is critical for supporting the aquatic food chain in fish farms. High TDS levels not only improve phytoplankton productivity but also enhance the overall stability of the pond ecosystem.

Finally, Pulsifer & Laws (2021) analyzed the temperature dependence of growth rates and grazing rates of both phytoplankton and zooplankton. Their research revealed virtually identical temperature dependencies, as predicted by the metabolic theory of ecology. This study highlights the role of temperature as a major environmental factor regulating the balance between phytoplankton growth and zooplankton grazing. By understanding these dynamics, fish farmers can better manage water temperature to ensure the sustainability of their aquaculture systems.

The study by Kumar et al. (2023) focused on the diversity and seasonal density variation of plankton in the Yamuna River, particularly in Auraiya district, Uttar Pradesh. The researchers monitored the river's plankton over multiple seasons, identifying different groups, including protozoa, rotifers, crustaceans, and insect larvae. Their findings showed seasonal fluctuations influenced by environmental factors, with higher zooplankton density in winter and summer compared to monsoon and autumn. The study highlighted how these variations are linked to changes in abiotic factors like temperature and water flow, which affect species succession and ecosystem dynamics.

In conclusion, the extensive body of research reviewed here underscores the critical importance of monitoring various environmental parameters, including pH, DO, TDS, temperature, and turbidity, to optimize plankton populations in aquaculture. Through real-time monitoring systems and advanced analytical techniques such as PCA and CCA, fish farmers can better understand the complex interactions between water quality and plankton abundance, enabling more effective management of aquatic ecosystems.

2.5 Machine Learning in Water Quality and Aquaculture Prediction

The integration of Machine Learning (ML) techniques into water quality prediction systems has significantly enhanced the accuracy and efficiency of environmental monitoring. The rise of Internet of Things (IoT) technologies further amplifies these advancements, allowing for real-time data collection and analysis in water bodies, particularly in aquaculture and public health sectors. Predicting water quality using models like Support Vector Machines (SVM), Random Forest (RF), Decision Trees (DT), and ensemble methods has emerged as an essential method to ensure sustainable water resources management. This literature review examines several recent studies that apply ML algorithms to predict various water quality parameters, focusing on their performance and utility in different settings.

Wang et al. (2023) conducted an extensive study comparing various machine learning models for predicting water quality parameters. They utilized algorithms like Support Vector Machines (SVM), Multilayer Perceptron (MLP), Random Forest (RF), XGBoost, and Long Short-Term Memory (LSTM). Their results showed that SVM performed exceptionally well, especially in predicting Dissolved Oxygen (DO), due to its strong generalization capabilities. MLP, known for its strength in modeling nonlinear relationships, also performed effectively across various water quality parameters. Interestingly, the study found that RF and XGBoost, which are typically used for their robustness in classification tasks, demonstrated relatively lower prediction accuracy for water quality in this specific study. However, LSTM, a recurrent neural network (RNN), excelled in processing time-series data, providing stable and accurate predictions for water quality parameters by capturing dynamic patterns in the data.

Rai et al. (2024) designed an IoT-based system that addressed existing water quality monitoring challenges, providing a modern alternative to conventional water monitoring systems. Their work focused on utilizing machine learning techniques such as XGBoost, Random Forest, AdaBoost, and Decision Tree models. These methods were employed to predict water quality parameters and exhibited strong predictive performance, reinforcing the idea that IoT combined with ML can offer powerful solutions for continuous water quality monitoring.

Yan et al. (2024) presented a comprehensive study on water quality prediction in coastal areas, focusing on challenges and advancements in using machine learning. The authors highlighted how algorithms like Random Forest, SVM, and Decision Tree were effectively applied to predict critical water quality parameters, such as chlorophyll-a, salinity, and DO. Despite the promising results, the study also addressed the limitations in improving the overall Water Quality Index (WQI) through these methods, emphasizing the ongoing need for reducing model uncertainties and refining ML architectures for more accurate predictions.

Sakaa et al. (2022) constructed a hybrid artificial intelligence model, combining Sequential Minimal Optimization-Support Vector Machine (SMO-SVM) with Random Forest (RF) as a benchmark for predicting water quality at the Wadi Saf-Saf river basin in Algeria. By using 15 physicochemical parameters, the study demonstrated that RF improved predictive accuracy while reducing the number of input variables. This hybrid approach showed that combining SVM and RF models could offer significant advantages for environmental monitoring systems.

In another study Akhlaq et al. (2024) address the challenges of monitoring water quality in glacial lakes and rivers due to contamination by heavy metals and other pollutants. The study applies models such as decision trees, KNN, MLP, SVM, and random forest, using the Boruta algorithm to select critical features. Random forest and SVM achieved the highest accuracy (88 percent) in predicting water quality. The findings highlight the need for real-time machine learning systems for better water management and environmental protection.

Nasir et al. (2022) examined various machine learning classifiers, including SVM, RF, Logistic Regression (LR), Decision Tree (DT), CATBoost, XGBoost, and Multilayer Perceptron (MLP), for classifying water quality data. The study also introduced stacking ensemble models to improve prediction accuracy. Among the models tested, CATBoost achieved the highest accuracy at 94.51%, followed by Random Forest and SVM, demonstrating the effectiveness of ensemble methods for water quality prediction.

Abbas et al. (2024) focused on water quality prediction using an IoT-based system, integrating AI forecasting models to enhance prediction accuracy while minimizing errors. The study employed multiple ML classifiers, with RF and Gradient Boosting models achieving the highest accuracy rates (95% and 96%, respectively). SVM followed closely at 92%, while KNN and Decision Trees lagged behind. This study highlighted the potential of combining IoT and ML to improve real-time water quality monitoring in aquaculture and other settings.

Nayan et al. (2024) explored the application of Long Short-Term Memory (LSTM) models in analyzing IoT-collected data for water quality prediction. The LSTM model achieved an average R^2 score of 93.33%, demonstrating its ability to provide accurate real-time forecasts. This study emphasizes the potential of LSTM for time-series prediction in dynamic environments like aquaculture, where continuous monitoring is crucial for maintaining optimal water conditions.

Hemal et al. (2024) introduced AquaBot, an IoT-based system designed for automated water quality monitoring in aquaculture. The system utilized ML algorithms, including RF, SVM, DT, KNN, and LR to evaluate water quality parameters in real time. Among these, Random Forest and the custom ensemble model provided the highest accuracy, suggesting that ensemble techniques offer robust solutions for water quality prediction. The application of machine learning in water quality prediction has evolved significantly with the integration of IoT technologies.

Abirami et al. (2023) addresses the problem of predicting water quality using machine learning techniques. It applies algorithms such as Random Forest, SVM, and Decision Tree to classify water quality based on various parameters in synthetic and real datasets. The study found that non-parametric models, especially Random Forest, delivered higher accuracy, highlighting the potential for improved automation in water quality management.

Models such as Random Forest, SVM, and LSTM have proven effective in predicting water quality parameters, with each model excelling under different conditions. While RF and SVM offer robust performance in classification tasks, LSTM has demonstrated exceptional abilities in handling time-series data. Hybrid and ensemble models further enhance prediction accuracy by combining the strengths of multiple algorithms. These advancements, particularly in IoT-based systems, hold great promise for improving water quality monitoring and ensuring sustainable water management practices across various sectors, including aquaculture and environmental protection.

Table 2.1: Comparative Analysis

Refs.	Parameters Considered *<pH, TDS, Temp, Turb, DO >*	Data Collection	Feature Selection	Multi-class	Plankton Abundance Analysis	Plankton Growth Detection	Cost Effective Solution
Odulate et al. (2017)	yes, yes, yes, yes, yes	yes	N/A	N/A	PCC	no	no
Anani et al. (2022)	yes, no, yes, no, yes	no	no	no	N/A	N/A	yes
Rajib et al. (2018)	yes, no, yes, no, no	yes	no	no	N/A	N/A	yes
Kumar et al. (2023)	yes, no, yes, no, yes	no	no	no	N/A	N/A	no
Hemal et al. (2024)	yes, yes, yes, yes, yes	yes	yes	yes	N/A	N/A	yes
Musa et al. (2022)	yes, no, yes, yes, no	yes	N/A	N/A	PCA, CCA	no	no
Khan et al. (2023)	yes, no, yes, yes, yes	yes	N/A	N/A	PCC, CCA	no	no
Rashid et al. (2021)	yes, yes, yes, no, no	yes	yes	no	N/A	N/A	yes
Nayan et al. (2021)	yes, no, yes, no, yes	yes	yes	yes	N/A	N/A	no
Aravinth et al. (2023)	yes, yes, yes, yes, yes	yes	no	no	Correlation	no	no
Duré et al. (2021)	yes, yes, yes, yes, yes	yes	yes	no	PCA	no	no
Proposed System	yes, yes, yes, yes, yes	yes	yes	yes	Correlation	yes	yes

Chapter 3

Design and Development

3.1 Project Management and Financial Activities

Project management is essential for the successful execution of projects by providing structure, direction, and control. It ensures clear goals and alignment among stakeholders, promotes efficient resource allocation, and mitigates risks to keep the project on track. Through effective scheduling and quality assurance, project managers maintain timelines and ensure high standards are met. Additionally, project management frameworks handle changes efficiently without disrupting progress. Delivering within scope, time, and budget, ensures stakeholder satisfaction and builds trust. Meanwhile, financial operations like budgeting play a crucial role in maintaining the project's long-term sustainability. It also enables better decision-making, supports risk mitigation, and ensures the project can adapt to unforeseen challenges while remaining financially viable. This section addresses how project management and finance strategies operate within the framework of an IoT-based phytoplankton and zooplankton growth monitoring system for smart aquaculture.

3.1.1 Work Breakdown Structure(WBS)

The Work Breakdown Structure (WBS) in Figure 3.1 outlines the major tasks involved in executing a project, providing a clear and organized view of each phase. It helps in managing the project from beginning to end. It begins with Project Control, which covers planning, requirement analysis, supply chain management, and cost estimation to ensure the project starts on a solid foundation.

The Design phase involves creating IoT device circuits, integrating hardware with cloud platforms and software, and developing the overall software architecture. Following this is Hardware Development, which includes building the physical components, conducting performance tests, and ensuring durability under various conditions. In parallel, Software Development focuses on creating the user interface, implementing the software, and ensuring its performance meets desired benchmarks.

The Data Collection and Processing phase emphasizes gathering data both from physical samples and IoT devices, preparing datasets, and training, testing, and deploying machine learning models for data analysis. Finally, Project Management ensures smooth operations by focusing on governance, ongoing maintenance, and quality control to maintain high standards throughout the project. This WBS serves as a structured guide for efficiently managing all aspects of a project, likely focusing on an IoT-based solution that involves both hardware and software components working together seamlessly.

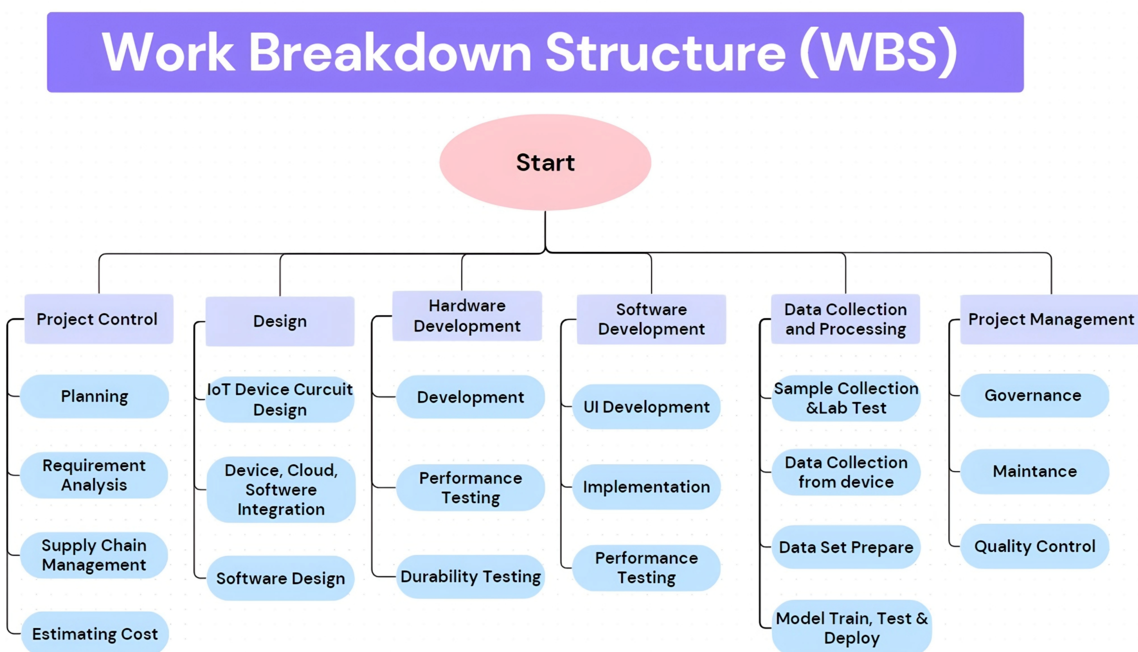


Figure 3.1: Work Breakdown Structure

3.1.2 Budget Plan

Budgeting: The budget for the IoT-based phytoplankton and zooplankton growth monitoring system for smart aquaculture project is structured to ensure efficient resource management and to cover all aspects, from research to deployment, in a cost-effective manner. In Table 3.1 a comprehensive, cost-effective, and efficient budget allocation has been shown.

Table 3.1: Cost Estimation Table

Sector	**Items**	**Cost**
Research, Laboratory, Utility	Sample Collection Instrument	4000TK
	Sample Logistical Bills	3000TK
	Lab Testing	15000TK
Hardware Procurement	pH Sensor	2500TK
	Temperature Sensor	200TK
	Total Dissolve Solid (TDS) Sensor	450TK
	Turbidity Sensor	900TK
	Arduino Uno	1000TK
	WiFi Module	300TK
	Bread Board	200TK
	Wires & Register	200TK
	Battery	200TK
Software Development	Backend Server Mobile Application	200,000TK (est.)
Deployment	Backend Server Hosting and Mobile Application Release (1 year)	20000TK (est.)
	IoT Hardware Cloud Integration (1 year)	5000TK (est.)
	Maintenance	100,000TK
Total		352,950 TK

In the Research, Laboratory, and Utility section, the budget includes expenses for sample collection instruments, logistical arrangements for transporting and managing samples, and lab testing to analyze water quality. These are essential for gathering baseline data, and the total cost of 22,000 TK ensures that the research phase is well supported. The testing and analysis will help fine-tune the sensors and the system as a whole. Hardware Procurement involves the acquisition of key components that form the backbone of the IoT system. This includes various sensors such as pH, temperature, and total dissolved solids (TDS) sensors, as well as other essential electronics like the Arduino Uno, WiFi modules, and breadboards. These components are crucial for the system to monitor water quality in real time. The total cost of 5,950 TK reflects a focus on affordability while ensuring that high-quality, reliable equipment is used. The sensors are expected to collect the necessary data for the machine learning models and analytics to make accurate predictions about water quality.

The budget for Software Development is significantly larger, estimated at 200,000 TK. This reflects the complexity and importance of developing both the backend server and the mobile application. The backend server will handle data from the IoT sensors, process it, and store it for further analysis, while the mobile application will provide users with real-time insights into water quality. The server will be designed for scalability, ensuring that as more sensors are added or more data is processed, the system will continue to perform efficiently. Finally, an allocation of 100,000 TK from the budget has been reserved for maintenance activities. This covers costs related to system monitoring, regular upkeep, and continuous support for the deployed solution, ensuring it remains fully operational and efficient in monitoring water quality parameters in aquaculture ponds. This budget ensures the system's long-term stability and effectiveness over time.

Deployment costs are primarily concerned with ensuring the system is up and running for the first year. This includes hosting the backend server and releasing the mobile application, with an estimated cost of 20,000 TK. Additionally, cloud services for integrating the IoT hardware with platforms like Amazon IoT and Digital Ocean will require an estimated 5,000 TK for the first year. These cloud services are vital for storing, managing, and analyzing the large volumes of data generated by the sensors, while also enabling the deployment of machine learning models for real-time predictions. The use of cloud services ensures reliability, scalability, and secure data management without the need for extensive on-premises infrastructure.

In total, the project is estimated to cost 352,950 TK, encompassing research, hardware procurement, software development, and deployment. This carefully planned cost-effective budget ensures that all necessary components are covered, from the early research and hardware acquisition to the final deployment of the IoT system and cloud integration.

3.2 Requirement Analysis

Requirement analysis is the process of identifying, documenting, and validating the needs of stakeholders to ensure a project aligns with both business goals and user expectations. This process plays a critical role in reducing risks, preventing costly rework, and improving communication among stakeholders. It ensures that the project scope is clear, requirements are actionable, and outcomes are aligned with objectives, leading to higher stakeholder satisfaction and better project management Reqtest (2018).

For IoT based Phytoplankton and Zooplankton Growth Monitoring Systems we have done requirement analysis and have identified the hardware and software requirements.

3.2.1 Hardware Requirements

- **pH Sensor:** A water pH sensor is a device designed to measure the acidity or alkalinity of water, providing real-time data essential for various industries. These sensors use a glass electrode and a reference electrode to detect the concentration of hydrogen ions in the solution, converting this into a pH value. Proper pH levels help safeguard aquatic life, maintain water quality, and prevent equipment corrosion in water systems Meacon Co., Ltd. (2023).

- **Temperature Sensor:** A water temperature sensor is a device used to measure the temperature of water in various environments, such as aquaculture systems, laboratories, and industrial processes. It operates by detecting changes in temperature through properties like electrical resistance or voltage shifts. These sensors are essential in maintaining water quality by ensuring optimal temperatures for aquatic organisms or chemical processes Atlas Scientific (2023).

- **Turbidity Sensor:** A turbidity sensor is a device used to measure the clarity or cloudiness of water, which results from the presence of suspended particles such as sediments, pollutants, or microorganisms. It works by emitting a light beam into the water and measuring how much light scatters when it encounters particles. Higher turbidity indicates more scattered light, suggesting poorer water quality. These sensors are essential for environmental monitoring, water treatment, and pollution control as they provide real-time data to detect contamination, and risks to aquatic ecosystems or public health Sensorex (n.d.).

- **TDS (Total Dissolved Solids) Sensor:** TDS (Total Dissolved Solids) sensor measures the concentration of dissolved substances, such as minerals, salts, and organic compounds, in water. This measurement is crucial for assessing water quality, as high TDS levels can affect taste, safety, and the performance of aquatic systems. The sensor works by gauging the electrical conductivity of water, since dissolved solids allow the current to flow more easily. It is often used in water treatment, aquariums, and environmental monitoring to ensure the water is within safe and acceptable TDS ranges for human or aquatic use AquaHow (2023); Osmotics (2023).

- **Arduino UNO:** The Arduino UNO is a microcontroller board based on the ATmega328 microcontroller, widely used for prototyping, education, and IoT applications. It features 14 digital input/output pins, 6 analog inputs, and supports serial communication via USB. Arduino UNO is valued for its ease of use, with support from the open-source Arduino IDE, which simplifies coding and uploading programs. Its versatility makes it suitable for applications ranging from robotics and home automation to sensor-based data logging and product development Hack the Developer (n.d.).

- **ESP8266 ESP-01 WiFi Module:** The ESP8266 ESP-01 is a compact and low-cost Wi-Fi module designed for providing wireless connectivity to electronic projects. Equipped with a built-in TCP/IP stack, this module allows microcontrollers to connect to Wi-Fi networks, enabling the transmission of data over the internet. It is particularly popular in Internet of Things (IoT) application Components101 (2024).

3.2.2 Software Requirements

- **Arduino IDE:** The Arduino Integrated Development Environment (IDE) is a software platform designed for programming Arduino microcontrollers. It simplifies the coding process by providing a user-friendly interface that allows users to write and upload code to their Arduino boards easily. The IDE supports programming in a language based on C/C++, making it accessible for beginners and experienced programmers alike. Its purpose is to facilitate the development of a wide range of applications, from basic projects to complex embedded systems, by offering tools such as syntax highlighting, code suggestions, and an integrated debugger Arduino (n.d.); Control.com (2022).

- **Google Colab:** Google Colab, short for Colaboratory, is a free cloud-based platform developed by Google that allows users to write and execute Python code within a Jupyter notebook environment. It is particularly advantageous for tasks related to machine learning and data analysis, as it provides free access to powerful hardware resources, including GPUs and TPUs. The platform also supports collaborative coding, enabling multiple users to work on the same notebook simultaneously, similar to Google Docs. Additionally, Colab's integration with Google Drive enhances accessibility, making it easy to save and share notebooks with others Marqo (2024); Google Research (2023).

- **Flutter & .NET CORE:** Flutter and .NET Core are widely used frameworks for application development. Flutter, developed by Google, is a UI toolkit that enables developers to create natively compiled applications for mobile, web, and desktop platforms from a single codebase Karunasena (2023). Conversely, .NET Core is a cross-platform, open-source framework developed by Microsoft, intended for building modern web applications, cloud-based services, and Internet of Things (IoT) solutions. It supports multiple programming languages, such as C#, F#, and VB.NET, allowing developers to create applications that run seamlessly across various platforms, including Windows, macOS, and Linux TrustRadius (2024).

- **Cirkit Designer:** Circuit Designer software is a tool used for creating and simulating electronic circuits, allowing users to visualize and analyze circuit designs before implementation. Its primary purpose is to facilitate the design process by providing a user-friendly interface where users can draft schematics, run simulations, and prepare layouts for printed circuit boards (PCBs). This software is beneficial for engineers, hobbyists, and students, enabling them to experiment with circuit configurations and improve their designs efficiently Altium (2023).

- **Amazon IoT Cloud server:** Amazon IoT Cloud Server, specifically AWS IoT Core, is a platform that enables secure and scalable connections between Internet of Things (IoT) devices and cloud applications. Its primary purpose is to facilitate the management and integration of vast networks of IoT devices, allowing for data collection, processing, and real-time analytics. AWS IoT Core supports various communication protocols and offers features like device authentication, data routing, and a rules engine, which collectively enhance the operational efficiency of IoT applications while ensuring security and ease of use Amazon Web Services (n.d.).

- **Digital Ocean:** DigitalOcean is a cloud infrastructure provider that focuses on simplifying the deployment, management, and scaling of applications. It offers a user-friendly interface designed for developers, enabling businesses—ranging from startups to large enterprises—to efficiently utilize cloud resources. The platform is suitable for tasks such as launching websites and running scalable applications. Its emphasis on performance and ease of use has established DigitalOcean as a popular choice for those looking to harness cloud computing capabilities Howtodojo (n.d.).

- **LaTex:** LaTeX is a typesetting system commonly used for producing scientific and mathematical documents due to its ability to handle complex formulas and structures efficiently. It allows users to create high-quality documents with precise formatting, making it a preferred choice in academia for articles, theses, and books. The primary purpose of LaTeX is to provide a consistent layout while enabling advanced typographical features, thereby enhancing the readability and professionalism of documents Overleaf (n.d.).

- **Microsoft Excel:** Microsoft Excel is a widely-used spreadsheet software developed by Microsoft that enables users to organize, analyze, and visualize data efficiently. Its primary purpose is to facilitate financial analysis, budgeting, and forecasting, making it essential for professionals in finance and accounting. Excel offers various features, including data entry, management, and powerful functions for calculations and data manipulation Corporate Finance Institute (2024).

- **LucidChart:** Lucidchart is a web-based diagramming application that allows users to create flowcharts, organizational charts, wireframes, mind maps, and other types of visual representations. It is widely used for its collaborative features, enabling teams to work together in real time on diagrams and designs. Lucidchart integrates with various other software tools, making it a versatile choice for project management, software development, education, and more Lucid Software Inc. (n.d.).

3.3 Adopted Tools and techniques

This section describes the selection criteria for required tools such as Sensors, microcontrollers, and Communication Modules for the IoT-based phytoplankton and zooplankton growth monitoring system for smart aquaculture.

3.3.1 Sensor Selection

In selecting sensors for an IoT-based phytoplankton and zooplankton growth monitoring system, the primary criteria included accuracy, sensitivity, and compatibility with water parameters like pH, temperature, turbidity, and total dissolved solids (TDS). Sensors must provide real-time data with minimal lag to ensure timely monitoring of environmental conditions. Power efficiency and wireless communication capabilities were also considerable criteria. Additionally, cost-effectiveness and ease of maintenance were considered especially for large-scale or long-term deployment in aquaculture systems.

pH sensor:

- **Model Name:** SEN-00239 pH Sensor
- **Measurement Range:** 0 to 14 pH
- **Accuracy:** ± 0.1 pH (at 25°C)
- **Selection Rationale:** The SEN-00239 pH sensor is ideal for aquaculture due to its wide measurement range (0-14 pH) and high accuracy (±0.1 pH), ensuring reliable water quality monitoring. It's durable, designed for long-term submersion, and offers real-time data, making it perfect for IoT-based systems. Its compatibility with Arduino simplifies integration into smart aquaculture setups.

(a) pH Sensor

(b) Tempareture Sensor

Figure 3.2: pH and Temperature sensor

Temperature sensor:

- **Model Name:** SEN-00072 Digital Temparature Sensor
- **Measurement Range:** -55°C to +125°C
- **Accuracy:** ±0.5°C (from -10°C to +85°C)
- **Selection Rationale:** The SEN-00072 digital temperature sensor is ideal for aquaculture monitoring due to its wide temperature range and reliable accuracy of ±0.5°C, ensuring precise water temperature tracking for optimal plankton growth. It is waterproof, durable, and designed for long-term use in aquatic environments, making it suitable for continuous IoT-based monitoring. It can be easily integrated with microcontrollers like Arduino UNO.

(a) Turbidity Sensor

(b) TDS Sensor

Figure 3.3: Turbidity and TDS sensor

Turbidity sensor:

- **Model Name:** SEN-00179 Analog Turbidity Sensor
- **Measurement Range:** 0 to 4000 NTU (Nephelometric Turbidity Units)
- **Accuracy:** ±5% of measured value
- **Selection Rationale:** The SEN-00179 analog turbidity sensor is well-suited for aquaculture systems due to its wide measurement range, covering low to high turbidity levels essential for monitoring water clarity and plankton growth. Its ±5% accuracy ensures reliable detection of suspended particles in water. The sensor is easy to integrate with microcontrollers for real-time IoT applications.

TDS sensor:

- **Model name:** SEN-00222 Analog TDS Sensor
- **Measurement Range:** 0 to 1000 ppm (parts per million)
- **Accuracy:** ±10% of measured value
- **Selection Rationale:** The SEN-00222 Analog TDS sensor is highly suitable for aquaculture as it effectively monitors total dissolved solids (TDS), which are crucial for water quality and plankton health. Its 0-1000 ppm range is suitable for most aquaculture water conditions, and its ±10% accuracy provides reliable measurement of dissolved substances. The sensor is durable and easy to integrate with microcontrollers.

3.3.2 Microcontroller and Communication Modules Selection

The primary selection criteria for choosing a microcontroller and WiFi module included compatibility, ensuring seamless integration between both components and processing power, where the microcontroller should handle the required tasks efficiently. Power consumption is crucial for energy efficiency, especially in IoT systems. Connectivity range and data transmission speed were key factors for the choosing WiFi module, ensuring stable and fast communication. Additionally, ease of programming, available libraries, and cost-effectiveness were also considered for smooth development and deployment in IoT applications.

- **Arduino UNO:** The Arduino UNO was an ideal choice for IoT IoT-based phytoplankton and zooplankton growth monitoring system due to its simplicity, versatility, and ease of use. It offered sufficient processing power and input/output pins to handle multiple sensors, such as those for pH, temperature, and turbidity, making it suitable for real-time environmental monitoring. Its wide range of libraries and community support enabled quick integration with various IoT components, including WiFi modules like the ESP8266. Additionally, the Arduino UNO is cost-effective, making it a practical option for scalable smart aquaculture applications.

- **ESP8266 ESP-01 WiFi Module:** The ESP8266 ESP-01 WiFi module was ideal for IoT-based phytoplankton and zooplankton growth monitoring system due to its low power consumption, affordable cost, and reliable wireless connectivity, enabling real-time data transmission from sensors monitoring plankton growth parameters. Its compact size and built-in TCP/IP stack made it efficient for IoT applications, allowing seamless communication between the monitoring system and remote servers or mobile devices. Additionally, its compatibility with Arduino UNO ensured easy integration, while its ability to support IoT protocols makes it well-suited for continuous environmental monitoring in smart aquaculture systems.

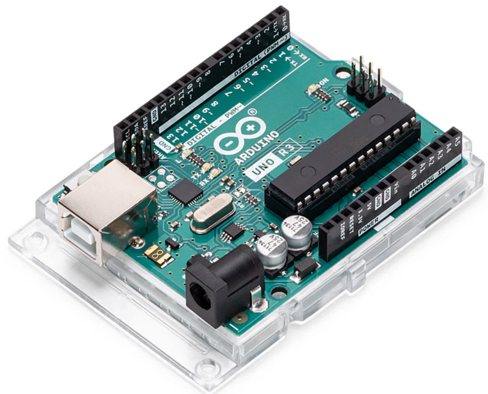


Figure 3.4: Arduino UNO

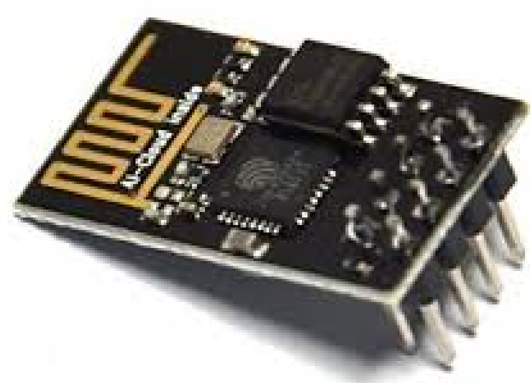


Figure 3.5: ESP8266 ESP-01 WiFi Module

3.4 Design the Solution

3.4.1 Innovative Aquaculture Solutions

- **Concept Generation:** In the ideation phase, the team explored various strategies to improve the monitoring of plankton growth in aquaculture. The initial concepts spanned from conventional sensor-based water quality measurements to advanced machine learning models for predictive analysis. The core objective was to develop a solution that could provide real-time monitoring and forecasting of plankton levels, offering fish farmers timely insights. The system needed to be flexible enough to adapt to different aquatic conditions while delivering meaningful data on plankton abundance.

- **Concept Reduction:** After a thorough evaluation focusing on criteria like accuracy, scalability, cost, and flexibility, the team narrowed down the options. Traditional methods such as Secchi disks, though considered, could not predict plankton population dynamics based on changing environmental factors. The final decision was to adopt an IoT-based approach combined with machine learning algorithms. This choice offered a more dynamic solution by predicting plankton levels through water quality indicators like pH, temperature, turbidity, and dissolved oxygen, thereby improving the management of fish farms.

- **Justification and Novelty:** The innovation of this solution lies in its combination of IoT devices and machine learning techniques, particularly the use of XGBoost, to predict plankton growth. This marks a departure from traditional manual methods, introducing the advantage of predictive capabilities. The system not only provides real-time insights but also forecasts future plankton levels, enabling proactive measures. This leads to better aquaculture management by helping farmers avoid critical issues such as oxygen depletion and harmful algal blooms, ultimately boosting fish production and profitability.

3.4.2 Methodology

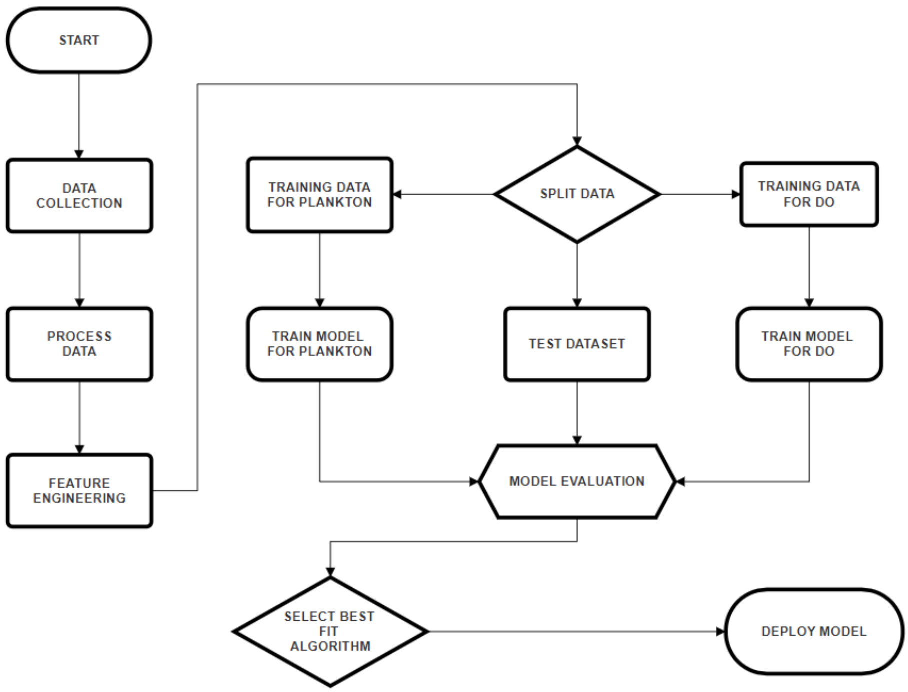


Figure 3.6: Methodology

The flowchart 3.6 shown in the figure outlines the methodology of IoT based Phytoplankton and Zooplankton Growth Monitoring System for Smart Aquaculture. It begins with data collection, followed by data processing and feature engineering to prepare the data. After the data is split into subsets for training and testing, machine learning models such as Random Forest, XGBoost, Gradient Boosting, KNN, CNN, RNN, and SVM are trained for Plankton and DO using their respective training datasets. After training, the models are evaluated using the test dataset, ensuring they perform well. The results are fed into the model evaluation step, where the performance of ML models is assessed. The best-performing algorithm which is XGBoost is then selected through the best fit algorithm step, and finally, the most suitable model will be deployed in the mobile application for use.

3.4.3 System Architecture

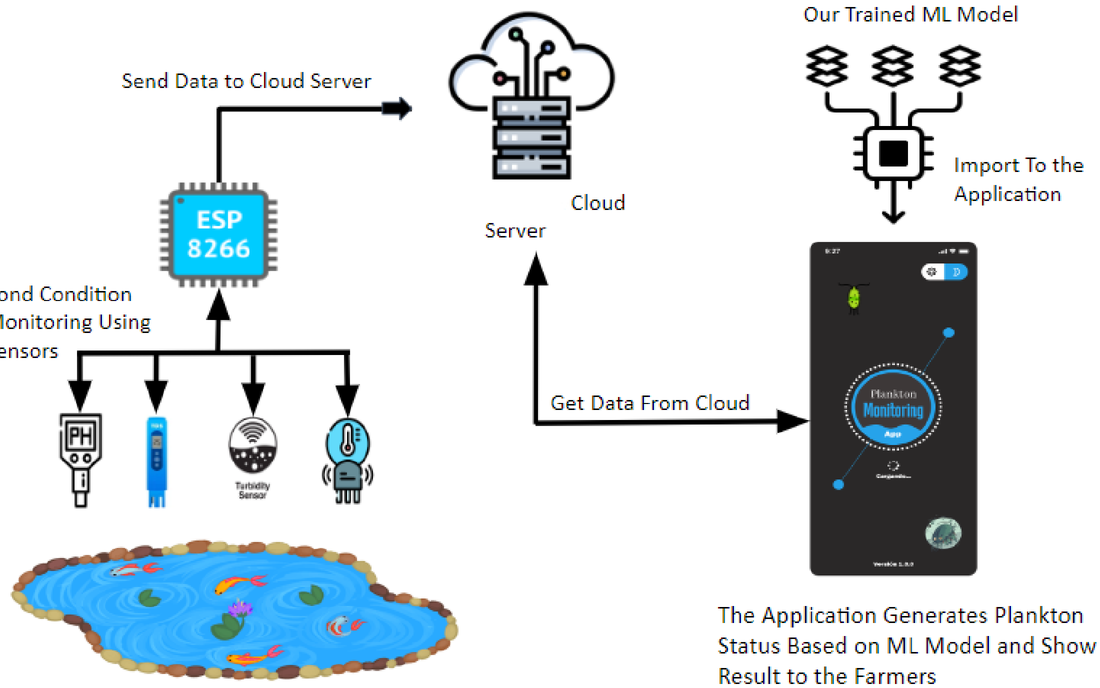


Figure 3.7: System Architecture

3.4.3.1 System Architecture Overview

The architecture of the IoT-based plankton growth monitoring system integrates sensor-based data collection, cloud computing, and machine learning to offer real-time pond condition analysis and plankton status monitoring. As illustrated in the image, the system starts by deploying multiple sensors in the pond to capture key environmental parameters essential for aquatic life. These sensors measure critical factors such as pH levels, water temperature, turbidity, and total dissolved solids (TDS). The collected data is then transmitted via an ESP8266 microcontroller, which serves as the communication bridge between the sensors and the cloud server.

Once the data is collected, it is sent to the Firebase Cloud server, where it is securely stored and processed. The system's machine learning component is housed in the cloud. This component involves a pre-trained machine learning (ML) model, which has been trained using various supervised algorithms to analyze the environmental data and predict the growth and status of plankton in the pond. The selected ML model, based on its optimal performance in terms of accuracy, recall, and other metrics, is utilized to predict whether the current pond conditions support healthy plankton growth or signal an imbalance.

The processed results from the ML model are sent back to the cloud server, which is then fetched by a dedicated mobile application. Farmers can access the application to view real-time insights, such as the status of plankton growth, pond conditions, and necessary corrective actions if required. The app allows farmers to make informed decisions and manage their ponds efficiently. The integration of cloud computing, machine learning, and sensor networks ensures continuous data flow, precise monitoring, and user-friendly access through the mobile platform.

3.4.3.2 Communication Protocols and Data Flow

The communication within the system architecture is engineered for efficient and secure data transmission. The IoT device employs the ESP8266 WiFi module to wirelessly transmit sensor data from the pond to the cloud server. Various sensors, including pH, temperature, turbidity, and dissolved oxygen sensors, continuously monitor pond conditions, ensuring real-time data collection.

The collected data is then sent to the Amazon IoT Cloud server, where it is processed and stored. The server integrates a trained machine learning (ML) model, which is used to analyze the data and predict plankton growth. This processed information is sent to the mobile application, where farmers can view the plankton growth status based on the ML model's predictions.

The server-to-application communication is secured with robust HTTP protocols, ensuring data integrity and protection against cyber threats. This end-to-end system ensures reliable monitoring of pond conditions and timely updates for farmers, enabling more effective aquaculture management.

3.4.4 Use Case Diagram

3.4.4.1 General User Use Case Diagram

The use case diagram shown in the figure 3.8 represents a detailed interaction framework for an IoT-based Phytoplankton and Zooplankton Growth Monitoring System, highlighting how a general user engages with various system functions. The process begins with the user creating an account, which allows them to sign in and access the system's features. Once logged in, the user can view and edit their profile, enabling them to manage personal information and preferences. A key functionality is the ability to view real-time pond status, where IoT sensors provide live data on the growth and conditions of phytoplankton and zooplankton in the monitored environment. This allows the user to track ecological trends and assess the health of the pond. Additionally, the system will send notifications to the user, alerting them to any critical updates or changes in the pond's ecosystem, ensuring proactive monitoring. Finally, the user can securely sign out, completing their interaction. Overall, the diagram emphasizes user-friendly and real-time monitoring capabilities, designed to facilitate ecosystem management through IoT technology.

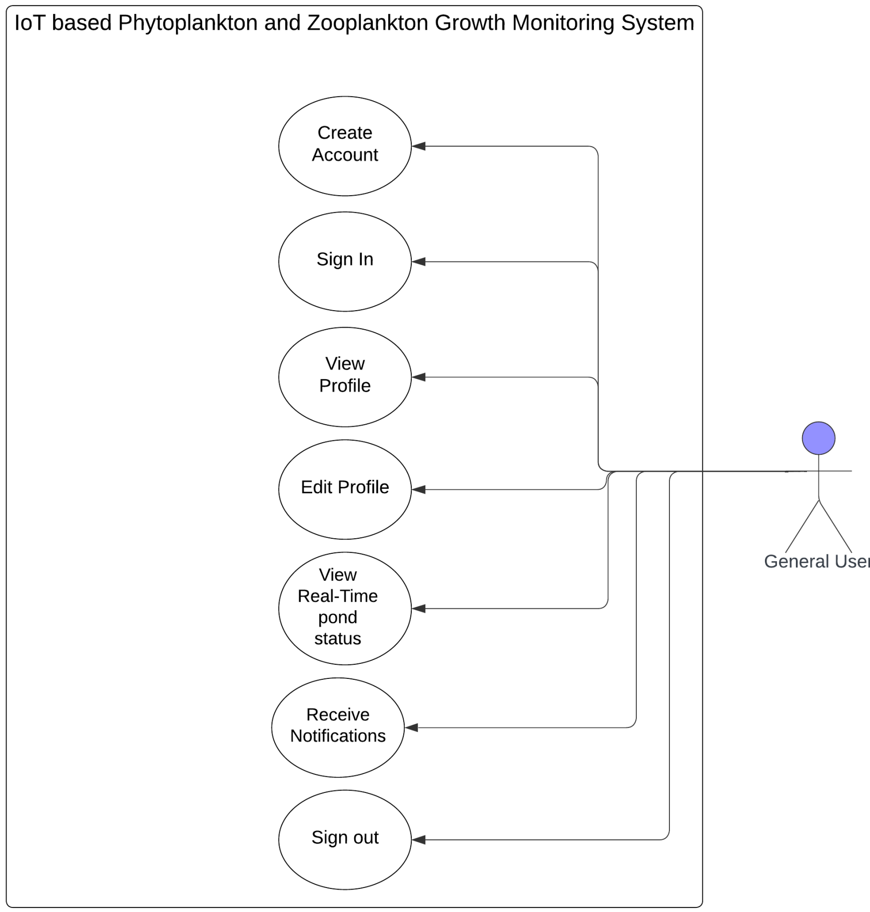


Figure 3.8: General user use case diagram

3.4.4.2 Admin Use Case Diagram

The diagram shown in the figure 3.9 represents the Admin Use Cases for an IoT-based Phytoplankton and Zooplankton Growth Monitoring System, depicts how the admin interacts with the system. The View Real-Time Status feature enables continuous monitoring of pond conditions and ecosystem health for admin. Monitor data collection use case allows admins to supervise the process of data collection from IoT devices deployed in the system. Admins can ensure that data is being recorded accurately and transmitted properly to the monitoring platform. They can also troubleshoot any issues related to data inconsistencies or transmission failures. Admins can adjust sensor settings through Configure Sensor Parameters, fine-tuning data collection and calibration. The Manage Application Users function ensures that admins can perform user management, providing role-based access control for different user types. Finally, View Analytics and Reports allows admins to generate detailed reports and analytical insights, enabling data-driven decisions to maintain a balanced ecosystem. Together, these functions equip admins with comprehensive tools for system oversight and operational efficiency.

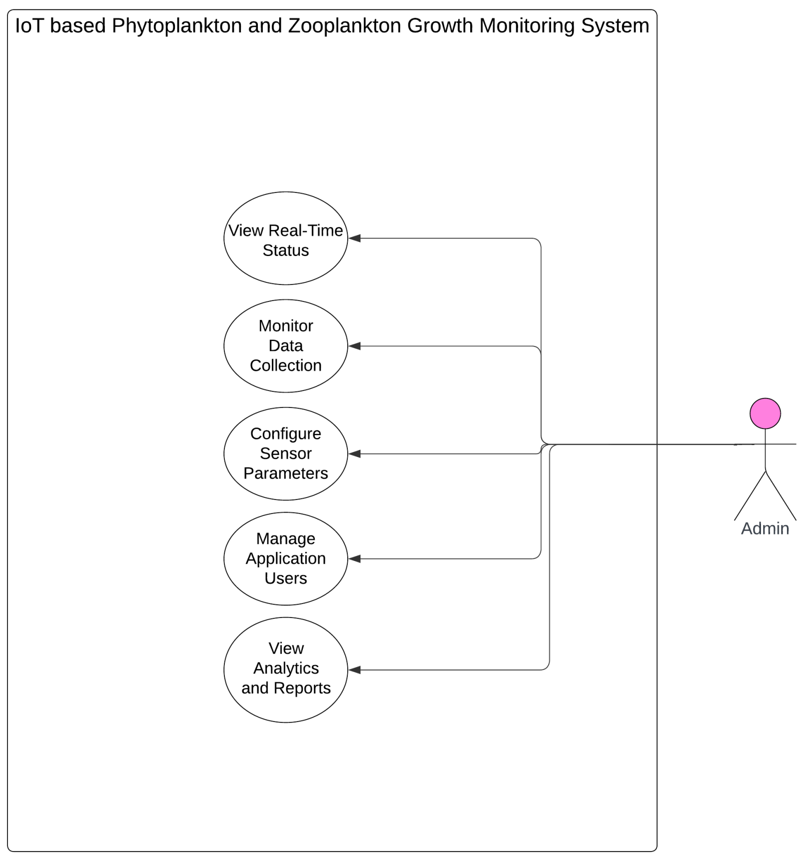


Figure 3.9: Admin Use Case Diagram

3.4.5 Activity Diagram

3.4.5.1 Admin Activity Diagram

In the admin activity diagram, the process begins with the login phase, where the admin enters credentials to gain access to the system. Once authenticated, the admin is presented with a range of options to manage the system effectively. One key responsibility is profile maintenance, where the admin oversees user management. This involves viewing the list of users, adding new users as needed, and ensuring that user details are kept up-to-date. Additionally, the admin has access to the device list, where they manage all devices linked to the system. This includes adding new devices, such as sensors, and performing regular maintenance to ensure that each device is functioning properly. These tasks are crucial to maintaining the system's integrity and ensuring smooth operation. After completing these administrative tasks, the admin can log out, thereby ending their session securely to prevent unauthorized access.

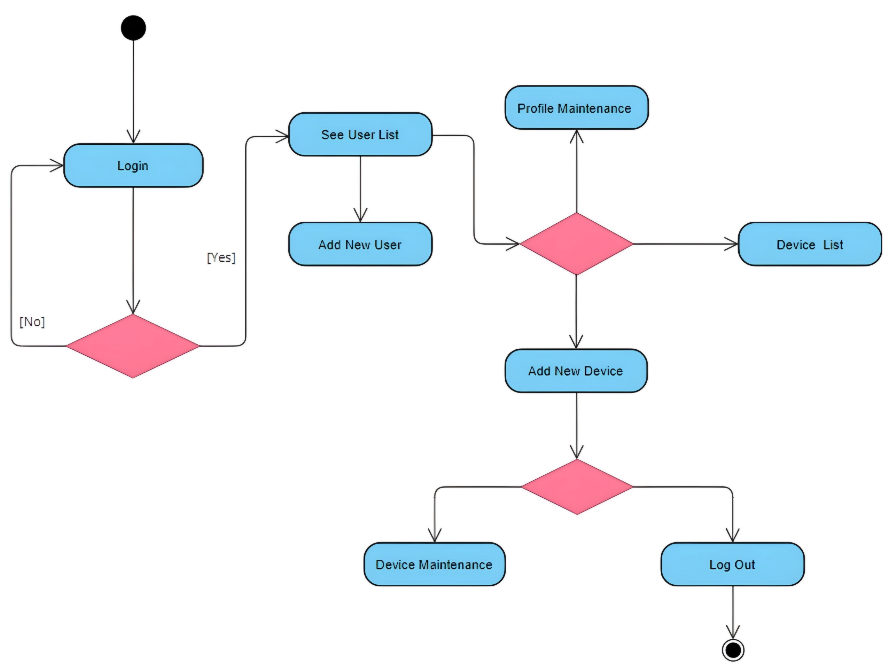


Figure 3.10: Admin Activity Diagram

3.4.5.2 User Activity Diagram

The user activity diagram, on the other hand, outlines a more focused set of tasks related to monitoring and interaction with pond data. After logging in, the user selects the pond they wish to monitor, allowing them to focus on specific pond conditions. The system displays real-time data for the selected pond, providing insights into water parameters and overall pond health. Users can also view notifications that alert them to any significant changes or conditions that may require attention. Additionally, the system offers access to historical data, enabling users to review past conditions and track trends over time. Alongside monitoring capabilities, the system allows users to view and edit their profiles, ensuring that their personal information is accurate. Once users have completed their tasks, they can log out, secure the system, and end their session.

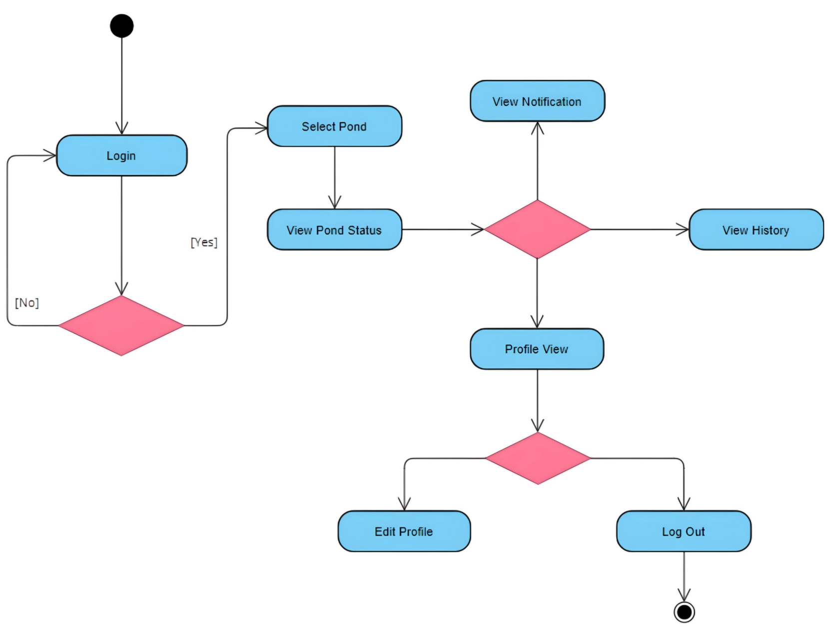


Figure 3.11: User Activity Diagram

In both diagrams, the flow is designed to ensure secure access, efficient system management, and real-time monitoring, with tailored roles and responsibilities for admins and users alike. The streamlined processes enhance the usability and functionality of the system, supporting effective pond management and device oversight.

3.4.6 Level-1 Data Flow Diagram

The provided Figure 3.12 illustrates a Data Flow Diagram (DFD) that outlines the interaction between different modules and users within the system, likely connected to aquaculture or environmental monitoring. It highlights the flow of information between users, services, and administrative functions to ensure smooth operation and effective management of the system.

The system begins with the Login (1.0) module, where both Users and Admin must authenticate to access their respective features. Once logged in, users can interact with various Services (2.0), such as monitoring pond water quality, viewing plankton levels, and receiving notifications. These services provide real-time insights into environmental parameters, helping users stay informed and take appropriate actions when needed. Users also can securely log out via the Logout (3.0) function, ensuring session management and preventing unauthorized access.

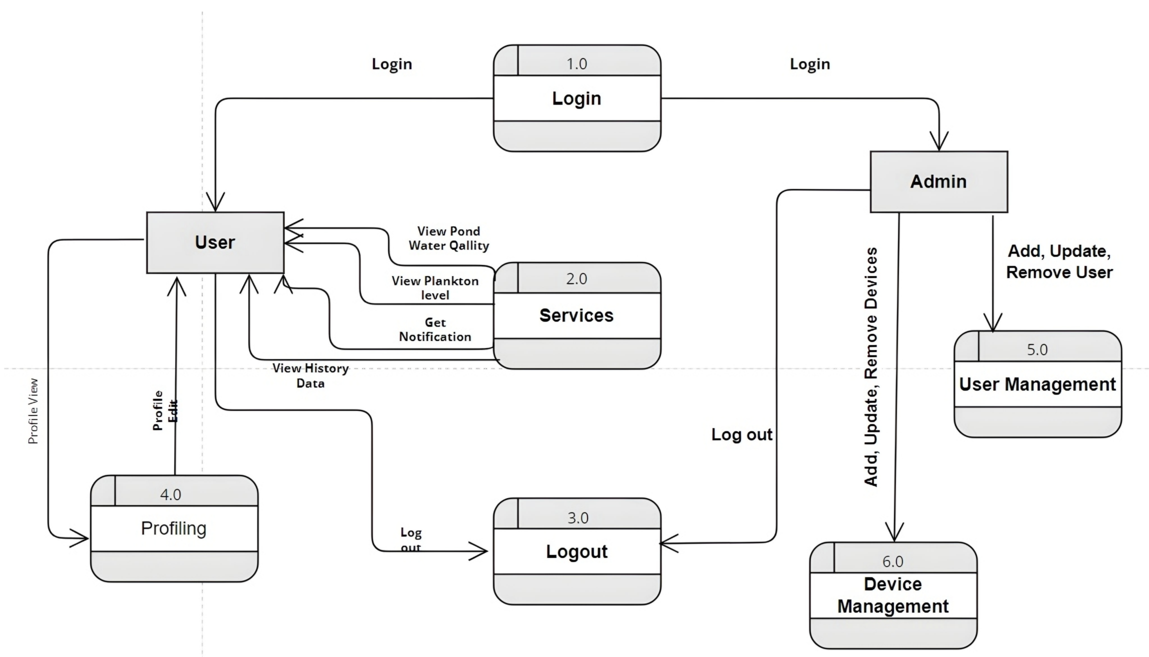


Figure 3.12: Level-1 Data Flow Diagram

Users can manage their personal information through the Profiling (4.0) module, which offers options to view and update their profiles. This feature ensures that users can keep their information current and aligned with system preferences. The Admin role, on the other hand, includes additional management responsibilities, such as User Management (5.0) and Device Management (6.0). Admins can add, update, or remove users, maintaining a well-organized user database. Similarly, they can manage the devices connected to the system by adding, updating, or removing IoT sensors and other hardware, ensuring the system functions smoothly.

The DFD shows how data flows between modules and distinguishes between the roles of Users and Admins. Users primarily engage with monitoring services and manage their data, while Admins are responsible for maintaining system operations by overseeing users and devices. The modular structure ensures secure access control, role-based permissions, and efficient management of system resources, contributing to the long-term sustainability and effectiveness of the platform.

3.5 Implementation

3.5.1 Hardware Implementation

This section describes the key components required for designing the Embedded IoT system for monitoring vital water quality parameters(pH, temperature, turbidity, TDS) to build a smart aquaculture for phytoplankton and zooplankton growth monitoring.

- **Arduino UNO:** The Arduino UNO is a microcontroller that acts as the system's brain. It controls the sensors, reads data from them, and sends/receives signals to and from other components, such as Wi-Fi modules or output devices.

- **ESP8266 ESP-01 Wi-Fi Module:** This module provides wireless connectivity, allowing the system to send sensor data over Wi-Fi to a remote server or cloud service. It is essential for enabling IoT functionality, where data from the sensors can be monitored remotely.

- **pH Sensor:** This sensor is used to measure the pH level of the water in the system. The pH level is crucial for monitoring the growth conditions for phytoplankton and zooplankton in aquaculture. It is connected to the Arduino to continuously monitor the water's acidity or alkalinity.

- **Temperature Sensor:** This sensor is responsible for measuring the water temperature. Temperature plays a significant role in aquaculture as it affects the growth of organisms.

- **Turbidity Sensor:** The turbidity sensor measures the clarity of the water, which is essential for assessing the water quality. Higher turbidity indicates more particles or microorganisms in the water, potentially affecting phytoplankton and zooplankton growth.

- **TDS (Total Dissolved Solids) Sensor:** The TDS sensor measures the concentration of dissolved solids in the water. It provides important information about the water's mineral content, which affects plankton and overall water quality. This sensor connects to the Arduino to monitor and transmit water quality metrics.

- **Battery Pack:** The battery pack powers the entire system. It ensures that the Arduino and the connected sensors can operate without a direct connection to a power outlet, making the system more portable and adaptable for remote environments.

- **Breadboard:** The breadboard is used to connect all components without soldering, allowing for easier prototyping and testing of the circuit. The power and ground lines are distributed through the breadboard, and it serves as the main hub for connecting the sensors to the Arduino.

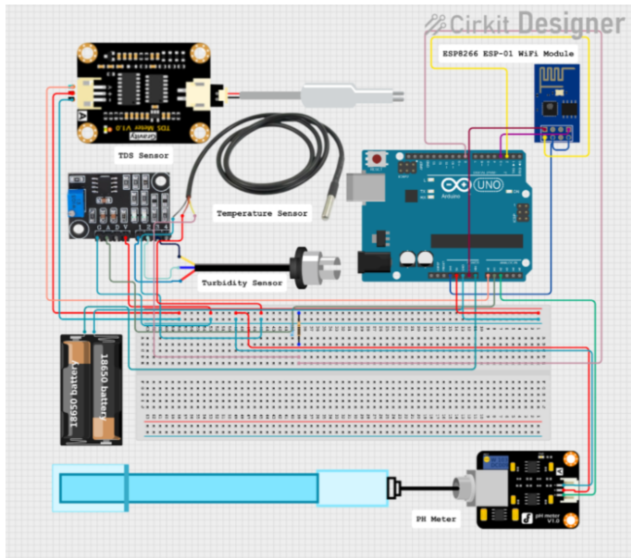


Figure 3.13: Circuit Diagram

Overall Circuit Connection Description: The pH, temperature, turbidity, and TDS sensors are all connected to the Arduino, and each sensor continually collects data from the water. Following that, Arduino processes the data. The ESP8266 Wi-Fi module enables the Arduino to send collected sensor data to a remote server or cloud platform, allowing for real-time monitoring of water quality indicators. This is critical for ensuring optimal plankton growth conditions in the smart aquaculture system. The battery pack powers the entire circuit, allowing for mobility and flexibility during field deployment in remote or aquatic situations.

3.5.2 Sensor Tuning

Sensor Tuning to the process of adjusting and fine-tuning sensors for optimal performance in a system, ensuring accurate and reliable data collection. This includes calibration to establish correct measurement baselines, filtering to minimize noise, and adjusting sensor sensitivity to the environment. Once tuned, sensors will be integrated into the system where they communicate with microcontrollers or IoT platforms, allowing collected data to be processed and used effectively in real-time applications like monitoring, control, or automation tasks.

1. **pH Sensor Tuning:** To calibrate the SEN-00239 pH Sensor, the system was warmed up for at least 15 minutes to stabilize its electronics. To eliminate

inaccuracies, we utilized two buffer solutions, typically pH 7.00 and one near the desired measurement range, while keeping both the sensor and the solutions at the same temperature. After cleaning with distilled water, we immersed the sensor in the pH 7.00 buffer, then waited for the reading to stabilize, and then set the meter to 7.00. We repeated the technique for the second buffer. We did calibration regularly, especially after extensive usage, contamination, or storage, to ensure accuracy.

2. **Temperature Sensor Tuning:** For the calibration of the Waterproof DS18B20 Digital Temperature Sensor (Model No: SEN-00072), we initially placed the sensor in a reference solution with a known temperature (such as an ice bath at 0°C or boiling water at 100°C) to ensure accuracy. After the sensor was immersed, the reference temperature was compared to the sensor's results. After that, the reported temperature was compared with the sensor's output, and any deviations were noted. To ensure accuracy and linearity, we performed the procedure at several temperature ranges. To ensure accurate temperature readings going forward, the acquired data was then utilized to modify the calibration variables in the system's programming.

3. **Turbidity Sensor Tuning:** To calibrate the SEN-00179 Analog Turbidity Sensor, we used pure water (0 NTU) as a baseline and standard turbidity solutions like 100 NTU and 400 NTU. The sensor was submerged in these solutions, and the corresponding analog voltage outputs were recorded. From the collected data, we generated a calibration curve to map voltage to turbidity in NTU. The microcontroller's code was then updated with the calibration formula to ensure accurate turbidity readings during operation, ensuring consistent performance in water quality monitoring.

4. **TDS Sensor Tuning:** For the calibration of the Analog TDS Sensor (Model No: SEN-00222), we began by using a reference solution with a known Total Dissolved Solids (TDS) value, typically a calibration solution with a defined ppm (such as 1413 ppm). The sensor's output voltage was measured after it was immersed in the solution. The difference between the measured output and the known TDS value was used to establish a calibration curve. To ensure accuracy over a range of values, this procedure was performed using other TDS reference solutions.

3.5.3 Software Implementation

- **Data Collection & Transmission:** Sensor data (pH, temperature, turbidity, TDS) collected by the Arduino UNO was transmitted wirelessly via the ESP8266 Wi-Fi module to a cloud server for storage and analysis.

- **Data Preprocessing:** Once stored in the cloud, the raw data underwent preprocessing to clean and normalize the values. Ensuring consistency and removing any missing data.

- **Machine Learning Model Development:** Various machine learning models, including XGBoost, CNN, and RNN, were developed and trained on the preprocessed data to predict phytoplankton and zooplankton abundance. These models used the sensor data to make accurate predictions.

3.5.4 System Integration and Data Collection Process

3.5.4.1 Integration Process

To integrate the pH, temperature, TDS, and turbidity sensors with the Arduino UNO, ESP8266 ESP-01 Wi-Fi Module, breadboard, and battery pack, we started by establishing the power supply. All sensors required a voltage of 5V, which we provided using the Arduino UNO's 5V pin. Each sensor's power input was connected to the 5V rail on the breadboard, and the ground (GND) pins of both the Arduino and breadboard were connected to ensure a common ground.

For signal connections, the pH sensor was connected to one of the Arduino's analog input pins to monitor its output. The temperature sensor (DS18B20) was interfaced using a digital pin, with a 4.7kΩ pull-up resistor placed between the data pin and the 5V supply. The TDS sensor's analog output was connected to another analog input pin on the Arduino, while the turbidity sensor's output was linked to a free analog pin.

The ESP8266 Wi-Fi Module required a 3.3V supply, so its VCC and CH_PD (enable) pins were connected to the 3.3V output from the Arduino. To safely interface the 3.3V signals with the Arduino's 5V logic for the RX and TX lines, we used a logic level shifter. Alternatively, if a level shifter was unavailable, we implemented a resistor divider for the TX signal from the Arduino to the ESP8266.

The RX pin of the ESP8266 was connected to the TX pin of the Arduino, and the TX pin of the ESP8266 was linked to the RX pin of the Arduino.

The battery pack was then connected to the power input (VIN) of the Arduino to provide continuous 5V power to the system. This setup allowed the Arduino to regulate the power of the sensors and the Wi-Fi module. We wrote the Arduino code to read the sensor values continuously from the analog and digital pins. This data was then sent to the ESP8266 module via serial communication, enabling it to transmit information to a server or cloud platform for real-time monitoring. This integration facilitated the collection of water parameter data, processed by the Arduino, and made accessible remotely through the ESP8266 Wi-Fi module.

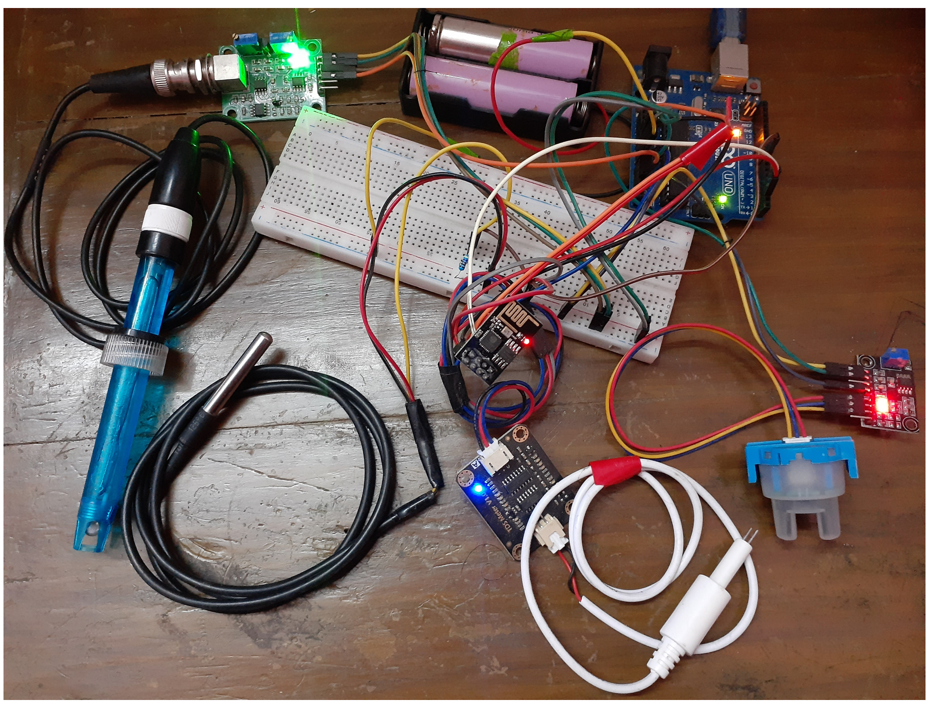


Figure 3.14: IoT Hardware System

3.5.4.2 Data Collection Process

Figure 3.15 shows the data collection in testing sites. We collected data from across 13 ponds situated in Magura, Noakhali, and the Botanical Garden in Dhaka, commencing on January 1st and concluding on July 10th. Over these seven months, approximately 4,500 samples were gathered, highlighting the extensive monitoring efforts undertaken in these diverse locations. The sensor data were recorded continuously, with measurements taken twice daily during peak hours, specifically between 9-10 AM and 5-6 PM. To streamline data logging and ensure real-time accessibility, the collected data were directly uploaded to a Google Sheet using an Arduino microcontroller integrated with a Wi-Fi module. This setup facilitated efficient data management and analysis, enabling it to monitor environmental conditions and track trends in water quality across the selected sites effectively.

Figure 3.15: Data Collection in testing sites

Chapter 4

Result analysis

4.1 Plankton Sampling and Abundance Measurement

Water samples for plankton analysis were collected biweekly from each pond and preserved with a formalin solution for later examination in the laboratory, as illustrated in Figure (4.1). In the lab, the samples were passed through a plankton net with a 25 μm mesh size for filtration.

Figure 4.1: Plankton Analysis in Lab

Plankton abundance was measured using a Sedgewick-Rafter Counting Chamber under a binocular microscope. The quantification process followed the method outlined previously. The plankton abundance was calculated using the formula shown in Equation (1).

$$N = \frac{A \times C}{V \times F \times L} \tag{1}$$

where,

- N represents the number of phytoplankton cells or units per liter of the original sample
- A denotes the total number of plankton counted
- C is the volume of the final concentrate of the sample in milliliters (mL)
- V indicates the volume of a field in milliliters (mL)
- F is the number of fields counted
- L is the volume of the original water sample in liters (L)

4.2 Summary of Water Quality Dataset Overview

The provided Table 4.1, shows a sample of the dataset collected during the study period. The dataset presented here contains water quality parameters collected from different ponds on various dates. These parameters include temperature, pH, TDS (Total Dissolved Solids), turbidity, and DO (Dissolved Oxygen). The dataset also includes two biological indicators: Phytoplankton and Zooplankton, which are crucial in determining the ecological health of the water bodies.

Significant Features of the Dataset:

- **Temperature (Temp):** Recorded in degrees Celsius, ranging from 29.8°C to 33.1°C, indicating the temperature variation between different ponds and times.
- **pH:** Reflecting the acidity or alkalinity of the water. Values vary from 7.4 to 10.2, showing different water conditions, from slightly alkaline to highly alkaline.

- **TDS (Total Dissolved Solids):** These values range significantly, from 83 ppm to 511 ppm, showing variability in the concentration of dissolved materials in the water.

- **Turbidity:** Measured in NTU (Nephelometric Turbidity Units), indicating how clear the water is. Values range from 8 to 58 NTU, where higher values signify more suspended particles.

- **DO (Dissolved Oxygen):** A critical parameter for aquatic life, values here vary between 4.3 mg/L and 5.2 mg/L, which can influence the biodiversity of the pond.

- **Phytoplankton and Zooplankton:** These two biological indicators show large fluctuations across different ponds and times, ranging from a few hundred to millions, reflecting varying levels of productivity and potential eutrophication in these water bodies.

The dataset was collected from various pond locations, including BG 1, BG 2, BG 3, and several ponds in the Magura region. The data spans different dates and times, specifically on January 3rd and May 10th, 2024, reflecting seasonal variations in water quality parameters. These temporal differences highlight the natural fluctuations in environmental conditions across the ponds, providing insight into how factors such as temperature, pH, and biological indicators change over time.

This dataset provides a comprehensive overview of the environmental and biological conditions of the ponds, useful for further analysis of water quality, pollution levels, and ecosystem health. The variability across locations and times reflects both natural fluctuations and possibly anthropogenic influences, making it suitable for ecological monitoring and environmental studies.

Table 4.1: Summary of Data Collection

Date	Time	Temp	pH	TDS	Turbidity	DO	Phytoplankton	Zooplankton	Pond
3/01/24	9:00 AM	29.8	7.6	83	17	4.3	1245	932	BG 1
3/01/24	10:00 AM	30.1	7.4	88	12	4.5	1251	495	BG 2
3/01/24	10:00 AM	30.8	7.4	100	8	5.1	1668	904	BG 3
5/10/24	10:00 AM	32.4	10.2	223	58	5.2	4123535	94340	Magura 1
5/10/24	10:00 AM	32.5	10.1	203	48	5.2	60601	16246	Magura 2
5/10/24	10:00 AM	32.8	9.0	291	55	4.7	83333	50000	Magura 3
5/10/24	10:00 AM	33.1	9.9	511	49	5.1	1883333	11904	Magura 4
5/10/24	10:00 AM	32.1	9.1	215	57	4.8	1133333	66666	Magura 5

4.3 Data Preprocessing

In this section, steps were taken to ensure the dataset was clean, with no missing or duplicate values, making it suitable for analysis. A data balancing technique was implemented alongside feature modification to minimize bias toward the target variable. The Phytoplankton and Zooplankton data were categorized into 10 distinct levels based on plankton counts. The procedure for creating these levels is outlined in Table 4.2.

Table 4.2: Plankton value leveling

Level	**Phytoplankton**	**Zooplankton**
Level 1	$\leq$ 10000	$\leq$ 1000
Level 2	$\leq$ 50000	$\leq$ 10000
Level 3	$\leq$ 100000	$\leq$ 20000
Level 4	$\leq$ 500000	$\leq$ 30000
Level 5	$\leq$ 1000000	$\leq$ 40000
Level 6	$\leq$ 2000000	$\leq$ 50000
Level 7	$\leq$ 3000000	$\leq$ 70000
Level 8	$\leq$ 4000000	$\leq$ 100000
Level 9	$\leq$ 5000000	$\leq$ 200000
Level 10	$>$5000000	$>$200000

In this study, the optimal range for phytoplankton was identified as 10,000 to 2,000,000 cells/L, and for zooplankton, it was 10,000 to 50,000 cells/L. Levels 3 to 6 were considered optimal for phytoplankton, while levels 7 and 8 indicated a warning, and levels 9 and 10 were deemed harmful. Levels below 3 were classified as neutral for phytoplankton. Similarly, for zooplankton, levels 2 to 6 were optimal, levels 7 and 8 served as a warning, levels 9 and 10 were harmful, and level 1 was regarded as neutral Aravinth et al. (2023) and Duré et al. (2021).

4.4 Environmental and Biological Correlation Analysis

Using the first 103 data samples, the correlations between environmental parameters and biological indicators, as shown in Figure 4.2, will provide a more robust and realistic understanding of the relationships within the given data.

Phytoplankton: Correlations between phytoplankton and environmental factors are generally moderate. The strongest positive correlation is with turbidity (0.48) meaning that phytoplankton levels increase as turbidity increases, possibly due to an increase in suspended particles that can provide nutrients. pH (0.43) also showed a positive correlation, suggesting that more alkaline conditions may support greater phytoplankton growth. Other positive but weak correlations include TDS (0.35) and temperature (0.15) indicating that higher TDS (total dissolved solids) and temperature slightly favor phytoplankton growth. However, the negative correlation with DO (-0.38) suggests that higher dissolved oxygen levels are associated with lower phytoplankton concentrations, possibly due to oxygen consumption by competition with other organisms.

Zooplankton: The zooplankton correlations show a similar pattern but are generally stronger. Turbidity (0.61) has the highest positive correlation, indicating that turbid conditions are highly favorable for zooplankton, possibly due to increased food availability from suspended organic matter. pH (0.47) also correlates positively, suggesting that more alkaline water supports higher zooplankton populations. Additionally, TDS (0.41) and Temperature (0.26) are positively correlated, although temperature's influence is relatively weaker. Like phytoplankton, zooplankton shows a negative correlation with DO (-0.4), indicating that lower oxygen levels may support larger zooplankton populations.

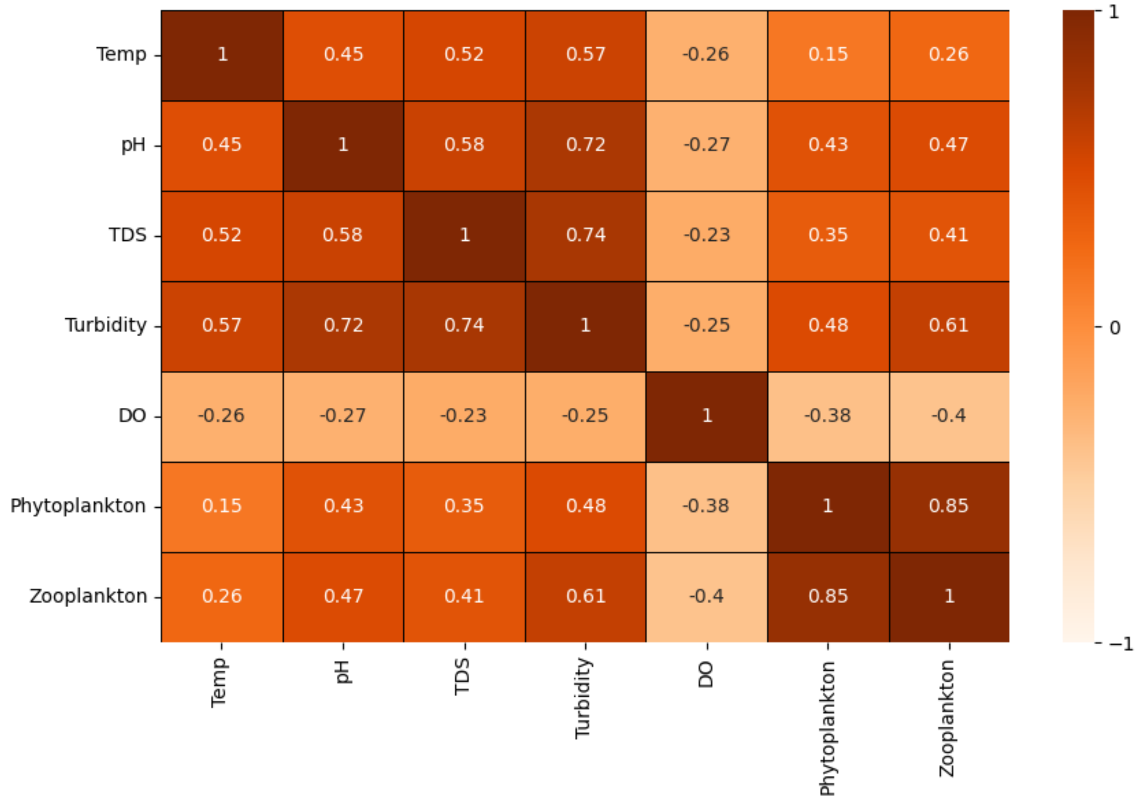


Figure 4.2: Correlation Matrix

4.5 Comparative Analysis of Plankton Classification Performance

This section provides a comparative analysis of different models' performance in classifying Phytoplankton and Zooplankton. High accuracy and robust f1-scores were observed for most classes. The analysis highlights the models' strengths in generalization and ability to manage multi-class classification effectively.

4.5.1 Support Vector Machines (SVM)

SVM is highly effective for classifying Phytoplankton and Zooplankton. SVM works by finding an optimal hyperplane that separates the classes and using the radial basis function (RBF) kernel, it can handle non-linear data well. Its strength lies in managing high-dimensional spaces and ensuring robust classification, even with smaller datasets. By tuning parameters like C and the kernel, SVM becomes a strong choice for achieving accurate classification in this multi-class problem.

Phytoplankton: Using parameters 'C': 10, 'kernel': 'rbf', the model achieved strong performance in classifying Phytoplankton across 10 levels, with an overall accuracy of 90.19%. The classification report shows excellent precision, recall, and f1-scores for the majority classes, particularly Level 1 and Level 4, both scoring near-perfect values. However, for some minority classes, such as Level 8, the model struggles with lower precision (0.45) and recall (0.29), indicating difficulty in correctly identifying this class.

The macro-average f1-score of 0.82 suggests that, although the model performs well across most levels, the performance of minority classes could be improved. Overall, the high weighted average f1-score of 0.90 reflects the model's robust ability to classify the majority of the samples correctly, while maintaining good generalization across different levels.

Zooplankton: With parameters 'C': 10, 'kernel': 'rbf', the model performed impressively in classifying Zooplankton levels, achieving an overall accuracy of 93.16%. The classification report shows outstanding precision, recall, and f1-scores for most classes, with Level 1, Level 2, and Level 10 scoring perfect 1.00 in all metrics. Levels such as 4 and 5 exhibit slightly lower performance, with f1-scores of 0.93 and 0.83, respectively, indicating some misclassifications. Notably, Level 6 has the lowest precision at 0.58 but compensates with a high recall of 0.93, which suggests the model tends to overclassify this level but manages to capture most of the true positives.

The overall macro-average f1-score of 0.92 reflects the balanced performance across all levels, while the high weighted average of 0.93 confirms the model's strong ability to generalize and provide accurate predictions across the dataset.

4.5.2 RandomForestClassifier (RF)

RandomForestClassifier's ability to manage multiple classes, even in an imbalanced dataset, makes it a valuable choice for this project.

Phytoplankton: The RF with parameters {'max_depth': 8, 'max_features': 0.6, 'max_samples': 1.0, 'n_estimators': 60} achieved an accuracy of 90.41% in classifying Phytoplankton. The model performs exceptionally well for major classes like Level 1 and Level 4, achieving f1-scores of 1.00 and 0.98, respectively. For minority classes such as Level 8, the model struggles, with a low f1-score of 0.27 due to poor recall (0.16). Overall, the macro-average f1-score of 0.80 and the weighted average of 0.89 indicate strong performance across the dataset with slight difficulties in handling the less frequent classes.

Zooplankton: The RF for Zooplankton, using the parameters {'max_depth': 8, 'max_features': 0.6, 'max_samples': 0.5, 'n_estimators': 120}, achieved a high accuracy of 93.05%. Major classes, like Level 1, Level 2, and Level 10, performed perfectly with an f1-score of 1.00. Most other levels also performed well, such as Level 4 and Level 9, with f1-scores of 0.94 and 0.99, respectively. However, Level 6 had a lower f1-score of 0.69 due to lower precision (0.58). Despite this, the overall macro-average and weighted average f1-scores of 0.92 reflect strong performance across all levels, demonstrating the model's robustness.

4.5.3 k-nearest neighbors (KNN)

KNN is a strong classifier for multi-class problems like Phytoplankton and Zooplankton classification, especially when fine-tuned with optimal parameters.

Phytoplankton: Using KNN for Phytoplankton classification, the model achieved an accuracy of 88.97%. The model performs exceptionally well for Level 1, Level 4, and Level 5, with f1-scores of 1.00, 0.98, and 0.97, respectively. However, for Level 8, the performance is lower, with an f1-score of 0.41 due to a precision of 0.35 and recall of 0.49. The overall macro-average f1-score is 0.80, indicating that while the model performs well for the majority classes, improvements are needed for the minority classes. The weighted average f1-score of 0.90 reflects strong general performance.

The best parameters for kNN in classifying Phytoplankton are {'n_neighbors': 3, 'weights': 'distance'}, which helped achieve an accuracy of 88.97% and strong performance for most classes, particularly Level 1, Level 4, and Level 5.

Zooplankton: The kNN model for Zooplankton, using the best parameters {'n_neighbors': 5, 'weights': 'distance'}, achieved an accuracy of 92.28%. The model performed perfectly for Levels 1, 2, and 10, each with f1-scores of 1.00. Strong results were also seen for Levels 9 and 8, with f1-scores of 0.99 and 0.91, respectively. However, Level 6 had a lower f1-score of 0.68 due to reduced precision (0.59). The overall macro-average f1-score of 0.91 and weighted average of 0.92 demonstrate the model's solid classification ability across most levels.

4.5.4 Convolutional Neural Networks (CNN)

Convolutional Neural Networks (CNN) are powerful models commonly used for image and pattern recognition tasks. They work by applying convolutional layers that detect spatial hierarchies and features in data through filters, followed by pooling layers that reduce the dimensionality. As the data passes through successive layers, CNN learns more abstract features. For classification tasks, the learned features are fed into fully connected layers to generate predictions for each class, which in this case allows the model to identify patterns among different Phytoplankton and Zooplankton levels based on input data.

Figures (4.3, 4.4, 4.5, and 4.6) represent the training and validation loss of the Phytoplankton and Zooplankton classification model as a function of epochs during training.

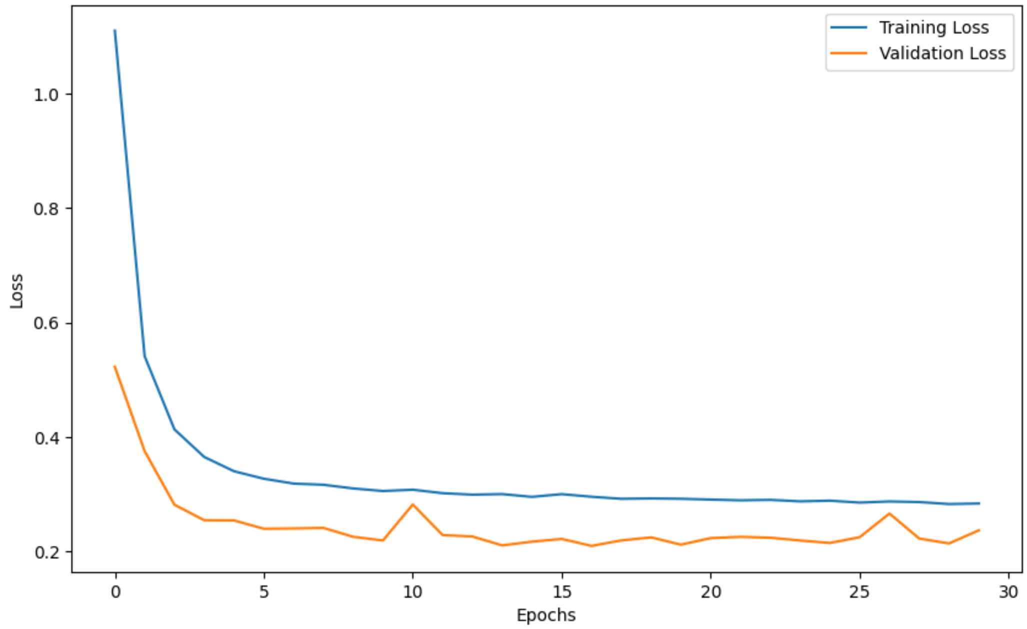


Figure 4.3: Phytoplankton Model Loss vs Epoch (CNN)

Figure 4.3 shows that both the training and validation loss decrease initially, indicating that the model is learning and improving. After around 10 epochs, both losses flatten out, suggesting that further training may not lead to significant improvements. Typically, overfitting occurs when the training loss decreases while the validation loss starts to increase. However, in this case, both training and validation loss stay consistent and low after stabilization, suggesting the model is not overfitting to the training data. The low values of both training and validation loss nearly (0.25) imply that the model is performing well in classifying the Phytoplankton data.

Training Loss (Blue Line): This line tracks the loss on the training dataset during each epoch. At the start (epoch 0), the training loss is high, but it decreases rapidly as the model learns, eventually stabilizing around 0.25 after about 10 epochs.

Validation Loss (Orange Line): This line tracks the loss on the validation dataset, which is used to evaluate how well the model generalizes to unseen data. Like the training loss, it decreases at the beginning and stabilizes around 0.25, remaining fairly consistent across epochs.

Figure 4.4 shows that the Phytoplankton classification model is performing well, with validation accuracy stabilizing around 90%. The lack of significant overfitting and the convergence of both training and validation accuracy implies that the model is well-tuned for this task.

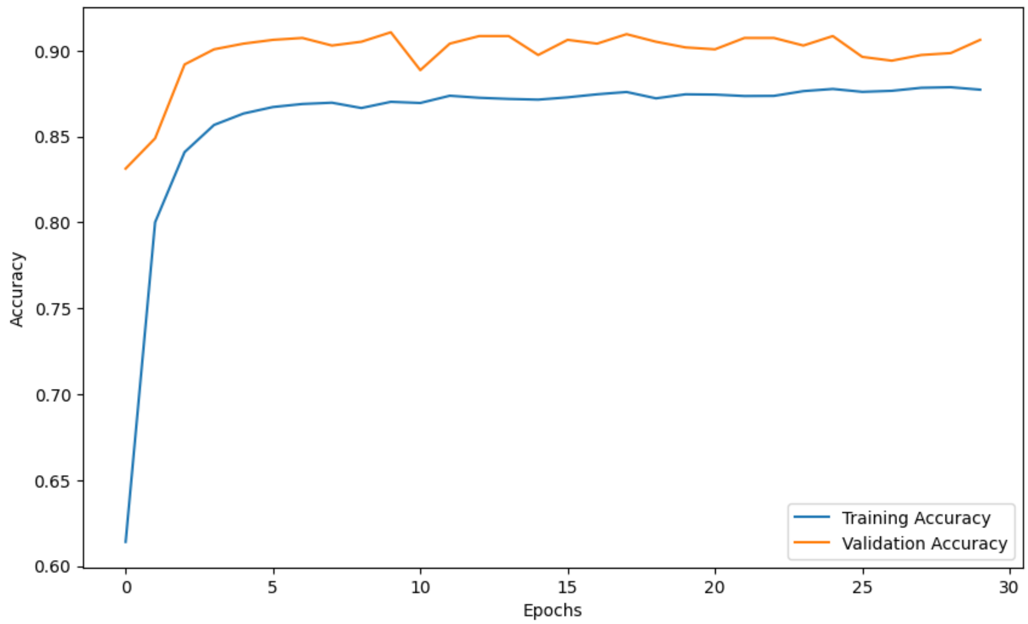


Figure 4.4: Phytoplankton Model Accuracy vs Epoch (CNN)

Training Accuracy (Blue Line): This line tracks the accuracy of the model on the training dataset. At the start (epoch 0), the training accuracy is low nearly 60%, but it increases rapidly and stabilizes around 85% after about 10 epochs.

Validation Accuracy (Orange Line): This line tracks the model's accuracy on the validation dataset, used to evaluate generalization on unseen data. Validation accuracy starts around 80% and rises quickly, stabilizing between 87% and 92% after the first few epochs.

Figure 4.5 suggests that the Zooplankton model is learning effectively without overfitting, as both the training and validation losses are low and consistent. The model achieves optimal performance after around 10 epochs, with further training showing minimal improvement. Both training and validation loss stabilize at a low value of nearly 0.2, indicating the model is performing well and has minimized the error in predictions for both the training and validation datasets.

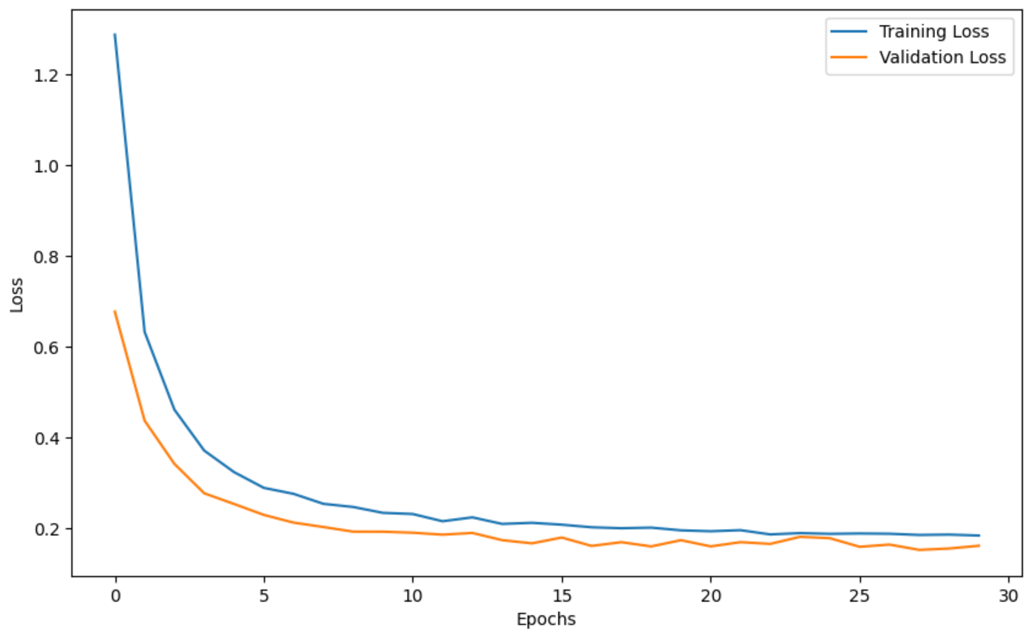


Figure 4.5: Zooplankton Model Loss vs Epoch(CNN)

Training Loss (Blue Line): This line tracks the loss on the training dataset. Initially, the training loss is high nearly 1.3, but it decreases sharply and stabilizes around 0.2 after approximately 10 epochs.

Validation Loss (Orange Line): This line tracks the loss on the validation dataset, which is used to evaluate how well the model generalizes. It follows a similar downward trend, starting around 0.6 and decreasing to approximately 0.2 after 10 epochs, with minimal fluctuation afterward.

Figure 4.6 represents the accuracy of the Zooplankton classification model using a Convolutional Neural Network (CNN) over 30 epochs. It shows both the training accuracy (in blue) and validation accuracy (in orange) over time.

Both training and validation accuracy increase rapidly during the first few epochs. By epoch 5, the validation accuracy is already close to 90%, and the training accuracy catches up shortly after. Around epoch 10, both training and validation accuracies begin to stabilize, indicating that the model is no longer learning significant new patterns.

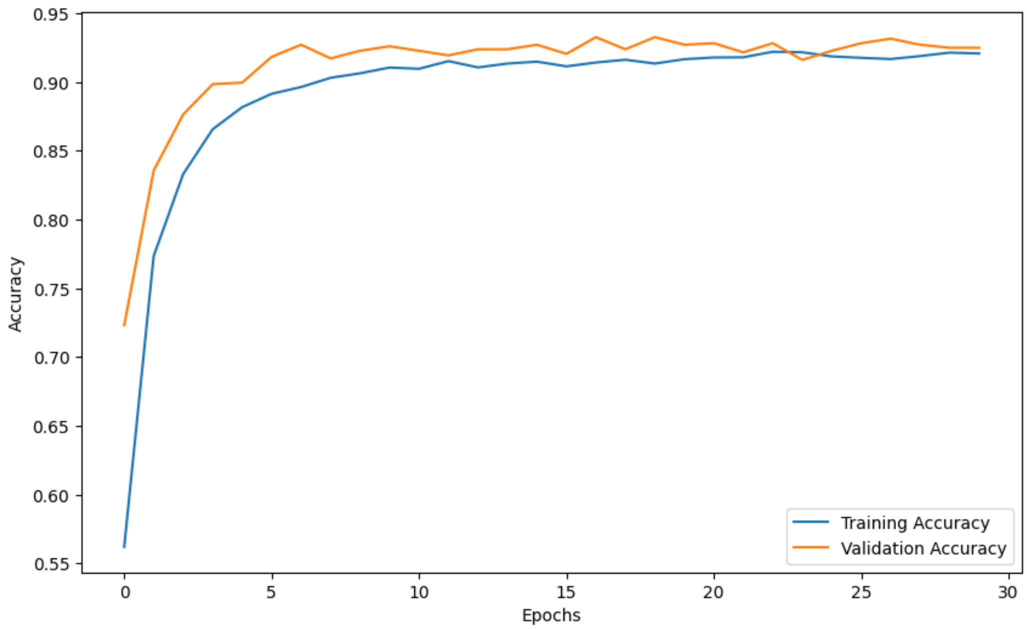


Figure 4.6: Zooplankton Model Accuracy vs Epoch (CNN)

After epoch 15, the two curves remain close, fluctuating slightly, but neither shows significant divergence, which is a good sign that the model is not overfitting. The model achieves around 92% accuracy for both training and validation at the end of the 30 epochs, which suggests that the model is well-optimized and generalizes well to unseen data.

Phytoplankton: The CNN model achieved a 90% accuracy for Phytoplankton classification, demonstrating strong overall performance across most levels. Levels 1, 4, and 5 performed exceptionally well, with f1-scores of 1.00, 0.98, and 0.97, respectively, indicating perfect or near-perfect precision and recall. Level 6 also exhibited high performance with an f1-score of 0.94. However, Level 8 showed lower performance with an f1-score of 0.39, primarily due to a recall of 0.31, suggesting that it had difficulty detecting true positives. Levels 7 and 10 achieved moderate f1-scores of 0.74 and 0.75, respectively, while Level 9 scored 0.82. Despite some variability in class performance, the macro-average f1-score of 0.82 and the weighted average f1-score of 0.90 indicate a well-performing model overall.

Zooplankton: The CNN model for Zooplankton classification achieved an overall accuracy of 93%, reflecting strong performance across most levels. Levels 1, 10, and 2 performed flawlessly, each attaining f1-scores of 1.00. Levels 8 and 9 also showed excellent results, with f1-scores of 0.92 and 0.98, respectively. Level 5 had moderate performance, achieving an f1-score of 0.83, while Level 6 exhibited some challenges with an f1-score of 0.66, due to lower precision. The overall macro-average f1-score of 0.91 and weighted average f1-score of 0.92 indicate a well-balanced model that handles both frequent and less frequent classes effectively.

4.5.5 Recurrent Neural Network (RNN)

Recurrent Neural Networks (RNNs) handle sequential data by maintaining information through loops in their architecture, making them ideal for tasks like time series and text classification. For multi-class problems, the RNN processes the sequence and passes the final hidden state through a dense layer with a softmax activation, which outputs class probabilities. The class with the highest probability is selected as the prediction. RNNs are trained using backpropagation through time (BPTT) to adjust weights and improve classification accuracy across multiple classes.

Figure (4.7,4.8, 4.9, 4.10) shows the loss over epochs for a Recurrent Neural Network (RNN) during training and validation. Here's a summary of what the plot indicates:

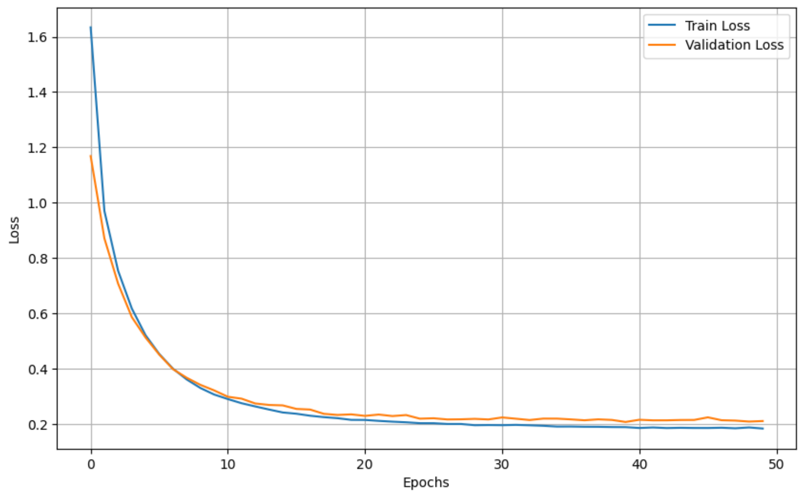


Figure 4.7: Phytoplankton Model Loss vs Epoch (RNN)

Figure 4.7 shows that both train Loss (blue curve) and Validation Loss (orange curve) decrease significantly as the number of epochs increases, indicating that the model is learning and the training process is effective. By around epoch 10, both losses stabilize at lower values, showing that the model has learned to a good extent.

The losses continue to decrease slightly but remain relatively stable after epoch 20. There is no visible indication of overfitting, as the validation loss closely follows the training loss, remaining consistent throughout the epochs.

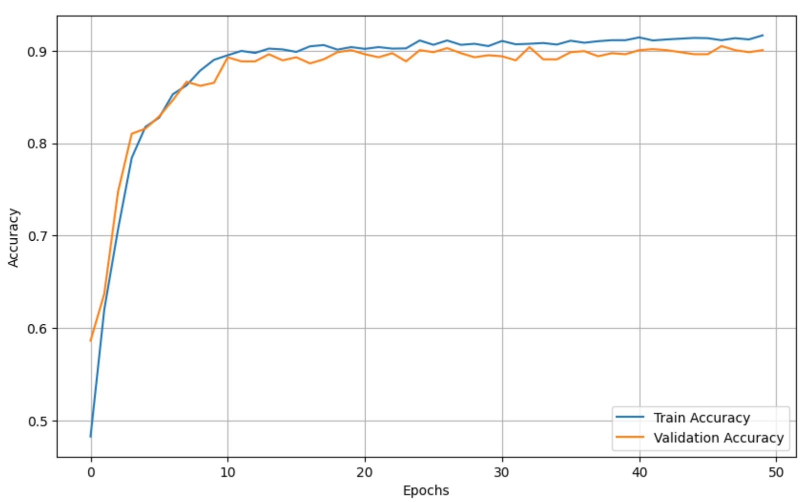


Figure 4.8: Phytoplankton Model Accuracy vs Epoch (RNN)

In Figure 4.8 the Accuracy vs. Epochs graph shows a rapid rise in both training and validation accuracy during the first 10 epochs, reaching around 0.9. After this, both accuracies stabilize with minimal fluctuations. The training accuracy remains slightly higher, but the small gap indicates minimal overfitting. Beyond 30 epochs, there is little improvement, suggesting the model has converged.

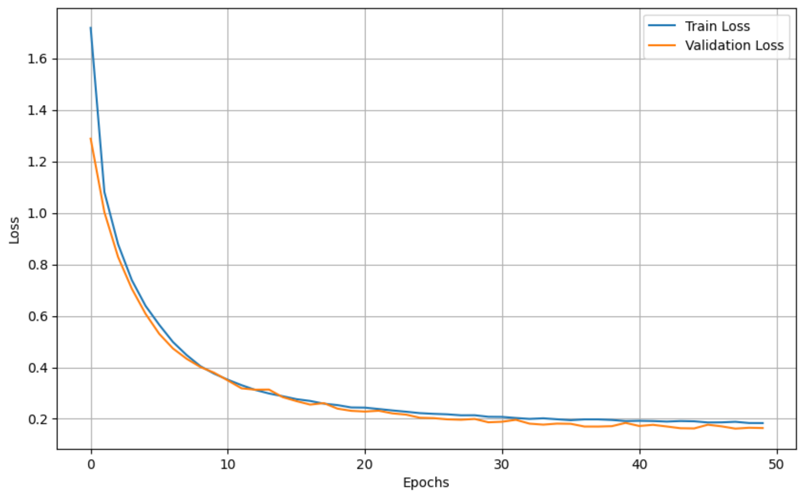


Figure 4.9: Zooplankton Model Loss vs Epoch (RNN)

In Figure 4.9 around 20 epochs the training and validation losses start stabilizing and converge, showing minimal fluctuation towards the later epochs (between 40 to 50 epochs). This suggests the model has learned well and is not overfitting, as the validation loss follows the training loss closely without increasing. The steady decrease and close alignment between training and validation losses indicate good generalization of the model. There is no significant overfitting since the validation loss remains low and mirrors the training loss, which is a positive outcome for the model.

The model training for Zooplankton classification is proceeding successfully, with both training and validation losses being minimized effectively over time.

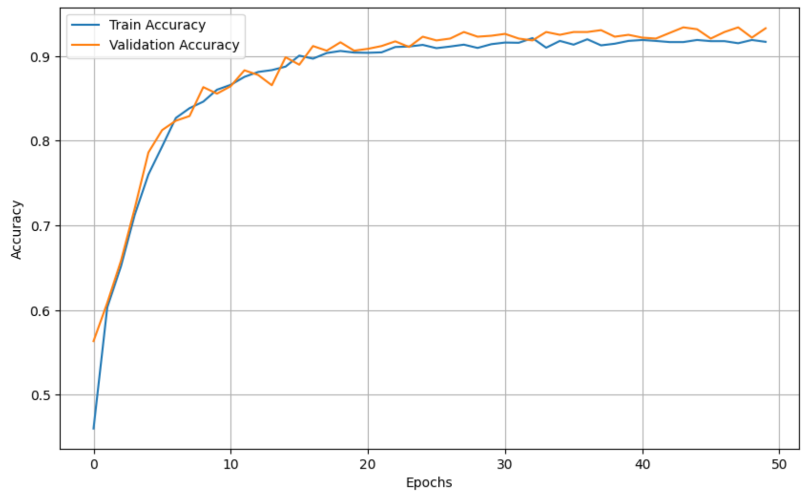


Figure 4.10: Zooplankton Model Accuracy vs Epoch (RNN)

In Figure 4.10 both training and validation accuracy start at relatively low values (below 60%) but improve rapidly during the initial epochs. After approximately 10 epochs, both accuracies exceed 85%, indicating the model is learning effectively. By around 20 epochs, accuracy stabilizes at a high value (around 90%), with both training and validation accuracies converging near 92–93%.

The accuracy continues to increase for both training and validation datasets but stabilizes after about 30 epochs. This close alignment between training and validation accuracy indicates that the model is generalizing well without overfitting.

The model achieves peak accuracy, with both training and validation accuracy remaining consistently above 90% from around epoch 30 onwards. There are slight fluctuations in validation accuracy, but it remains tightly aligned with the training accuracy, which suggests the model is robust and performing well on unseen data.

Phytoplankton: Overall accuracy of 93%. Levels 1, 4, 5, and 6 showed excellent performance with f1-scores above 0.97, while Level 10 and Level 9 also performed moderately well. However, the model struggled with Level 8, which had a low f1-score of 0.45, mainly due to poor recall. Overall, the model demonstrated strong predictive capabilities, although improving performance for less represented levels (like Level 8) could enhance its accuracy further.

Zooplankton: Zooplankton classification performed very well, achieving an overall accuracy of 93%. Levels 1, 2, and 10 demonstrated perfect performance with f1-scores of 1.00, while Level 9 also achieved near-perfect results with an f1-score of 0.99. Levels 3, 4, 5, and 7 showed solid results, with f1-scores ranging from 0.82 to 0.92. However, Level 6 had a lower f1-score of 0.74 due to a recall of 0.93 but weaker precision. Overall The model effectively classified most levels, with minor room for improvement in balancing precision and recall for certain levels like Level 6.

4.5.6 Gradient Boosting

Gradient Boosting tackles multi-class classification by extending its binary classification framework through strategies like ”One-vs-All”. In this method, it creates separate binary classifiers for each class.

The algorithm employs a multi-class loss function, such as softmax or logistic loss, to compute class probabilities. During training, it iteratively builds decision trees that minimize the residual errors of previous trees, enhancing accuracy with each iteration.

For final predictions, the output from all trees is combined to calculate probabilities for each class, and the class with the highest probability is selected as the final prediction. This iterative and corrective approach makes Gradient Boosting highly effective for multi-class classification tasks.

Phytoplankton: The performance analysis of the Phytoplankton classification using the Gradient Boosting Classifier reveals a commendable accuracy of 89.86%, indicating solid classification capabilities for a multi-class problem. The model excels in predicting Level 1 and Level 4, achieving perfect scores for Level 1 and high precision (0.99) and recall (0.97) for Level 4, leading to excellent F1-scores of 1.00 and 0.98, respectively. However, it struggles with Level 2 and Level 8, showcasing low precision (0.38) and recall (0.33), which highlights areas for improvement. The macro average F1-score of 0.78 suggests room for enhancement across classes, while the weighted average metrics indicate strong performance in the majority of classes. To boost overall performance, especially for the underperforming classes, strategies such as addressing class imbalance through oversampling or under-sampling, enhancing feature engineering, and further hyperparameter tuning should be considered.

Zooplankton: The Zooplankton classification using the Gradient Boosting Classifier achieved excellent results with a 92.50% accuracy. Level 1, Level 10, and Level 2 had perfect precision, recall, and F1-scores (1.00), while Level 9 also performed exceptionally (F1-score of 0.97). Moderate performance was seen in Level 4 (F1-score of 0.88) and Level 7 (0.87). Weaker results were observed for Level 5 (F1-score of 0.80) and Level 6 (0.74). Overall, the model performed strongly, but improving predictions for Levels 5 and 6 could enhance its accuracy further.

4.5.7 XGBoost

For both Phytoplankton and Zooplankton classification, XGBoost employed a structured approach to handle the complexity of multi-class classification. By using a moderate depth and controlling the learning process, the model was able to effectively capture relationships within the dataset. The classifier managed class imbalances, ensuring that even minority classes received adequate focus. This strategic setup allowed XGBoost to provide accurate and reliable predictions across diverse levels in ecological datasets, making it a powerful tool for classifying intricate biological data.

Phytoplankton: With 91% accuracy for Phytoplankton classification, the model demonstrated robust results. Levels 1 and 4 showed excellent performance with high f1-scores (1.00 and 0.98, respectively). However, Level 8 had a lower f1-score of 0.38 due to a low recall of 0.25. The overall macro average f1-score is 0.80, and the weighted average f1-score is 0.90, indicating strong overall performance with some variation in the lower-performing levels like Level 8.

Zooplankton: At 93.49% accuracy for Zooplankton classification, the model excelled in predicting several levels. Levels 1, 2, and 10 achieved perfect f1-scores of 1.00. Levels 4, 5, and 6 showed moderate performance, with f1-scores of 0.94, 0.80, and 0.74, respectively. The overall macro average f1-score is 0.92, and the weighted average f1-score is 0.94, indicating a solid performance across most levels.

4.6 Comparative Analysis of Top Model in Phytoplankton Classification

Table 4.3 presents a performance comparison of various machine learning algorithms for Phytoplankton classification, including SVM, Random Forest, KNN, CNN, RNN, Gradient Boosting, and XGBoost, evaluated on metrics such as accuracy, precision, recall, and f1-score. RNN has the highest accuracy 0.93. Among the models, KNN and CNN achieve an accuracy at 0.910, while Gradient Boosting has the lowest at 0.898. RNN stands out with the highest precision (0.85) and recall (0.93), indicating its strong capability to identify true positives effectively. In terms of f1-score, RNN also excels with a score of 0.90, highlighting its balanced performance. The macro and weighted averages show consistency across models, with RNN leading in overall metrics. Despite Gradient Boosting respectable accuracy, its lower precision and f1-scores suggest room for improvement. All models perform well, RNN emerges as the most effective choice for Phytoplankton classification with Random Forest and XGBoost also demonstrating strong performance suitable for scenarios requiring robustness and interpretability.

Table 4.3: Performance comparison among various ML algorithms for Phytoplankton

Model Name	**Accuracy**	**Precision**		**Recall**		**f1-score**	
		Macro Avg	**Weighted Avg**	**Macro Avg**	**Weighted Avg**	**Macro Avg**	**Weighted Avg**
SVM	0.901	0.80	0.90	0.86	0.90	0.82	0.90
Random Forest	0.904	0.84	0.92	0.83	0.90	0.80	0.89
KNN	0.910	0.80	0.90	0.81	0.91	0.80	0.90
CNN	0.910	0.81	0.91	0.86	0.91	0.82	0.90
RNN	0.930	0.85	0.93	0.83	0.93	0.83	0.92
Gradient Boosting	0.898	0.77	0.90	0.79	0.90	0.78	0.90
XGBoost	0.905	0.82	0.91	0.82	0.91	0.80	0.90

The provided Table 4.4 compares the performance of various machine learning algorithms for the classification of Zooplankton, featuring models such as SVM, Random Forest, KNN, CNN, RNN, Gradient Boosting, and XGBoost. In terms of accuracy, XGBoost achieves the highest score at 0.934, closely followed by KNN and SVM, both at 0.931. All models exhibit strong precision, with SVM, Random Forest, and XGBoost leading with scores of 0.92. The recall scores are also robust, highlighting the models' effectiveness in identifying true positive instances, with SVM, Random Forest, and XGBoost achieving a recall of 0.94. The f1-scores reflect a similar trend, with SVM, Random Forest and XGBoost achieving the highest f1-scores of 0.93. The macro and weighted averages indicate consistent performance across the models, with all maintaining similar metrics. All models perform exceptionally well in classifying Zooplankton, with XGBoost emerging as the top performer due to its superior accuracy, precision, and f1-score, while SVM and Random Forest also demonstrate reliable capabilities.

Table 4.4: Performance comparison among various ML algorithms for Zooplankton

Model Name	Accuracy	Precision		Recall		f1-score	
		Macro Avg	Weighted Avg	Macro Avg	Weighted Avg	Macro Avg	Weighted Avg
SVM	0.931	0.92	0.94	0.93	0.93	0.92	0.93
Random Forest	0.930	0.92	0.94	0.92	0.93	0.92	0.93
KNN	0.931	0.91	0.93	0.92	0.93	0.91	0.93
CNN	0.930	0.91	0.93	0.92	0.93	0.91	0.93
RNN	0.930	0.92	0.94	0.92	0.93	0.92	0.93
Gradient Boosting	0.925	0.91	0.93	0.91	0.93	0.91	0.93
XGBoost	0.934	0.92	0.94	0.93	0.93	0.92	0.94

4.7 ML Cross-Validation Performance for Phytoplankton and Zooplankton

Table 4.5 presents the cross-validation performance of various machine learning models for both Phytoplankton and Zooplankton classification, showing the accuracy scores across five folds and the mean cross-validation (CV) scores. For Phytoplankton, KNN, CNN, RNN, and XGBoost achieve strong mean CV scores of 0.90 or higher, with CNN slightly outperforming the others at 0.90, showing its consistency across folds. SVM demonstrates slightly lower but stable performance with a mean CV score of 0.87. Gradient Boosting performs reasonably well with a mean score of 0.89, while Random Forest shows solid results at 0.89, with slightly higher scores in some folds.

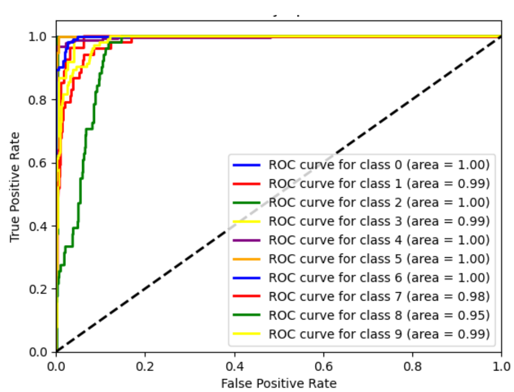
For Zooplankton, XGBoost achieves the highest mean CV score at 0.92, displaying consistent performance across folds. SVM, KNN, and CNN also perform well with mean CV scores of 0.91, showing strong predictive capability. Random Forest and Gradient Boosting demonstrate reliable but slightly lower performance with mean CV scores of 0.90 and 0.91, respectively. RNN also performs strongly, with a mean CV score of 0.91, indicating robustness across multiple folds.

Table 4.5: Cross-Validations Performance

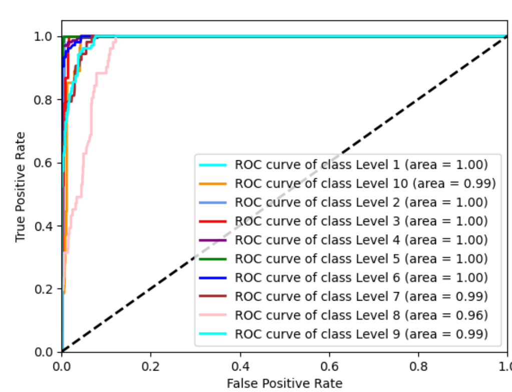
Models	**Plankton**	**CV Accuracy Scores**					**Mean CV Score**
		Fold 1	**Fold 2**	**Fold 3**	**Fold 4**	**Fold 5**	
SVM	Phyto	0.87	0.87	0.86	0.87	0.87	0.87
	Zoo	0.91	0.92	0.92	0.91	0.92	0.92
RF	Phyto	0.88	0.90	0.89	0.89	0.90	0.89
	Zoo	0.91	0.90	0.89	0.90	0.90	0.90
KNN	Phyto	0.88	0.91	0.91	0.91	0.90	0.90
	Zoo	0.91	0.91	0.92	0.91	0.90	0.91
CNN	Phyto	0.90	0.90	0.88	0.89	0.93	0.90
	Zoo	0.92	0.92	0.91	0.89	0.89	0.91
RNN	Phyto	0.90	0.89	0.90	0.90	0.92	0.90
	Zoo	0.90	0.90	0.91	0.91	0.92	0.91
Gradient Boosting	Phyto	0.89	0.90	0.89	0.88	0.91	0.89
	Zoo	0.91	0.92	0.90	0.91	0.91	0.91
XGBoost	Phyto	0.89	0.90	0.91	0.91	0.91	0.91
	Zoo	0.92	0.91	0.91	0.92	0.90	0.92

4.8 ROC Analysis of Model Performance in Phytoplankton Classification

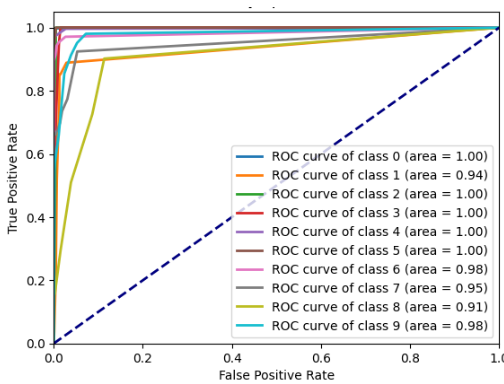
The ROC curve shown in Figure 4.12 is another metric used to compare the performance of ML models in terms of phytoplankton. The One vs Rest method has been used to draw the ROC curve for multi-class classification. From the curve, it can be stated that the RNN model performed best. The Area score of the RNN model is between 0.97 to 1.00. On the other hand, it can be stated that KNN is the worst-performed ML model in differentiating classes for this dataset as the Area score is between 0.91 to 1.00. Level 8 has demonstrated lower classification performance compared to other levels in the dataset, despite respectable ROC-AUC scores across various models. While models like SVM (0.95), CNN (0.96), RNN (0.97), and XGBoost (0.96) show strong results, the overall performance for Level 8 lags behind. Random Forest and KNN models both recorded an ROC-AUC score of 0.91. Gradient Boosting, with a score of 0.95, also performed well, but not enough to mitigate the lower general classification results. This indicates that while the models achieve good ROC-AUC scores, the classification challenges for Level 8 persist, likely due to the complexity in its features.



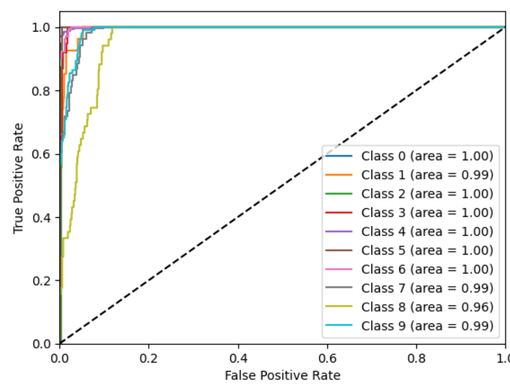
SVM



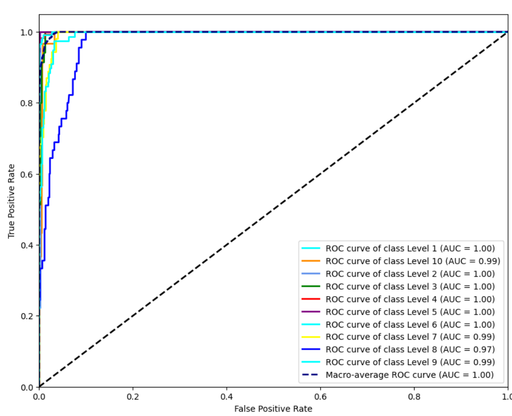
RF



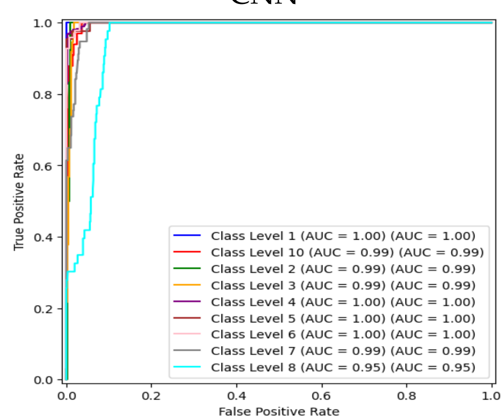
KNN



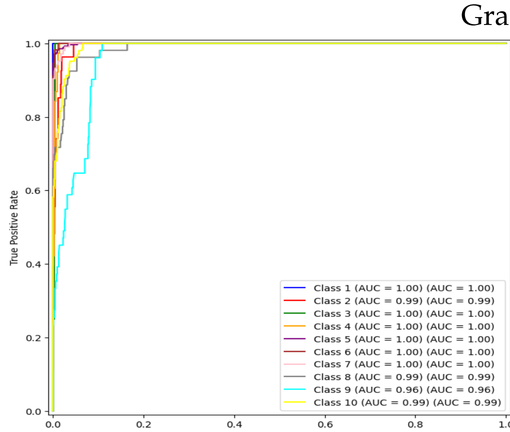
CNN



RNN



Gradient Boosting

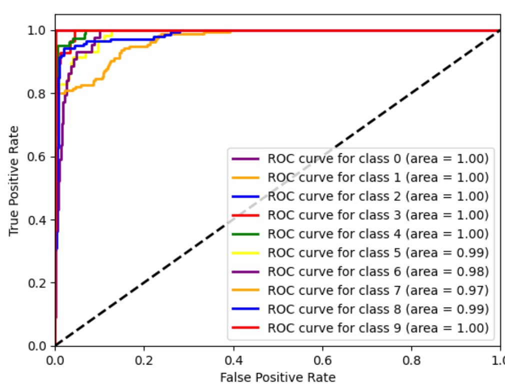


XGBoost

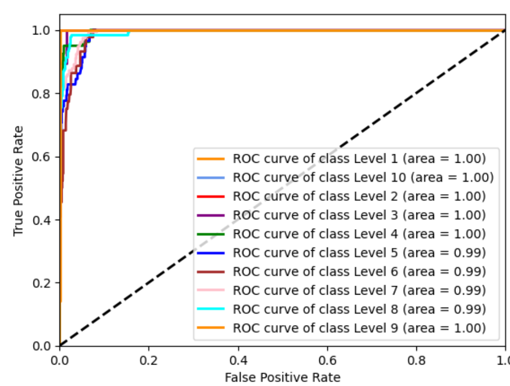
Figure 4.12: ROC Curve for Phytoplankton

4.9 ROC Analysis of Model Performance in Zooplankton Classification

The ROC curve shown in Figure 4.14 is an analysis of various machine learning models for Zooplankton classification revealing impressive performance across the board. The Support Vector Machine (SVM) achieved outstanding results, scoring 1.00 for most levels, with only minor reductions to 0.97 and 0.98 for levels 7 and 6, respectively. Similarly, Random Forest (RF) exhibited strong results, with several levels also scoring 1.00 and slight dips to 0.99 for levels 5, 6, and 7. K-Nearest Neighbors (KNN) performed well, reaching 1.00 for levels 1, 10, and 2, while scores for other levels ranged from 0.96 to 0.99. Convolutional Neural Network (CNN) and Recurrent Neural Network (RNN) maintained robust performance with many levels scoring 1.00, and minor decreases to 0.99 for some levels. Gradient Boosting mirrored RNN's results, achieving high scores of 1.00 across multiple levels with slight dips to 0.99. XGBoost also demonstrated similar effectiveness, reinforcing its strong classification capabilities.



SVM



RF

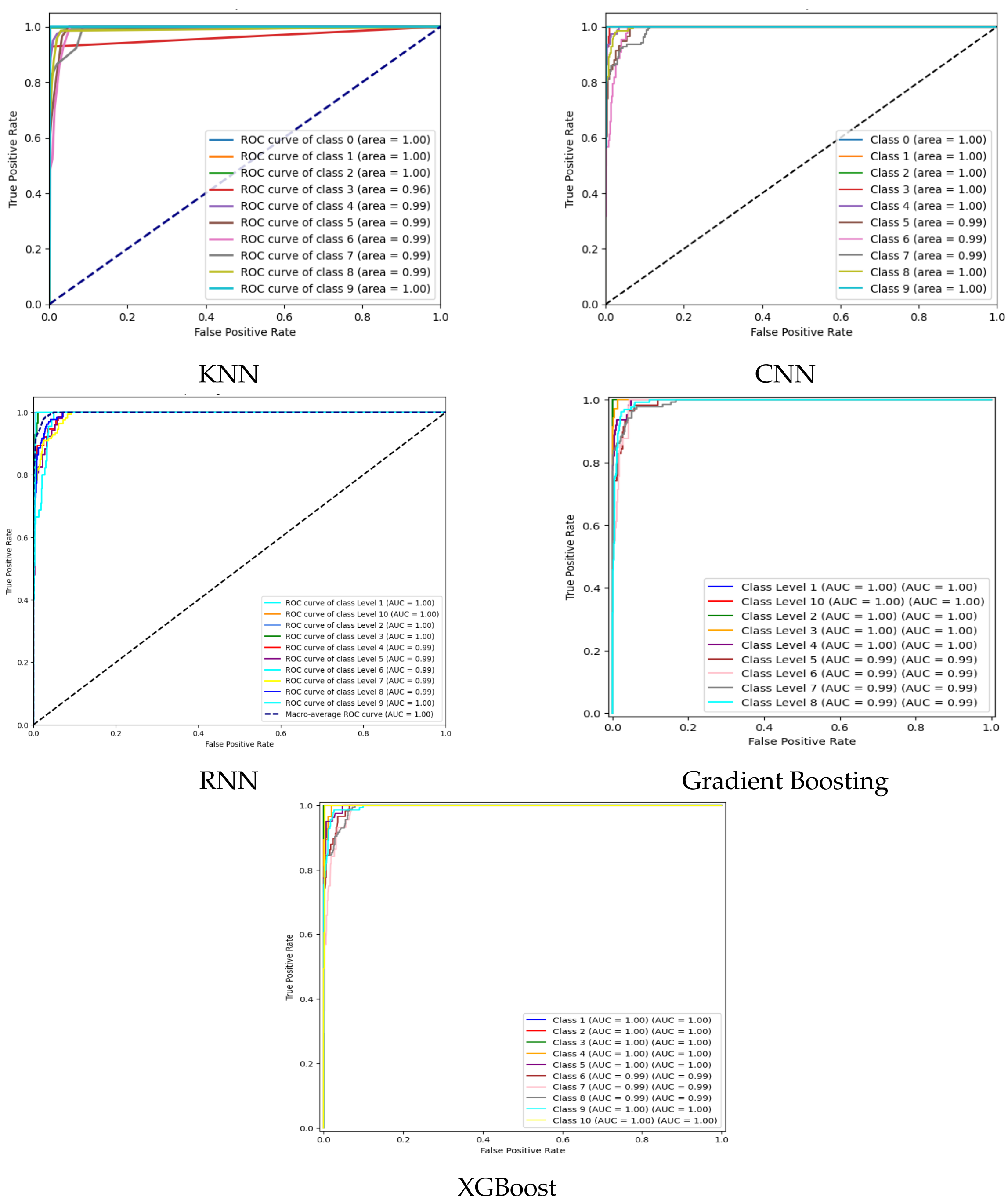
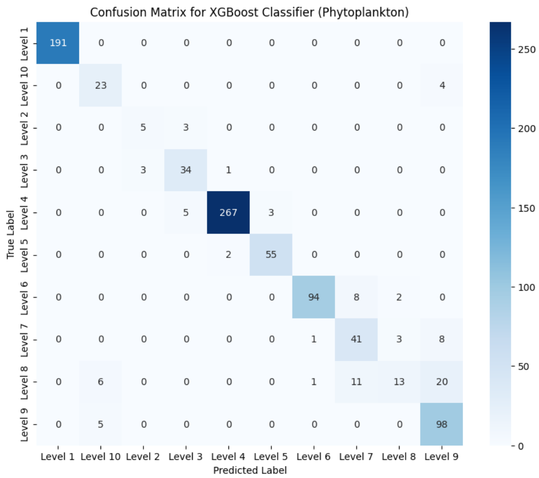
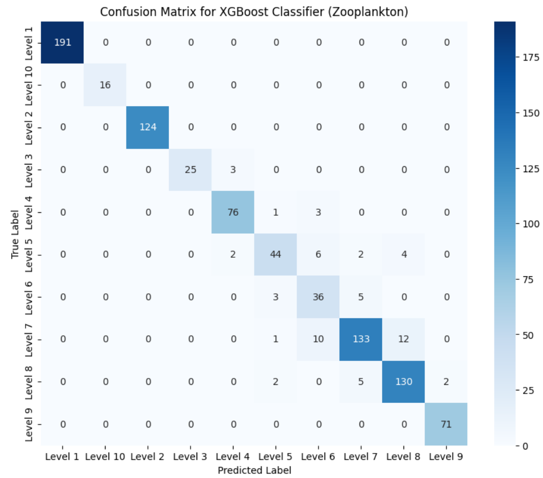


Figure 4.14: ROC Curve for Zooplankton

4.10 Balanced and Efficient Model for plankton Prediction



Confusion Matrix of XGBoost Model for Phyto



Confusion Matrix of XGBoost Model for Zoo

Figure 4.15: Confusion Matrix

The XGBoost model proves to be a highly effective choice for predicting both Phytoplankton and Zooplankton. For Phytoplankton, XGBoost achieved an accuracy of 91%, with impressive precision, recall, and F1-scores across multiple classes, demonstrating its ability to handle the multi-class classification task effectively. Similarly, for Zooplankton, XGBoost outperformed other models, achieving the highest accuracy of 93.4%. It also showed strong precision, recall, and F1-scores, indicating its capability to accurately classify the various levels of Zooplankton. XGBoost's advantages include high accuracy, robustness against overfitting, and computational efficiency, making it particularly suitable for predictions where computational cost is a concern. Overall, XGBoost stands out as a strong candidate for both Phytoplankton and Zooplankton prediction due to its high performance and ability to manage complex classification tasks effectively.

Chapter 5

Conclusions

5.1 Societal, health, safety, legal and cultural aspects

This IoT-embedded phytoplankton and zooplankton growth monitoring system for smart aquaculture can impact the project's societal, health, safety, legal, and cultural aspects. These issues arise due to the data collection, automation, and real-time predictive features enabled by machine learning algorithms. Below is an evaluation of each aspect and a proposed course of action to mitigate potential concerns.

1. Societal Impact

- **Economic Growth and Job Creation:** The system has the potential to boost aquaculture productivity, resulting in increased economic growth in the fisheries sector. However, automation may lower the demand for conventional labor, displacing workers.

  Mitigation: Improving workers' skills in technology management and data analysis can be a good solution to cope with this societal impact. Providing training programs to help farmers and workers so that they can migrate into roles that require system maintenance, data analysis, or farm management.

2. Health Impact

- **Food Safety:** The system's ability to monitor water quality parameters (pH, temperature, TDS, turbidity) in real-time ensures healthier environments for aquatic life, reducing the risk of disease in aquaculture products. This improves food safety for consumers.

Mitigation: Ensuring continuous monitoring and timely alerts for dangerous water conditions that may affect plankton growth or cause the spread of diseases can be a way to cope with this health impact. Creating protocols for immediate intervention when hazardous levels are detected might be beneficial too.

- **Consumer Health:** By ensuring that aquaculture water quality is maintained at optimum levels, the system reduces the risk of harmful microorganisms or pollutants entering the human food chain, improving overall public health.

  Mitigation: Working with local health authorities to maintain high standards of water quality and promoting safe farming practices that align with food safety guidelines will be an ideal solution for consumer health protection.

3. Safety Issues

- **System Reliability:** If the IoT sensors or machine learning algorithms fail, incorrect predictions could result in poor water quality management, leading to the destruction of aquatic life or contamination of the water system.

  Mitigation: Implement redundancy in sensor systems such as having backup sensors or alternative data sources. Regular maintenance schedules and performance testing should be enforced to ensure system accuracy and reliability. Additionally, manual overrides can be included in the system in case of system failure. These steps can help to mitigate the safety issues related to the system.

- **Risk of Data Breaches:** The collection and transmission of real-time data through wireless networks pose a cybersecurity risk. If data is intercepted or tampered with, it could disrupt aquaculture operations or lead to false water quality readings.

  Mitigation: Using secure communication protocols like encryption, and firewalls to protect the system from unauthorized access can be proven as a good solution. Conducting regular cybersecurity audits and implementing fail-safe mechanisms to detect and block cyber-attacks will mitigate the data breach issue.

4. Legal Issues

- **Data Privacy and Protection:** As the system collects sensitive environmental data, it must comply with legal standards on data privacy and protection. If the system shares data with third-party providers like cloud services, strict data handling protocols must be in place.

  Mitigation: Ensuring compliance with data protection regulations such as GDPR (General Data Protection Regulation) or CCPA (California Consumer Privacy Act), depending on the geographical region must be done. Obtaining consent from users for data collection, and clarifying how the data will be used should be done.

- **Environmental Regulations:** The system must comply with local and international environmental regulations regarding water quality standards and safe practices in aquaculture.

  Mitigation: The ways of dealing with this issue can be, working sincerely with local environmental agencies to ensure the system adheres to legal water quality parameters (pH, temperature, TDS, turbidity) for sustainable aquaculture. Performing regular environmental audits to ensure that operations remain within legal parameters.

- **Intellectual Property Rights:** The technology, algorithms, or data collected might be subject to intellectual property (IP) laws. If any proprietary technologies are used in the system, legal challenges may arise regarding their use or modification.

  Mitigation: Ensuring that all software, algorithms, and hardware are either open-source, properly licensed, or original developments to avoid infringing on any patents or IP rights will prevent this legal issue.

5. Cultural Issues

- **Cultural Acceptance of Technology:** In some regions, there may be resistance to adopting new technologies in traditional aquaculture practices. Cultural views on automation and machine learning could hinder the widespread implementation of the system.

Mitigation: Engaging with local communities through workshops and demonstrations to showcase the benefits of the system and tailoring the system to respect cultural values, ensuring that it complements traditional aquaculture practices rather than replacing them.

- **Local Ecosystem Knowledge:** Different regions may have unique aquatic ecosystems that require specialized monitoring techniques. The system's generalized approach may not account for local variations in water chemistry or species behavior.

  Mitigation: A possible solution could be customizing the system to incorporate local ecosystem knowledge, such as adjusting machine learning algorithms to account for region-specific variables that affect plankton growth.

5.2 Impact on Environment and Sustainability

The IoT-embedded phytoplankton and zooplankton growth monitoring system for smart aquaculture significantly impacts the environment and sustainability. By leveraging real-time data collection and machine learning algorithms to predict plankton growth, the system will promote a more sustainable approach to aquaculture while protecting and enhancing the aquatic ecosystem. Below are key points detailing the positive environmental and sustainability impacts:

- **Improved Water Quality Management:** Continuous monitoring of critical water parameters (pH, temperature, TDS, turbidity) helps maintain optimal aquaculture conditions, preventing issues like eutrophication, which can harm aquatic ecosystems.
- **Sustainable Plankton Management:** The system helps to optimize plankton growth, which is essential for aquatic food chains, while reducing resource waste like fertilizers, nutrients, etc. This ensures sustainable management of plankton populations, promoting ecosystem balance.
- **Energy Efficiency:** Automation of processes like feeding, aeration, and water control reduces the need for manual labor and minimizes energy consumption, thereby lowering the overall carbon footprint of aquaculture operations.
- **Mitigation of Environmental Risks:** Real-time monitoring helps detect harmful conditions such as harmful algal blooms (HABs) or water quality imbalances early and allows quick intervention to prevent environmental damage.
- **Biodiversity Conservation:** By maintaining healthy water conditions, the system supports plankton, which are crucial to marine ecosystems. This fosters biodiversity and reduces reliance on wild fish populations, promoting sustainable aquaculture practices.
- **Reduction in Chemical Use:** Real-time data reduces the need for excessive chemical treatments like antibiotics or algaecides, minimizing pollution and protecting surrounding aquatic ecosystems from harmful chemical runoff.

- **Sustainable Fish Farming:** The system supports more efficient management of resources such as water and feed, which minimizes the environmental footprint of fish farming operations and promotes long-term sustainability.

- **Data-Driven Sustainability:** The system's long-term data collection informs decisions on environmental conservation and aquaculture policies. This helps to guide sustainable practices and ensures compliance with environmental regulations.

5.3 Ethical and professional principles

Several ethical concepts and professional ethics were and will be carefully incorporated into the IoT-based phytoplankton and zooplankton growth monitoring system for smart aquaculture to assure safety, compliance, and environmental sustainability. This project is consistent with basic engineering obligations, prioritizing public safety, environmental protection, and adherence to legal standards.

1. **Ethical Responsibility to the Public and Environment**

   The primary ethical duty of an engineer is to ensure that the solution is safe for both the public and the environment. For this project, the following steps were and will be taken to meet these obligations:

   - **Public Safety:** The system will continuously monitor water parameters (pH, temperature, TDS, turbidity) and detect unsafe levels early, preventing unhealthy conditions in aquaculture and minimizing risks to human health from contaminated products.

   - **Environmental Protection:** The system is designed to prevent harmful practices, such as the overuse of chemicals or fertilizers, which can damage ecosystems. By using real-time data and machine learning predictions, farmers can maintain optimal water conditions, preventing events like algae blooms or oxygen depletion, which could devastate local ecosystems

2. **Design for Safety and Hazards and Failure Analysis**

   To ensure the system is robust and safe to use, Design for Safety methods and Hazards and Failure Analysis were and will be incorporated into the design process.

   - **Hazards Identification:** Potential hazards such as sensor failures, incorrect data processing, or communication breakdowns were identified. These hazards could lead to incorrect predictions, poor water quality, or plankton die-off.

   - **Failure Analysis:** A detailed failure analysis was conducted to assess potential risks. For example, sensor malfunction could lead to inaccurate data, which could disrupt plankton growth predictions or lead to harmful environmental conditions. To mitigate these risks, sensor redundancy was implemented and automated system checks would be incorporated.

   - **Safety Measures:** Redundant sensors and regular maintenance schedules were incorporated to ensure accurate data collection. In the case of system errors, manual overrides will be included to allow for immediate human intervention. Additionally, the system is intended for real-time alerts to prevent escalation of issues and minimize environmental harm.

3. **Addressing Legal Requirements**

   Both local and international legal requirements will be carefully considered in the development of the project. The following steps are and will be considered to ensure compliance with regulatory standards:

   - **Environmental Regulations:** The system monitors water quality in real-time and ensures compliance with local and international water quality standards. The parameters tracked (pH, temperature, TDS, turbidity) align with water quality thresholds set by regulatory bodies for safe aquaculture operations.

   - **Data Privacy Laws:** Since the system will collect and transmit real-time data, it will adhere to data privacy and protection laws, such as the General Data Protection Regulation (GDPR) in Europe or the California

Consumer Privacy Act (CCPA) in the US. Data will be encrypted during transmission, and user consent will be taken for data collection.

- **Intellectual Property:** The project will comply with intellectual property laws, ensuring that proprietary technology or software used in the system will be properly licensed or developed in-house. This will protect the project from potential legal disputes.

4. **Incorporation of Safety into the Design**

   Safety was a core principle in every step of the design process and how safety was and will be incorporated into the design is described below:

   - **Redundancy and Backup Systems:** To ensure reliability, backup sensors and data processing redundancies will be included. This will prevent any single point of failure from causing critical issues in monitoring or predictions.

   - **Real-Time Alerts:** The system is designed to send alerts to operators when water parameters exceed safe thresholds, allowing for quick corrective action to ensure the safety of both plankton populations and the broader environment.

   - **Manual Override and Control:** In case of system malfunctions or extreme conditions, manual overrides will allow aquaculture operators to regain control of the system and make adjustments as needed, ensuring human oversight and safety.

5. **Addressing Ethical Concerns**

   In addition to safety and legal compliance, ethical principles were and will be incorporated to ensure that the system benefits society at large:

   - **Sustainability:** By optimizing water quality and plankton growth, the system will support sustainable aquaculture practices, reducing the need for harmful chemicals and minimizing waste. This aligns with ethical considerations for environmental sustainability and responsible resource management.

- **Fair Access:** Ethical considerations were also made to ensure that the system can be scaled and adapted for small-scale or local farmers, providing access to technology that promotes sustainable aquaculture, even in resource-limited settings.

5.4 Brief Summary

The IoT-based phytoplankton and zooplankton growth monitoring system for smart aquaculture is designed to achieve the following outcomes:

1. **Real-time Monitoring:** The system will provide real-time tracking of critical water parameters such as pH, temperature, TDS, and turbidity through embedded sensors. This allows fish farm owners to continuously monitor water conditions and respond promptly to changes that could affect plankton growth.

2. **Accurate Plankton Growth Prediction:** Machine learning algorithms such as RF, SVM, CNN, RNN, KNN, XGB Boosting, and Gradient Boosting have already been implemented to analyze the relationship between the collected water data and plankton growth. The system can predict plankton growth with a high degree of accuracy based on these correlations, ensuring optimal management of aquaculture ecosystems.

3. **Data-Driven Decision Making:** By leveraging the data collected, fish farmers will be able to make informed decisions regarding water quality management, ensuring sustainable plankton growth and minimizing the risks of overfeeding, chemical overuse, or environmental imbalances.

4. **Mobile App Development:** The system will eventually integrate with a mobile application, which will provide users with real-time alerts, data visualizations, and predictive insights. This will make managing aquaculture systems more accessible and efficient for fish farmers.

5. **Sustainability and Resource Optimization:** The system is expected to promote sustainable aquaculture practices by optimizing resource usage such as feed and chemicals, reducing waste, and preventing harmful environmental impacts like algal blooms or oxygen depletion.

Overall, the project aims to enhance productivity in aquaculture by ensuring optimal plankton growth conditions while maintaining environmental sustainability through smart, data-driven insights.

5.5 Future works

The Future Works for the IoT-Embedded Phytoplankton and Zooplankton Growth Monitoring System can be but not limited to:

1. **Mobile Application Development:** Create a mobile app that interfaces with the system and provides real-time data visualization, warnings, and plankton growth projections. The software will provide farmers with an easy-to-use interface for remote monitoring of water parameters and system operation.

2. **Enhanced Data Analytics:** Refine the machine learning algorithms by adding additional data and increasing model accuracy. This might include adding more water characteristics, such as salinity, and investigating new correlations to improve plankton growth predictions.

3. **Scalability and System Optimization:** Test the system at various scales of aquaculture operations, from small to large. Optimize the system to ensure reliability and accuracy in a variety of situations and climates.

4. **Integration with Control Systems:** Develop features that allow the system to autonomously control water conditions by adjusting parameters such as aeration, nutrient levels, or water filtration, minimizing manual intervention.

5. **Regulatory Compliance and Standards:** Verify that the system conforms with local, national, and global laws concerning data privacy, water quality, and aquaculture. Update as needed to comply with changing environmental and regulatory standards.

6. **IoT Device Enhancements:** Enhance the battery life, accuracy, and dependability of the sensor for long-term, low-maintenance operation. For improved performance, consider either introducing new kinds of sensors or updating communication protocols.

7. **Field Testing and User Feedback:** Conduct field trials in diverse aquaculture environments to validate the system's performance under real-world conditions. Gather user feedback to improve usability, efficiency, and system features.

8. **Machine Learning Model Integration:** Integrate additional machine learning models, such as time series forecasting, to predict long-term plankton growth trends and offer advanced recommendations for water management and resource optimization.

By addressing these future works, the project will continue evolving into a fully automated, scalable, and sustainable solution for smart aquaculture management, ensuring both environmental and economic benefits.

References

Abbas, F., Cai, Z., Shoaib, M., Iqbal, J., Ismail, M., Arifullah, ... Albeshr, M. F. (2024, 3). Machine learning models for water quality prediction: A comprehensive analysis and uncertainty assessment in mirpurkhas, sindh, pakistan. *Water*, *16*, 941. doi: 10.3390/w16070941

Abdel-Wahed, R. K., Shaker, I. M., Elnady, M. A., & Soliman, M. A. M. (2018, 5). Impact of fish- farming management on water quality, plankton abundance and growth performance of fish in earthen ponds. *Egyptian Journal of Aquatic Biology and Fisheries*, *22*, 49-63. doi: 10.21608/ejabf.2018.7705

Abirami, K., Radhakrishna, P. C., & Venkatesan, M. A. (2023, 4). Water quality analysis and prediction using machine learning. In *2023 ieee 12th international conference on communication systems and network technologies (csnt)* (p. 241-245). IEEE. doi: 10.1109/CSNT57126.2023.10134661

Akhlaq, M., Ellahi, A., Niaz, R., Khan, M., Sammen, S. S., & Scholz, M. (2024, 7). Comparative analysis of machine learning algorithms for water quality prediction. *Tellus A: Dynamic Meteorology and Oceanography*, *76*, 177-192. doi: 10.16993/tellusa.4069

Altium. (2023). *Best circuit design software — schematic capture.* https://resources.altium.com/.

Amazon Web Services. (n.d.). *How aws iot works - aws iot core.* https://aws.amazon.com/iot-core/how-it-works/. (n.d.)

Anani, O. A., Adetunji, C. O., Olugbemi, O. T., Hefft, D. I., Wilson, N., & Olayinka, A. S. (2022). Iot-based monitoring system for freshwater fish farming: Analysis and design. In (p. 505-515). Elsevier. doi: 10.1016/B978-0-12-823694-9.00026-8

AquaHow. (2023). *How tds meters work.* https://www.aquahow.com.

Araujo, G. S., Pacheco, D., Cotas, J., da Silva, J. W. A., Saboya, J., Moreira, R. T., & Pereira, L. (2022, 5). Plankton: Environmental and economic importance for a sustainable future. IntechOpen. doi: 10.5772/intechopen.100433

Aravinth, A., Kannan, R., Chinnadurai, G., Manickam, N., Raju, P., Perumal, P., & Santhanam, P. (2023, 3). Temporal changes in plankton diversity in relation to hydrographical characteristics at perumal lake, cuddalore district, tamil nadu, india. *The Journal of Basic and Applied Zoology*, *84*, 13. doi: 10.1186/s41936-023-00337-7

Arduino. (n.d.). *The arduino software (ide).* Arduino Documentation. (n.d.)

Atlas Scientific. (2023). *What is a temperature sensor?* https://atlas-scientific.com.

Brander, K. (2017). *Plankton and fisheries* (Vol. 1). Oxford University Press. doi: 10.1093/oso/9780199233267.003.0008

Components101. (2024). *Esp8266 esp-01 wifi module: Features and applications.* https://components101.com.

Control.com. (2022). *Introduction to arduino: History, hardware, and software.* Control.com.

Corporate Finance Institute. (2024). *Excel definition - what is microsoft excel? overview, definition.* https://corporatefinanceinstitute.com/resources/knowledge/finance/excel-definition/.

Das, A. N., & Sharma, D. K. (2022, 5). Analysis of planktonic abundance and its correlation to fish diversity in dhir beel (oxbow lake), assam, india. *Egyptian Journal of Aquatic Biology and Fisheries*, *26*, 725-743. doi: 10.21608/ejabf.2022.246433

Ding, S., Chen, M., Gong, M., Fan, X., Qin, B., Xu, H., . . . Zhang, C. (2018, 6). Internal phosphorus loading from sediments causes seasonal nitrogen limitation for harmful algal blooms. *Science of The Total Environment*, *625*, 872-884. doi: 10.1016/j.scitotenv.2017.12.348

Duré, G. A. V., Simões, N. R., de Souza Magalhães Braghin, L., & Ribeiro, S. M. M. S. (2021, 11). Effect of eutrophication on the functional diversity of zooplankton in shallow ponds in northeast brazil. *Journal of Plankton Research*, *43*, 894-907. doi: 10.1093/plankt/fbab064

El-Feky, M. M. M., Alprol, A. E., Heneash, A. M. M., Abo-Taleb, H. A., & Omer, M. Y. (2019, 1). Evaluation of water quality and plankton for mahmoudia canal in northern west of egypt. *Egyptian Journal of Aquatic Biology and Fisheries*, *22*, 461-474. doi: 10.21608/ejabf.2019.26384

Filstrup, C. T., Wagner, T., Oliver, S. K., Stow, C. A., Webster, K. E., Stanley, E. H., & Downing, J. A. (2018, 3). Evidence for regional nitrogen stress on chlorophyll a in lakes across large landscape and climate gradients. *Limnology and Oceanography*, *63*. doi: 10.1002/lno.10742

Google Research. (2023). *Colaboratory: Jupyter notebooks in the cloud.* https://colab.research.google.com.

Hablützel, P., Rombouts, I., Dillen, N., Lagaisse, R., Mortelmans, J., Ollevier, A., ... Deneudt, K. (2021, 12). Exploring new technologies for plankton observations and monitoring of ocean health. *Oceanography*, 20-25. doi: 10.5670/oceanog.2021.supplement.02-09

Hack the Developer. (n.d.). *What is an arduino uno? a beginner's guide.* https://hackthedeveloper.com/. (n.d.)

Hemal, M. M., Rahman, A., Nurjahan, Islam, F., Ahmed, S., Kaiser, M. S., & Ahmed, M. R. (2024, 6). An integrated smart pond water quality monitoring and fish farming recommendation aquabot system. *Sensors*, *24*, 3682. doi: 10.3390/s24113682

Howtodojo. (n.d.). *What is digitalocean?* https://www.howtodojo.com/digitalocean/. (n.d.)

Islam, S. M. R., Sefat, M. M. H., Habib, M. A., & Kibria, M. G. (2020). Internet of things-based household water quality monitoring system using wireless sensor. In (p. 567-576). doi: 10.1007/978-981-15-3607-6_45

Karunasena, D. (2023). *What is flutter and why it's used for mobile app development.* Cult by Honeypot, https://cult.honeypot.io.

Khan, S., Sultana, S., Hena, S. M., Ahmed, M. S., Sultana, M. S., Akhi, M. S. N., ... Hossain, M. M. (2023). Seasonal dynamics of zooplankton in a eutrophic fish pond of bangladesh in relation to environmental factors. *Journal of Aquaculture Marine Biology*, *12*, 129-136. doi: 10.15406/jamb.2023.12.00365

Kumar, S. M. R., Mohammed, T. K., Rao, S. S., P, D. A. R., Bakhare, R., Kumar, A., & Sarkar, S. (2023, 2). Iot based fish pond monitoring system to enhance its productivity. In *2023 3rd international conference on innovative practices in technology and management (iciptm)* (p. 1-6). IEEE. doi: 10.1109/ICIPTM57143.2023.10117894

Lomartire, S., Marques, J. C., & Gonçalves, A. M. (2021, 10). The key role of zooplankton in ecosystem services: A perspective of interaction between zooplankton and fish recruitment. *Ecological Indicators*, *129*, 107867. doi: 10.1016/j.ecolind.2021.107867

Lucid Software Inc. (n.d.). *Lucidchart*. https://www.lucidchart.com/pages/what-is-lucidchart. (n.d.)

Manoj, M., Kumar, V. D., Arif, M., Bulai, E.-R., Bulai, P., & Geman, O. (2022, 3). State of the art techniques for water quality monitoring systems for fish ponds using iot and underwater sensors: A review. *Sensors*, *22*, 2088. doi: 10.3390/s22062088

Marqo. (2024). *Getting started with google colab: A beginner's guide.* https://www.marqo.ai/blog/getting-started-with-google-colab-a-beginners-guide.

Meacon Co., Ltd. (2023). *What is a ph sensor & how does it work?* https://www.meaconsensor.com.

Musa, M., Thoyibah, A. A., Puspitaningtyas, D. A., Arsad, S., Mahmudi, M., Lusiana, E. D., ... Huda, A. S. (2022, 11). The impact of water quality on the availability of phytoplankton and growth of litopenaeus vannamei. *Journal of Water and Land Development*, 127-135. doi: 10.24425/jwld.2023.143753

Nasir, N., Kansal, A., Alshaltone, O., Barneih, F., Sameer, M., Shanableh, A., & Al-Shamma'a, A. (2022, 8). Water quality classification using machine learning algorithms. *Journal of Water Process Engineering*, *48*, 102920. doi: 10.1016/j.jwpe.2022.102920

Nayan, A.-A., Khan, M. S., Ferdaous, J., Mozumder, A. N., Alam, M. K., & Kibria, M. G. (2024). Drinking water quality analysis and prediction using lstm: Safe drinking water for school children. In (p. 499-513). doi: 10.1007/978-981-99-8937-9_34

Nayan, A.-A., Kibria, M. G., Rahman, M. O., & Saha, J. (2020, 11). River water quality analysis and prediction using gbm. In *2020 2nd international conference on advanced information and communication technology (icaict)* (p. 219-224). IEEE. doi: 10.1109/ICAICT51780.2020.9333492

News18. (2020). Lack of oxygen kills over 600 metric tonnes of fish worth rs 10 crore in bangladesh.

Odulate, D., Omoniyi, I., Alegbeleye, W., George, F., & Dimowo, B. (2017). Water quality in relation to plankton abundance and diversity in river ogun, abeokuta, southwestern nigeria. *International Journal of Environmental Health Engineering*, *6*, 3. doi: 10.4103/ijehe.ijehe_31_13

Osmotics. (2023). *Understanding tds meters.* https://www.osmotics.co.uk.

Overleaf. (n.d.). *What is latex?* https://www.overleaf.com/learn/What_is_LaTeX%3F. (n.d.)

Pulsifer, J., & Laws, E. (2021, 6). Temperature dependence of freshwater phytoplankton growth rates and zooplankton grazing rates. *Water*, *13*, 1591. doi: 10.3390/w13111591

Rai, S., Poduval, D. S., Anand, U., Verma, V., & Banerjee, S. (2024, 8). An effective smart water quality monitoring and management system using iot and machine learning. *SN Computer Science*, *5*, 846. doi: 10.1007/s42979-024-03208-2

Rashid, M. M., Nayan, A.-A., Simi, S. A., Saha, J., Rahman, M. O., & Kibria, M. G. (2021). Iot based smart water quality prediction for biofloc aquaculture. *International Journal of Advanced Computer Science and Applications*, *12*. doi: 10.14569/IJACSA.2021.0120608

Reqtest. (2018). *Requirements analysis – understanding the process & techniques.* https://reqtest.com.

Rui, L., & Qiu-jin, X. (2013). Effects of various total dissolved solids (tds) on the growth of phytoplankton. *Research of Environmental Sciences*, *26*, 1072-1078.

Saha, S., Rajib, R. H., & Kabir, S. (2018, 10). Iot based automated fish farm aquaculture monitoring system. In *2018 international conference on innovations in science, engineering and technology (iciset)* (p. 201-206). IEEE. doi: 10.1109/ICISET.2018.8745543

Sakaa, B., Elbeltagi, A., Boudibi, S., Chaffaï, H., Islam, A. R. M. T., Kulimushi, L. C., . . . Wong, Y. J. (2022, 7). Water quality index modeling using random forest and improved smo algorithm for support vector machine in saf-saf river basin. *Environmental Science and Pollution Research*, *29*, 48491-48508. doi: 10.1007/s11356-022-18644-x

Sensorex. (n.d.). *Understanding the science behind turbidity sensors and how they work.* https://sensorex.com. (n.d.)

Smith, D. W., & Piedrahita, R. H. (1988, 2). The relation between phytoplankton and dissolved oxygen in fish ponds. *Aquaculture*, *68*, 249-265. doi: 10.1016/0044-8486(88)90357-2

TrustRadius. (2024). *Compare flutter by google vs .net.* https://www.trustradius.com.

Wang, X., Li, Y., Qiao, Q., Tavares, A., & Liang, Y. (2023, 8). Water quality prediction based on machine learning and comprehensive weighting methods. *Entropy*, *25*, 1186. doi: 10.3390/e25081186

Yan, X., Zhang, T., Du, W., Meng, Q., Xu, X., & Zhao, X. (2024, 1). A comprehensive review of machine learning for water quality prediction over the past five years. *Journal of Marine Science and Engineering*, *12*, 159. doi: 10.3390/jmse12010159