

ABDELHAMID IBN BADIS UNIVERSITY - MOSTAGANEM

Faculty of Exact Sciences and Computer Science
Department of Mathematics and Computer Science
Sector: Computer Science

COMPUTER SCIENCE MASTER'S PROJECT REPORT

Specialty: **Artificial Intelligence for Internet
of Things (AI for IoT)**

THEME:

**“IOT-Based Intelligent Monitoring System for
Aquaculture”**

Student: Belhamiti Nada Zahra

Supervisor: Mr. Mechaoui Moulay Dris

Academic Year: 2024-2025

Abstract

Aquaculture has become an essential contributor to global food security and sustainable protein production. However, many aquaculture environments continue to rely on manual monitoring techniques, which are time-consuming, imprecise, and reactive in nature. This thesis addresses the need for real-time environmental monitoring by developing a data acquisition system based on Internet of Things (IoT) technologies.

The system is designed for small to medium-scale freshwater fish farms and focuses on the continuous collection of key water quality parameters, including temperature, pH, and total dissolved solids. Using a Raspberry Pi microcontroller and a set of low-cost sensors, the system enables real-time data capture and remote access, providing fish farmers with more timely and informed decision-making tools.

The work begins with an overview of aquaculture practices and the limitations of traditional monitoring methods, followed by a review of existing intelligent aquaculture systems. It then presents the hardware and software setup used for data acquisition and introduces the theoretical framework for applying artificial intelligence (AI), including machine learning (ML) and deep learning (DL) models. These technologies may be used to analyze collected data and support intelligent decision-making within the scope of this thesis.

This project lays the groundwork for modern, data-driven aquaculture management by introducing accessible digital monitoring tools and exploring the potential of IoT and AI integration.

Résumé

L'aquaculture est devenue un pilier essentiel de la sécurité alimentaire mondiale et de la production durable de protéines. Cependant, de nombreuses exploitations aquacoles continuent de s'appuyer sur des techniques de surveillance manuelle, souvent chronophages, imprécises et réactives. Ce mémoire vise à répondre à ce besoin en développant un système d'acquisition de données en temps réel basé sur les technologies de l'Internet des Objets (IoT).

Le système proposé est conçu pour les fermes piscicoles d'eau douce de petite et moyenne taille. Il permet la collecte continue de paramètres clés de la qualité de l'eau,

notamment la température, le pH et matières dissoutes totales. Grâce à un microcontrôleur Raspberry Pi et à des capteurs à faible coût, le système permet la capture et l'accès à distance aux données, offrant ainsi aux pisciculteurs des outils d'aide à la décision plus rapides et plus fiables.

Le travail commence par une présentation des pratiques aquacoles et des limites des méthodes traditionnelles de surveillance, suivie d'une revue des systèmes intelligents existants. Il détaille ensuite l'architecture matérielle et logicielle du système proposé, tout en introduisant les bases théoriques de l'intelligence artificielle (IA), y compris l'apprentissage automatique (ML) et l'apprentissage profond (DL). Ces technologies pourraient être utilisées pour analyser les données collectées et appuyer la prise de décision intelligente dans le cadre de ce travail.

Ce projet jette les bases d'une gestion aquacole moderne et fondée sur les données, en introduisant des outils numériques accessibles et en explorant le potentiel d'intégration de l'IoT et l'IA.

ملخص

العديد من مزارع الأسماك تعتمد على أساليب المراقبة اليدوية، وهي أساليب تستغرق وقتاً طويلاً وتفتقر إلى الدقة، وغالباً ما تكون استجابتها متأخرة. يهدف هذا البحث إلى معالجة هذه المشكلة من خلال تطوير نظام لجمع البيانات في الوقت الحقيقي باستخدام تقنيات إنترنت الأشياء (IoT).

تم تصميم النظام ليتناسب مع مزارع الأسماك الصغيرة والمتوسطة الحجم في المياه العذبة، ويركز على جمع بيانات مستمرة حول معايير جودة المياه مثل درجة الحرارة، ودرجة الحموضة (pH)، وإجمالي المواد الصلبة الذائبة. يعتمد النظام على وحدة تحكم Raspberry Pi ومجموعة من المجسات منخفضة التكلفة، ويتيح الوصول إلى البيانات عن بُعد، مما يساعد المزارعين على اتخاذ قرارات أفضل وفي الوقت المناسب.

يبدأ العمل بمقدمة حول تطور ممارسات تربية الأسماك، متبوعة بمراجعة للأنظمة الذكية المستخدمة في هذا المجال. كما يعرض مكونات النظام من الناحيتين المادية والبرمجية، ويقدم الأساس النظري لاستخدام الذكاء الاصطناعي (AI)، بما في ذلك التعلم الآلي (ML) والتعلم العميق (DL). وقد تُستخدم هذه التقنيات لتحليل البيانات المجمعة ودعم اتخاذ القرار الذكي ضمن نطاق هذا المشروع.

يهدف هذا المشروع إلى تمهيد الطريق نحو إدارة ذكية ومبنية على البيانات في قطاع تربية الأحياء المائية، من خلال إدخال أدوات رقمية فعالة واستكشاف إمكانية دمج تقنيات إنترنت الأشياء و الذكاء الاصطناعي.

Acknowledgements

I would like to express my sincere gratitude to my thesis supervisor, Mr. MECHAOUI Moulay Driss, for their invaluable guidance, support, and continuous encouragement through this work.

I am also deeply thankful to Professor Mr. MOUSSA Mohamed, whose valuable advice, technical insights, and generosity with their time greatly contributed to the progress and clarity of this project.

A heartfelt thank you goes to my family and close friends, whose constant encouragement, patience, and understanding gave me strength during challenging moments.

Finally, I would like to thank everyone who contributed, directly or indirectly, to the completion of this thesis and to the development of my academic and personal growth during this time.

List of Abbreviations

Abbreviation	Full Expression	Page
IoT	Internet of Things	1
AI	Artificial Intelligence	1
RAS	Recirculating Aquaculture Systems	8
IMTA	Integrated Multi-Trophic Aquaculture	10
RPi	Raspberry Pi	14
ML	Machine Learning	16
SVM	Support Vector Machines	16
GBM	Gradient Boosting Machines	16
DSS	Decision Support Systems	16
LoRa	Long Range	17
PCB	Printed Circuit Board	17
ORP	Oxidation-reduction Potential	19
IAMCS	IoT-Based Aquaculture Monitoring and Control System	20
RTT	Round-Trip Time	20
QoS	Quality of Service	20
RF	Random Forest	22
NN	Neural Network	22
DO	Dissolved Oxygen	22
GPIO	General Purpose Input/output	24
ARM	Advanced RISC Machine	24

RISC	Reduced Instruction Set Computer	24
RAM	Random-access Memory	24
USB	Universal Serial Bus	25
HVAC	Heating, Ventilation, and Air Conditioning	26
ADC	Analog to Digital Converter	30
SSH	Secure Shell	31
AIoT	Artificial Intelligence of Things	38
DL	Deep Learning	42
PCA	Component Analysis	41
RL	Reinforcement Learning	41
ANN	Artificial Neural Networks	42
CNN	Convolutional Neural Networks	43
RNN	Recurrent Neural Networks	43
DQN	Deep Q-Networks	43
DDPG	Deep Deterministic Policy Gradient	43
CPU	Central Processing Unit	45
GPU	Graphics Processing Unit	45

List of Figures

Figure Number	Figure's Title	Page
Figure 1	Different types of aquaculture	4
Figure 2	Site Selection in Aquaculture	7
Figure 3	IoT-based Water Quality Monitoring System	7
Figure 4	RAS System Used for Stocking Density Experiments	8
Figure 5	Insect-based Protein Production for Sustainable Aquafeeds	9
Figure 6	Disease Management Strategies	9
Figure 7	Recirculating Aquaculture Systems (RAS)	10
Figure 8	Conceptual Diagram of the Integrated Multi-trophic Aquaculture (IMTA) System	10
Figure 9	Aquaponics Technologies	11
Figure 10	Tilapia Production Analytics	11
Figure 11	Harvesting Operations	12
Figure 12	Regulatory Costs by Type of Cost	12
Figure 13	Technology and Innovation in Aquaculture	13
Figure 14	Block Diagram	15
Figure 15	Raspberry pi 4 Model B	25
Figure 16	Temperature Sensor: DS18B20	26
Figure 17	pH Sensor: ph-4502c	27
Figure 18	TDS Meter V1.0	28
Figure 19	Breadboard	28
Figure 20	Resistor	29

Figure 21	Jumper Wires	30
Figure 22	ADS1015 Converter	30
Figure 23	PuTTY Emulator Logo	32
Figure 24	VS Code Logo	32
Figure 25	SSH Command to use IP Address of RPi	34
Figure 26	Aquaculture Intelligent System	35
Figure 27	Example of a Dataset from our Intelligent System	36
Figure 28	Conceptual integration of IoT with AI for intelligent systems	39
Figure 29	Machine Learning Process	40
Figure 30	Types of Machine Learning	42
Figure 31	Deep Learning's Layers.	43
Figure 32	Machine Learning Paradigms	44
Figure 33	Data Loading and Preprocessing	50
Figure 34	True vs Predicted Survival Rate using Random Forest Regressor	51
Figure 35	True vs Predicted Average Fish Weight using Random Forest Regressor	52
Figure 36	Confusion Matrix of Random Forest Classifier for Health Status Classification	53
Figure 37	Feature Importance of Random Forest Classifier in Health Status Prediction	54
Figure 38	Confusion Matrix of XGBoost Classifier for Health Status Classification	55
Figure 39	Confusion Matrix of SVM Classifier for Health Status Classification	56
Figure 40	Anomaly Detection in Temperature and Dissolved Oxygen using Isolation Forest	57
Figure 41	Observed vs Forecasted Water Temperature using ARIMA	57

Figure 42	True vs Predicted Dissolved Oxygen using LSTM	58
-----------	---	----

List of Tables

Table Number	Table's Title	Page
Table 1	Comparison of Existing Systems	22
Table 2	ML and DL comparison	45
Table 3	Comparative Summary of Machine Learning and Deep Learning Model Performance on the Tilapia Aquaculture Dataset.	59

Contents

General Introduction	1
Chapter 1: Aquaculture	3
1.1 Introduction	3
1.2 Aquaculture Definition	3
1.3 Types of Aquaculture	4
1.4 The Role of Aquaculture in National Policies	4
1.4.1 Food Security and Nutritional Improvements	4
1.4.2 Trade and Foreign Exchange	5
1.4.3 Producers' Incomes	5
1.4.4 Employment opportunities	5
1.5 Traditional Fish Farming Management Practices	5
1.5.1 Species selection based on local climate and resources	5
1.5.2 Integration with agriculture	5
1.5.3 Mixed Fish Culture	6
1.5.4 Low-Tech Water Quality Control	6
1.5.5 Feeding with Organic Inputs	6
1.5.6 Use of Community-Shared Wisdom	6
1.6 Modern Fish Farming Management Practices	6
1.6.1 Site Selection	6
1.6.2 Water Quality Management	7
1.6.3 Stocking Density Regulation	8
1.6.4 Nutritional Feeding Practices	8
1.6.5 Health Management and Disease Control	9

1.6.6 Environmental Sustainability	10
1.6.7 Technology Integration	11
1.6.8 Efficient Harvesting and Processing	12
1.6.9 Regulatory Compliance	12
1.6.10 Continuous Learning and Innovation	13
1.7 Problem Statement	13
1.8 Scope and Limitations	14
1.9 Research Methodology	14
1.9.1 System Methodology	14
1.9.2 Methodology Diagram	15
1.10 Conclusion	15
Chapter 2: Literature Review	16
2.1 Introduction	16
2.2 Existing IoT-Based Intelligent Aquaculture Systems	16
2.2.1 IoT-Based Environmental Control System for Fish Farms	16
2.2.2 Smart Aquaponics with AI and IoT Integration	17
2.2.3 IoT-Based Automated Smart Water Quality Control System for Tilapia Culture	18
2.2.4 IoT-Based Intensive Aquaculture Monitoring with Automatic Water Correction	19
2.2.5 Aquaculture Monitoring and Control System: An IoT-Based Approach	19
2.2.6 IoT-Based Aquaculture Monitoring and Control System	20
2.3 Comparative Analysis of Existing IoT-Based Aquaculture Monitoring Systems	22
2.4 Conclusion	23
Chapter 3: Data Acquisition	24
3.1 Introduction	24
3.2 Internet of Things Definition	24
3.3 Hardware and Software Tools	24

3.3.1 Raspberry Pi	24
3.3.2 IoT Sensor	25
3.3.3 Temperature Sensor: DS18B20	26
3.3.4 pH Sensor: PH-4502C	27
3.3.5 TDS Meter V1.0: KS0429	28
3.3.6 Breadboard	28
3.3.7 Resistors	29
3.3.8 Jumpwires	30
3.3.9 ADC Converter	30
3.3.9.1 ADS1015 Definition	30
3.3.9.2 ADS1015 Pin Configuration	31
3.3.10 Putty Emulator	31
3.3.11 Visual Code Studio	32
3.4 The Process	33
3.4.1 Raspberry Pi Setup	33
3.4.1.1 Hardware Used	33
3.4.1.2 Installing Raspberry Pi OS	33
3.4.2 Remote Development Setup	34
3.4.3 Sensor Integration and Wiring	34
3.4.3.1 DS18B20 – Water Temperature Sensor (Digital, 1-Wire)	34
3.4.3.2 pH-4502C – pH Sensor (Analog via ADS1015)	34
3.4.3.3 TDS Meter V1.0 – Conductivity Sensor (Analog via ADS1015)	35
3.4.4 Development and Data Logging	36
3.5 Conclusion	37
Chapter 4: Chapter 4 : Model Implementation	38
4.1 Introduction	38

4.2 4.2 Artificial Intelligence Definition	38
4.3 Artificial Intelligence for Internet of Things	38
4.4 Machine Learning	39
4.4.1 Definition	39
4.4.2 ML Types	40
4.4.2.1 Supervised Learning	40
4.4.2.1.1 Classification	40
4.4.2.1.2 Regression	41
4.4.2.2 Unsupervised Learning	41
4.4.2.2.1 Clustering	41
4.4.2.2.2 Association	41
4.4.2.3 Semi-Supervised Learning	41
4.4.2.4 Reinforcement Learning	41
4.4.3 Machine Learning Models	42
4.5 Deep Learning	42
4.5.1 Definition	42
4.5.2 DL in ML Paradigms	43
4.5.2.1 Supervised Learning	43
4.5.2.2 Unsupervised Learning	43
4.5.2.3 Reinforcement Learning	43
4.5.3 Comparison Between Machine Learning and Deep Learning	45
4.6 Chosen Dataset	46
4.6.1 Justification for Dataset Selection	46
4.6.2 Dataset Used	46
4.7 ML and DL Algorithms Choices	47
4.7.1 Random Forest Regressor	47

4.7.2 Random Forest Classifier	47
4.7.3 XGBoost	48
4.7.4 Support Vector Machine (SVM)	48
4.7.5 Isolation Forest	48
4.7.6 Autoregressive Integrated Moving Average (ARIMA)	48
4.7.7 Long Short Term Memory (LSTM)	49
4.8 Preprocessing	49
4.8.1 Python	49
4.8.2 Packages Used	49
4.9 Results and Plots	51
4.9.1 Survival Rate Prediction (Random Forest Regressor)	51
4.9.2 Average Fish Weight Prediction (Random Forest Regressor)	52
4.9.3 Health Status Classification (Random Forest Classifier)	52
4.9.3.1 Confusion Matrix	52
4.9.3.2 Feature Importance	53
4.9.4 Health Status Classification (XGBoost Classifier)	54
4.9.5 Health Status Classification (SVM Classifier)	55
4.9.6 Anomaly Detection (Isolation Forest)	56
4.9.7 Temperature Forecasting (ARIMA)	57
4.9.8 Multivariate Time Series Forecasting (LSTM)	58
4.10 Summary of ML and DL Models Applied to the Tilapia Dataset	59
4.11 Conclusion	60
General Conclusion	61
Perspectives and Future Enhancements	62
Bibliography	63

General Introduction

There's an old saying in fishkeeping: *“Good fish deserves good water.”* This simple yet profound phrase captures the essence of what underlies successful aquaculture practices: water quality. For aquatic organisms, the surrounding water is not just habitat but also the medium through which they breathe, feed, and excrete waste. Maintaining optimal water conditions is therefore essential not only for fish health but also for overall farm productivity and sustainability.

As aquaculture continues to expand globally, the need for efficient and reliable monitoring of environmental parameters has become increasingly urgent. Traditional methods, which rely on manual sampling and visual inspection, are often inadequate in addressing the dynamic nature of water quality. These methods are time-consuming, labor-intensive, and reactive, often leading to delayed responses that can negatively impact fish health and survival rates.

In this context, the integration of modern technologies such as the Internet of Things (IoT), artificial intelligence (AI), and embedded systems represents a transformative opportunity. These technologies enable real-time monitoring, data-driven decision-making, and eventually, automated control of aquaculture environments. However, in many regions, including Algeria, the adoption of such tools remains limited or nonexistent. This thesis aims to bridge that gap by developing an IoT-based data acquisition system tailored to small and medium-scale aquaculture settings, serving as a foundation for future intelligent aquaculture solutions.

This report is structured into four chapters:

Chapter 1 provides a general introduction to aquaculture, its evolution from traditional to modern practices, and the importance of water quality management. It also presents the problem statement, research objectives, and limitations of the study.

Chapter 2 reviews related works and existing systems in the field of aquaculture monitoring. It analyzes several IoT-based approaches, highlights their features, and provides a comparative analysis between each one of them.

Chapter 3 focuses on the development of the data acquisition system. It introduces the concept of the Internet of Things in aquaculture, describes the selected sensors and hardware, and explains how the components are connected and how data is collected.

Chapter 4 introduces the principles of artificial intelligence, and explores its potential applications in intelligent aquaculture monitoring and decision-making.

Chapter 1: Aquaculture

1.1 Introduction

In this chapter, we are presenting a general introduction about aquaculture and the contextual background of the research. Aquaculture has become an increasingly important sector for global food security, economic development, and sustainable resource use. As the demand for aquatic products continues to rise, countries around the world are integrating aquaculture into their national strategies to enhance nutrition, create employment, and strengthen rural economies. This chapter begins by exploring the historical evolution of aquaculture, from traditional farming practices rooted in indigenous knowledge to modern, technology-driven systems. It highlights the role of aquaculture in policy-making, the challenges associated with conventional monitoring methods, and the need for intelligent, automated solutions. The chapter also outlines the motivation behind this research and defines its objectives and scope.

1.2 Aquaculture Definition

Aquaculture is the controlled cultivation of marine and freshwater fish, molluscs, aquatic plants, algae and other organisms for a variety of purposes. Aquaculture involves growing aquatic organisms in a controlled environment, rather than relying on wild sources for their extraction or production. Although generally intended for food production, aquatic organisms can also be bred for use in aquaponics systems, to restore endangered species, to enhance valuable wild species or to produce non-food items such as pearls, biofuels and bioplastics. [1]

1.3 Types of Aquaculture

According to El-Sayed [2], he defines aquaculture as the cultivation of aquatic organisms, including fish, molluscs, crustaceans, and aquatic plants. The categories of aquaculture are simplified in the following figure.

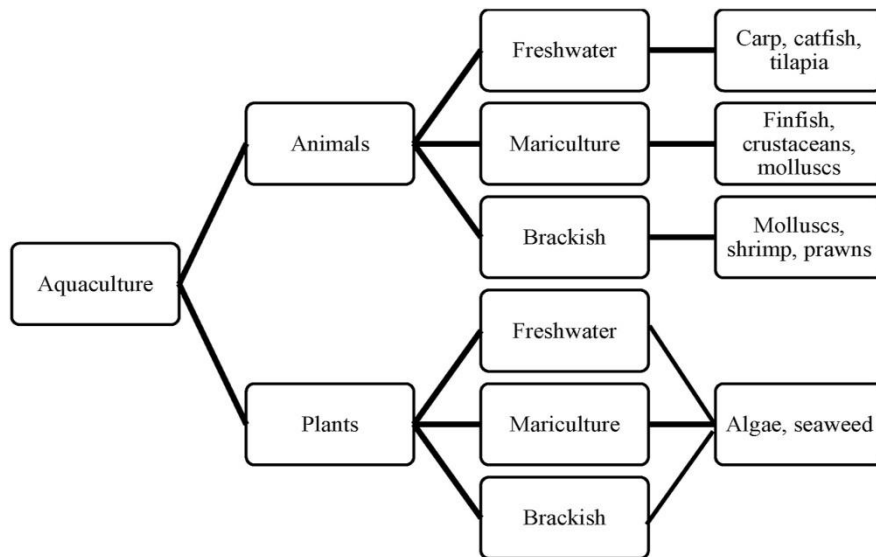


Figure 1: Different types of aquaculture. [2]

1.4 The Role of Aquaculture in National Policies

Aquaculture has become an integral component of national development strategies, contributing to various socio-economic objectives. Its multifaceted role encompasses enhancing food security, generating foreign exchange, increasing producers' incomes, and creating employment opportunities. [3]

1.4.1 Food security and nutritional improvements

Governments prioritize aquaculture to bolster domestic food supplies and improve nutrition, especially in regions with limited access to wild fish stocks. By facilitating the local production of aquatic protein, aquaculture addresses dietary deficiencies and supports public health initiatives.

1.4.2 Trade and foreign exchange

In countries where food security is less pressing, aquaculture serves as a means to earn foreign exchange through the export of high-value species. This approach not only contributes to national income but also integrates these countries into global seafood markets.

1.4.3 Producers' incomes

Enhancing the profitability of aquaculture ventures is a common policy objective. By ensuring that aquaculture operations are economically viable, governments aim to improve the livelihoods of fish farmers and encourage sustainable industry growth.

1.4.4 Employment opportunities

While aquaculture is not inherently labor-intensive, it offers employment prospects in associated sectors such as feed production, processing, and distribution. Policies often focus on these ancillary benefits to stimulate rural economies and reduce unemployment.

In summary, aquaculture's role in national policies is multifaceted, addressing critical areas of food security, economic development, and employment. Its strategic importance is expected to grow as global demand for aquatic products increases and capture fisheries face sustainability challenges.

1.5 Traditional Fish Farming Management Practices

Traditional fish farming in India [4] has been shaped by centuries of cultural, ecological, and economic practices. Farmers typically rely on natural water sources such as ponds, lakes, rivers, and floodplains, using local species adapted to regional conditions. These methods require minimal technological input and make use of indigenous knowledge, with decisions guided by experience and seasonal cues.

Key traditional management strategies include:

1.5.1 Species selection based on local climate and resources: often featuring fish like catla, rohu, and mrigal.

1.5.2 Integration with agriculture: where pond water is used to irrigate crops and agricultural waste serves as natural feed.

1.5.3 Mixed fish culture: allowing different species to coexist and utilize different ecological niches.

1.5.4 Low-tech water quality control: including bamboo aerators and natural additives to maintain pond health.

1.5.5 Feeding with organic inputs: like rice bran, oilcakes, or kitchen waste instead of manufactured feed.

1.5.6 Use of community-shared wisdom: especially for disease prevention and harvesting cycles.

Traditional systems such as rice–fish farming, pokkali culture, and composite fish culture promote sustainability by minimizing environmental impact, preserving biodiversity, and using locally available resources. However, these systems also face challenges, including limited yields, dependency on natural conditions, and market access difficulties.

1.6 Modern Fish Farming Management Practices

Modern fish farming practices have evolved to enhance productivity, fish health, and environmental sustainability by adopting structured management approaches and scientific innovations. These practices are built on principles of precision, efficiency, and ecological responsibility, and they reflect the transition from subsistence-level aquaculture to commercial-scale production. Key components of modern management include:

1.6.1 Site Selection: The success of a fish farm begins with choosing an appropriate site. Factors such as water temperature, dissolved oxygen levels, proximity to markets, land availability, and overall environmental compatibility are carefully evaluated to support optimal farm performance and reduce ecological risks.



Figure 2: Site Selection in Aquaculture. [5]

1.6.2 Water Quality Management: Water quality plays a vital role in fish health and growth. Farmers regularly monitor parameters like pH, ammonia, temperature, and oxygen concentration. Techniques such as aeration, filtration, and water exchange are applied to maintain balanced aquatic environments.

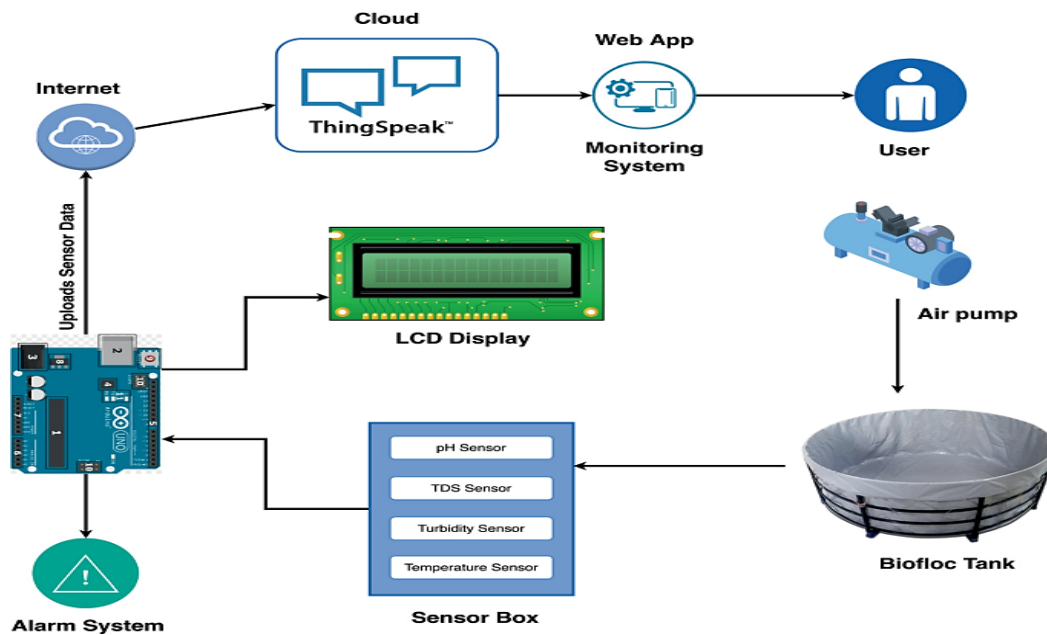


Figure 3: IoT-based Water Quality Monitoring System. [6]

1.6.3 Stocking Density Regulation: Overcrowding can lead to stress, disease, and inhibited growth. Modern aquaculture systems calculate and adjust stocking densities based on the biological needs of the species being farmed to ensure sufficient space and minimize competition for resources.

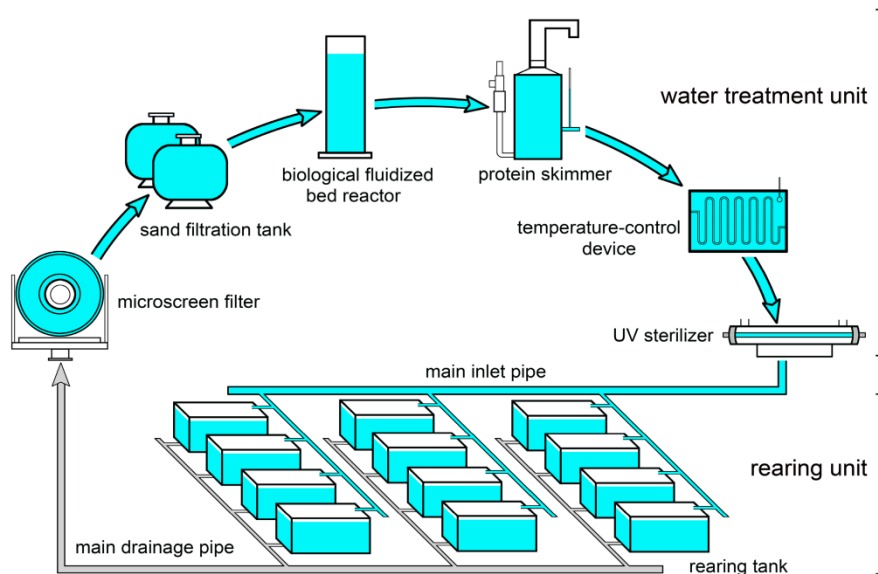


Figure 4: RAS System Used for Stocking Density Experiments. [7]

1.6.4 Nutritional Feeding Practices: Formulated feeds enriched with proteins, vitamins, minerals, and essential nutrients are now commonly used to support rapid and healthy fish development. Feeding schedules and quantities are optimized to prevent waste, improve feed conversion ratios, and maintain water quality.

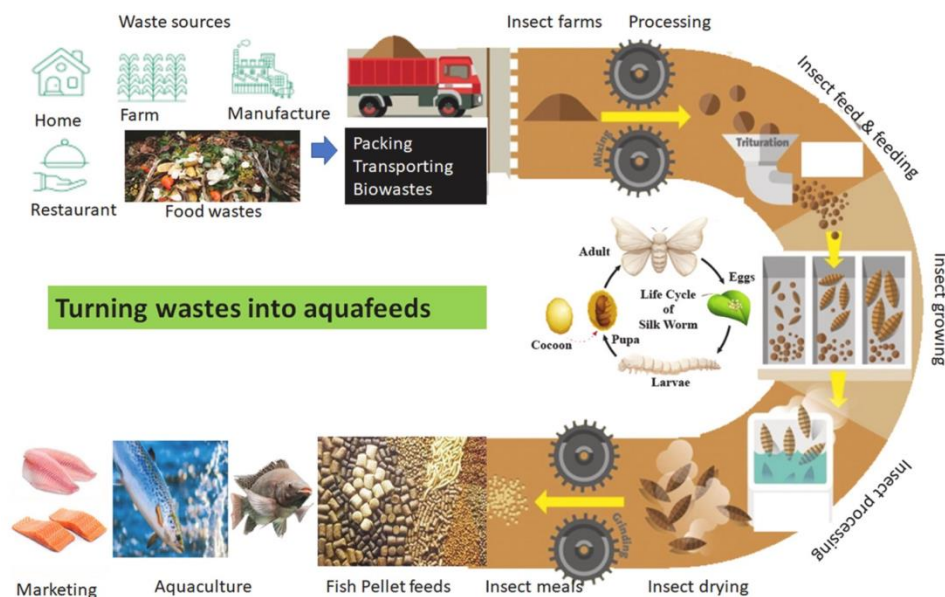


Figure 5: Insect-based Protein Production for Sustainable Aquafeeds. [8]

1.6.5 Health Management and Disease Control: Biosecurity protocols, vaccinations, regular health screenings, and controlled treatment processes are implemented to prevent disease outbreaks. Early detection and prompt intervention help maintain a healthy fish population and reduce losses.

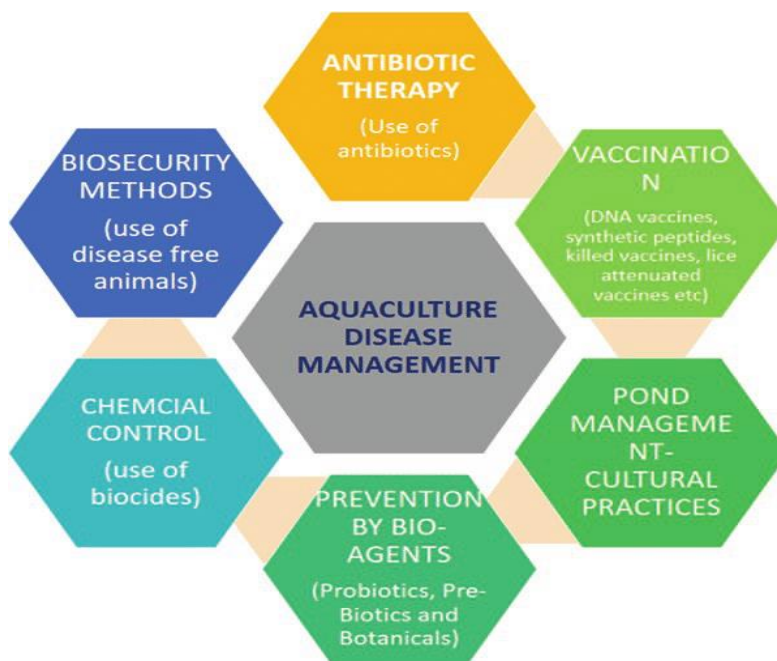


Figure 6: Disease Management Strategies. [9]

1.6.6 Environmental Sustainability: Modern systems are designed to minimize pollution and ecological disruption. Technologies like Recirculating Aquaculture Systems (RAS) and Integrated Multi-Trophic Aquaculture (IMTA) reduce water use and process nutrient waste more efficiently, contributing to lower environmental footprints.

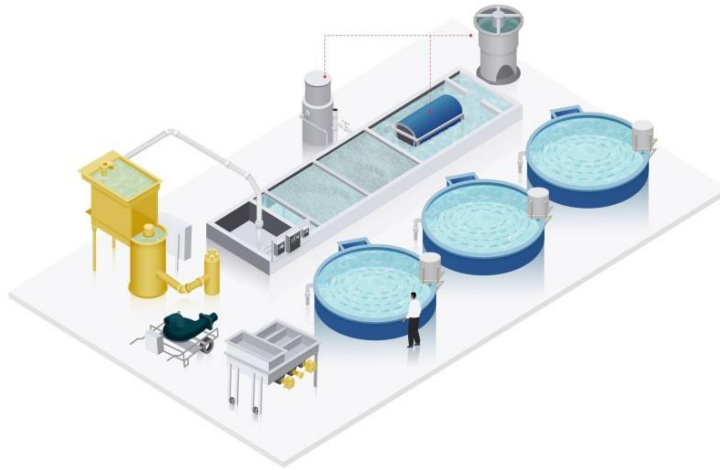


Figure 7: Recirculating Aquaculture Systems (RAS). [10]

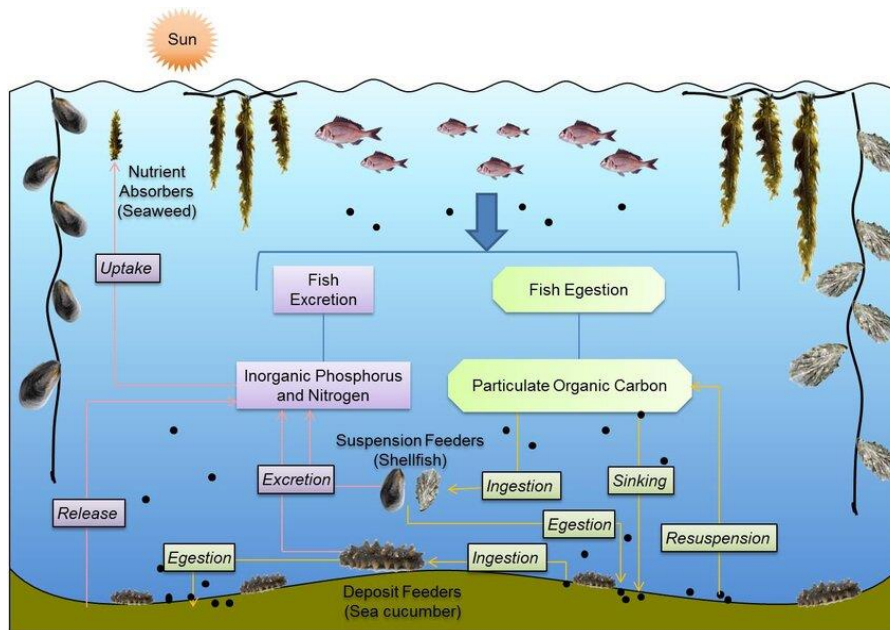


Figure 8: Conceptual Diagram of the Integrated Multi-trophic Aquaculture (IMTA) System. [11]

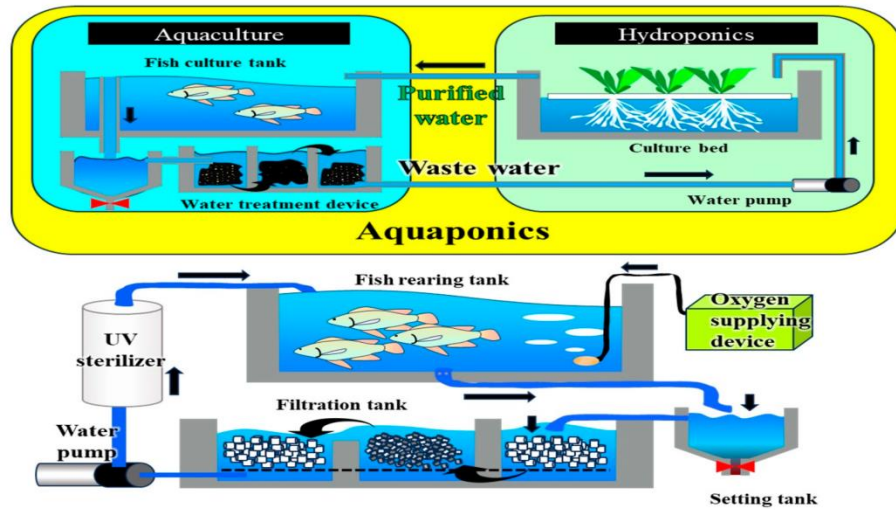


Figure 9: Aquaponics Technologies. [12]

1.6.7 Technology Integration: Automated systems now support feeding, monitoring, and environmental control. The use of sensors, real-time data analytics, and remote management tools allows farmers to make evidence-based decisions and optimize operations in real-time.

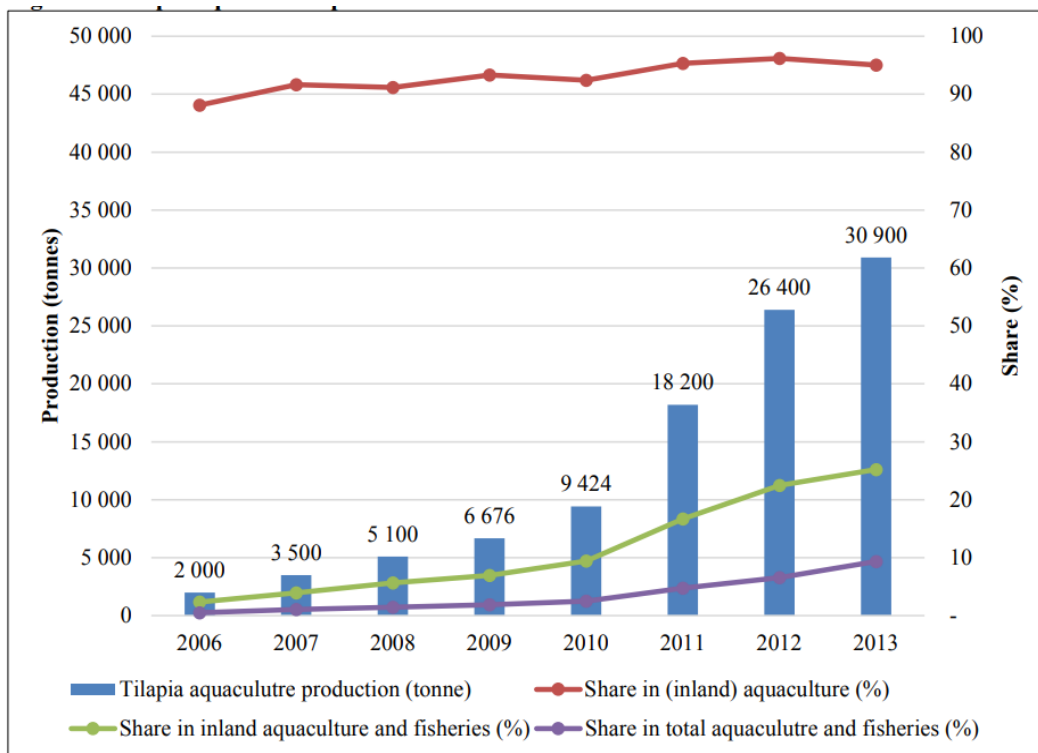


Figure 10: Tilapia Production Analytics. [13]

1.6.8 Efficient Harvesting and Processing: The harvesting process is designed to reduce fish stress and preserve quality. Equipment such as grading systems and live transport units ensure humane handling, while modern processing and storage methods maintain freshness until the product reaches market.



Figure 11: Harvesting Operations. [14]

1.6.9 Regulatory Compliance: Responsible aquaculture today aligns with government regulations concerning licensing, environmental safety, and animal welfare. Farmers are expected to meet sustainability benchmarks and adhere to industry guidelines.

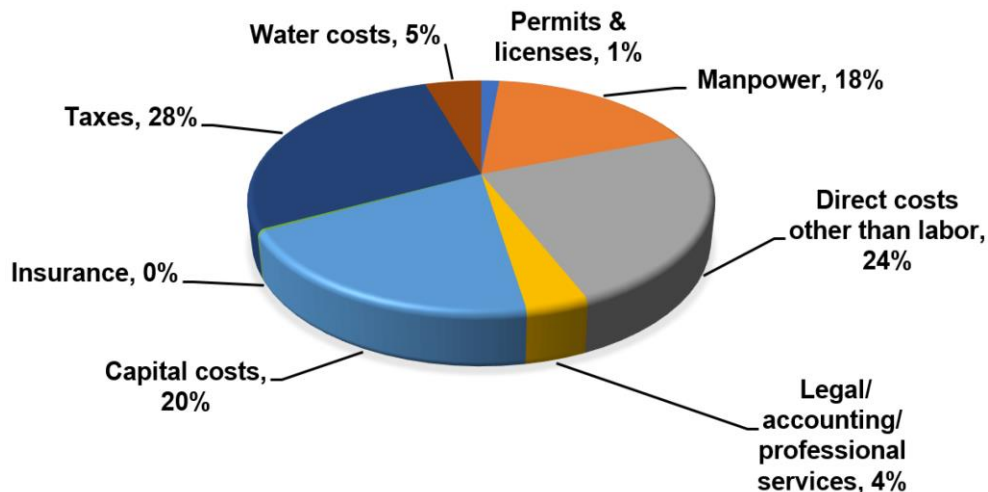


Figure 12: Regulatory Costs by Type of Cost. [15]

1.6.10 Continuous Learning and Innovation: Ongoing education and the adoption of emerging technologies are critical to staying competitive in the sector. Farmers regularly update their skills through training, workshops, and applied research to improve their farming practices and respond to evolving challenges.



Figure 13: Technology and Innovation in Aquaculture. [16]

While modern systems offer higher yields and export potential, they often require significant investment, skilled labor, and robust infrastructure. However, with government support and increasing market demand, these systems are rapidly gaining ground across India and globally.

1.7 Problem Statement

Aquaculture remains largely dependent on manual practices for monitoring and managing environmental conditions such as pH, temperature, and total dissolved solids levels. These in-person methods are time-consuming, labor-intensive, and often reactive rather than preventive. As a result, farmers face frequent challenges including delayed responses to critical changes in water parameters, increased operational costs, and reduced productivity. Despite global advancements in aquaculture technology, the local sector has yet to adopt intelligent systems that leverage real-time data, automation, or remote control. There is a clear need for a low-cost, adaptable, and user-friendly

monitoring system that can introduce digital transformation to aquaculture. Such a system would reduce human dependency, improve decision-making, and enhance sustainability in local fish farming operations.

1.8 Scope and Limitations

The scope of this research focuses on freshwater aquaculture environments, particularly small- to medium-scale fish farms, aiming to introduce smart aquaculture tools for monitoring key parameters including temperature, pH, and total dissolved solids. While the system is designed for reliability and modularity, limitations include the dependency on sensor calibration, network connectivity, and adaptability to environmental changes.

1.9 Research Methodology

1.9.1 System Methodology

The proposed intelligent aquaculture monitoring system is structured into three primary components: input, processing, and output. The input stage consists of three key sensors such as temperature sensor, pH sensor, and total dissolved solids sensor, they are responsible for collecting real-time water quality data from the aquaculture pond. These data streams are transmitted to the processing unit, which is implemented using a Raspberry Pi microprocessor. The RPi handles data acquisition, preprocessing, and decision-making processes based on predefined thresholds or trained models. Finally, in the output stage, the processed information is transmitted through an IoT platform, enabling remote monitoring and data visualization by the user in the future. This system architecture ensures continuous environmental surveillance, early anomaly detection, and informed decision-making, thereby contributing to sustainable aquaculture practices.

1.9.2 Methodology Diagram

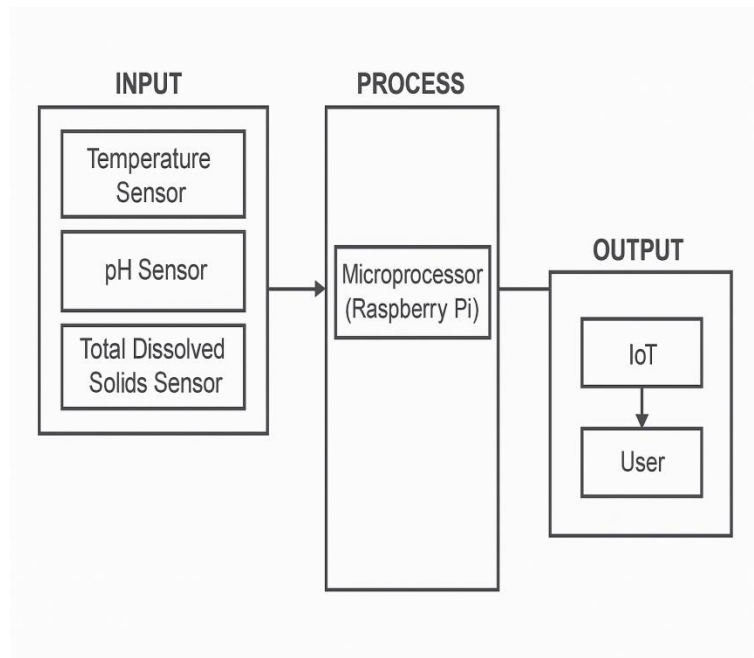


Figure 14: Block Diagram.

1.10 Conclusion

This chapter has presented the historical development and current state of aquaculture, highlighting the shift from traditional to modern practices. It has also outlined the challenges faced by aquaculture systems today and the role of emerging technologies in addressing them. The next chapter reviews the state-of-the-art systems that have attempted to bridge this gap through IoT-based monitoring and AI-enhanced control strategies.

Chapter 2: Literature Review

2.1 Introduction

In recent years, the aquaculture industry has witnessed a paradigm shift toward technology-driven sustainability, driven by the urgent need to improve productivity while reducing environmental impact. Key advancements in this transformation include the adoption of Internet of Things (IoT) frameworks, machine learning (ML) algorithms, and automated environmental control systems. These technologies offer aquaculture farmers the tools to monitor critical water quality parameters, detect anomalies in real-time, and implement data-driven decision-making strategies to improve stock health and resource efficiency. This chapter delivers an in-depth evaluation of recent scholarly and technological developments related to intelligent aquaculture solutions. The focus is placed on IoT-based sensor platforms, AI-enabled decision support systems, and integrated aquaponics designs. By analyzing a diverse set of experimental implementations from wireless sensor networks to cloud-integrated dashboards and predictive ML models, this review aims to identify current strengths, expose technological gaps, and provide a solid foundation for the intelligent monitoring solution proposed in this thesis.

2.2 Existing IoT-Based Intelligent Aquaculture Systems

2.2.1 IoT-Based Environmental Control System for Fish Farms

Dhinakaran et al. [17] made one of the most creative contributions to the field of smart aquaculture where they suggested an IoT-based environmental control system combined with machine learning decision support mechanisms for fish farms. Their system integrates wireless sensor network to monitor vital environmental parameters such as water temperature, pH levels, humidity, and fish behavior in real-time. The data are processed using methods like data imputation, outlier detection, and feature engineering, leading to reliable datasets for predictive modeling.

At the heart of their system are four machine learning algorithms, each responsible for specific aspects of aquaculture management, which are; Random Forests keep an eye on the water parameters and optimize them to maintain an optimal water quality, Support Vector Machines (SVMs) early detection of diseases and parasites in fish, Gradient Boosting Machines (GBMs) dynamically adjust feeding schedules to minimize waste and finally Neural Networks govern the operation of equipment such as water pumps and heating systems automatically.

This multi-model ML integration grants real-time, automated decision-making that leads to an improved fish health, decreases resource use, and enhances overall system sustainability. Moreover, the system offers a real-time monitoring interface for manual control and historical data analysis, empowering farmers with both automation and flexibility.

2.2.2 Smart Aquaponics with AI and IoT Integration

Gayam et al. [18] presented a comprehensive framework for smart aquaponics that integrates IoT-enabled environmental sensing with AI-driven decision support systems (DSS) to optimize yield and sustainability in aquaculture. Their proposed system architecture uses LoRa-based wireless sensor networks to monitor critical parameters such as pH, water temperature, total dissolved solids, light intensity, humidity, and water level. Two sensor nodes collect aquatic and environmental data, which are wirelessly transmitted to a centralized IoT gateway. This gateway forwards data to the cloud for storage and analysis, while also enabling real-time display and control.

The AI component leverages a deep neural network model trained on historical data exported from the IoT platform. The dataset is preprocessed and visualized to identify meaningful correlations (e.g., between air temperature and water temperature, or pH and optimal conditions). A custom-defined “Optimal Level” metric serves as the target variable to assess system health and productivity. The trained AI model, implemented using Keras, achieves high prediction accuracy and is capable of guiding real-time interventions such as automated pump activation or feeding schedules.

Importantly, the paper emphasizes hardware-level integration, including custom PCB design, secure LoRa communication, and on-board actuation through relays.

This edge-driven approach reduces reliance on continuous internet access and enhances system reliability. Compared to conventional aquaponics systems, this design improves resource efficiency, remote manageability, and ecological sustainability, showcasing how AI and IoT can work together to transform food production in controlled ecosystems.

2.2.3 IoT-Based Automated Smart Water Quality Control System for Tilapia Culture

Amrita's et al. study [19] developed a low-cost IoT-based automated system for real-time monitoring and control of water temperature and dissolved oxygen in tilapia culture tanks, aimed at reducing fish mortality and improving growth outcomes. The system integrates LM35 temperature sensors and DFRobot analog dissolved oxygen sensors, interfaced with an Arduino UNO microcontroller and NodeMCU, to gather and transmit environmental data to the Blynk mobile application in real time. When sensor readings deviate from optimal thresholds, corresponding actuators such as a heater, fan, or air pump are triggered automatically using relay circuits, ensuring environmental conditions remain within safe bounds.

The prototype was deployed in a controlled experimental setup over 21 days, where results demonstrated improved growth rates in automated tanks (0.2334g/day) compared to traditional, manually monitored systems (0.1810g/day). Importantly, the study also found a strong inverse correlation (-0.9918) between temperature and dissolved oxygen levels, reinforcing the importance of dynamic control systems in aquaculture. Through automation, the system achieved consistent parameter regulation and significantly reduced human intervention, showcasing the feasibility and reliability of smart aquaculture using embedded IoT systems.

This work provides a robust, field-tested model for real-time water quality control and is a valuable contribution toward more resilient and efficient aquaculture infrastructure, particularly for developing countries where resource constraints and labor limitations challenge manual operations.

2.2.4 IoT-Based Intensive Aquaculture Monitoring with Automatic Water

Correction

Tolentino et al. [20] proposed an advanced IoT-enabled aquaculture monitoring system designed to address the limitations of manual water quality management in intensive fish farming environments. The system monitors six critical water parameters: temperature, pH, oxidation-reduction potential (ORP), turbidity, salinity, and dissolved oxygen. Using a multi-sensor architecture connected to an Arduino Mega and Raspberry Pi 3B+, the system performs real-time data acquisition and automatic correction based on pre-programmed thresholds. Data transmission is handled via LoRaWAN, ensuring low-power, long-range communication ideal for aquaculture environments.

Corrective mechanisms include an aquarium heater, water pumps for flushing, and sodium bicarbonate dispensing motors to adjust pH levels. The entire system is visualized and controlled via a custom web application, which also tracks fish growth parameters like average weight and length. Their experimental setup compared the performance of the automated system against a conventional control tank. The automated system resulted in a 46.88% improvement in growth rate, with better environmental consistency and reduced fish stress. The study demonstrated that automatic water correction significantly boosts fish survival and yield quality by maintaining optimal water conditions with minimal human intervention.

By addressing a broader set of water parameters and automating both monitoring and correction, this work makes a major contribution to the design of intelligent aquaculture systems that are scalable, cost-effective, and environmentally adaptable. It also sets a precedent for future systems incorporating data analytics, predictive control, and integrated fish health monitoring.

2.2.5 Aquaculture Monitoring and Control System: An IoT-Based Approach

Preetham et al. [21] proposed a modular, IoT-driven architecture for remote aquaculture monitoring and control, addressing critical challenges in manual water quality management and environmental instability in fish farming. Their system, known

as the IoT-Based Aquaculture Monitoring and Control System (IAMCS), utilizes a Raspberry Pi 3B as the central controller and incorporates commercial-grade sensors to monitor pH, temperature, turbidity, and salinity in aquaculture ponds. The system is integrated with ThingSpeak cloud services, enabling real-time data transmission, visualization, and alerting.

The IAMCS architecture comprises four core modules: power (with DC-DC converters and battery backup), sensor, controller, and output (mobile app). The mobile interface provides real-time alerts and enables farmers to remotely control water flow and pumps in response to threshold breaches. A key innovation of this system is its compatibility with aquaponics; where wastewater from aquaculture is cycled through hydroponic grow beds for pH and ammonia neutralization before being recirculated, effectively reducing water consumption by up to 90%.

Experimental results over a 24-hour period validated the system's effectiveness, where deviations in turbidity, pH, or temperature automatically triggered alerts and corrective actions. Furthermore, the authors quantified system responsiveness using Round-Trip Time (RTT) analysis under various QoS (Quality of Service) configurations, demonstrating high communication efficiency with minimal latency. The use of Python for backend data processing, Android for the mobile app, and SPI/I2C for sensor interfacing illustrates the system's lightweight, low-cost design suitable for small-scale and rural aquaculture.

This study significantly contributes to accessible smart aquaculture by offering a scalable, modular, and energy-efficient IoT platform that can enhance water quality control, reduce human intervention, and support sustainability in integrated fish and plant farming systems.

2.2.6 IoT-Based Aquaculture Monitoring and Control System

Rosaline and Sathyalakshimi [22] proposed a multi-sensor, IoT-based aquaculture monitoring and control system designed to reduce manual labor and improve environmental control in shrimp and fish farming operations. The system features six water quality sensors for temperature, ammonia, salinity, dissolved oxygen, water level, and distance connected to an Arduino and NodeMCU ESP8266 platform for

real-time monitoring. Sensor data are collected every 5 seconds and transmitted wirelessly to a cloud-based Ubidots IoT platform, where it is visualized on dashboards, stored, and used to trigger automated alerts via SMS, voice call, or email.

The design includes a multiplexed analog input interface (74HC4067) to overcome the NodeMCU's single analog pin limitation, enabling the connection of multiple analog sensors. The authors emphasize the system's scalability, modularity, and affordability, noting its potential to significantly reduce the operational overhead of small to medium-sized aquafarms. The solution allows continuous remote observation and provides early warnings when any parameter exceeds predefined safety thresholds such as high ammonia or low dissolved oxygen which can severely impact aquatic health.

This platform bridges the gap between manual pond testing (e.g., handheld meters and lab sampling) and fully automated water quality regulation. It demonstrates how affordable sensor networks and cloud integration can increase aquaculture sustainability while minimizing input costs, improving growth rates, and preventing major environmental failures. The paper also outlines plans for future improvements, including complete underwater sensor submersion and additional environmental variable tracking before harvest.

2.3 Comparative Analysis of Existing IoT-Based Aquaculture Monitoring Systems

Study	Software Used	Hardware Used	Methods Used	Key Results
[17]	IoT sensor network, actuators	ML algorithms (RF, SVM, GBM, NN)	Predictive analytics for water quality & fish health	Improved fish health & sustainability
[18]	IoT sensors, AI-based DSS	Decision support system (AI)	Smart aquaponics for water & nutrient optimization	Increased crop & fish yield, reduced fertilizers
[19]	Arduino, temperature & DO sensors	Blynk mobile app	Automated temperature & DO control	Reduced fish mortality, improved efficiency
[20]	Arduino, Raspberry Pi, sensors	Web app, LoRaWAN	Automated water correction & quality monitoring	Prevented fish kills, improved efficiency
[21]	Raspberry Pi, sensors	ThingSpeak, Android app	IoT-based real-time monitoring & aquaponics	Water conservation, real-time alerts
[22]	Arduino, NodeMCU, sensors	Ubidots cloud storage	IoT-based low-cost monitoring	Improved control, reduced labor costs

Table 1: Comparison of Existing Systems.

2.4 Conclusion

This reviewed literature highlights significant progress in the development of intelligent aquaculture monitoring systems, particularly through the integration of IoT sensors, cloud platforms, and AI-driven analytics. Collectively, these technologies have enabled real-time data acquisition, remote monitoring, and in some cases, automated environmental control. Current systems demonstrate notable improvements in water quality management, energy efficiency, and labor reduction. Despite these advances, challenges remain in terms of long-term system reliability, adaptability across different aquaculture environments, and the seamless integration of sensing, processing, and actuation. Many systems still operate in isolated silos or rely on limited datasets for decision-making, which constrains their ability to scale or generalize. Furthermore, the lack of standardized architectures and robust validation under diverse real-world conditions limits broader adoption. These observations reinforce the need for a holistic, modular, and intelligent framework that is not only responsive but also predictive, capable of adapting dynamically to changing aquatic circumstances while maintaining sustainability and operational efficiency.

Chapter 3: Data Acquisition

3.1 Introduction

This chapter presents the core of the project's current stage: the design and setup of a real-time data acquisition system for aquaculture monitoring. The system is intended to replace traditional, manual methods of environmental observation with a digital platform capable of collecting and transmitting data continuously. We introduce the concept of the Internet of Things (IoT) in aquaculture, define the parameters being monitored, and explain the role of each hardware component used, including the Raspberry Pi and the selected sensors. This chapter also outlines how the components are connected, how the data is transmitted, and how it can be accessed. Emphasis is placed on modularity, cost-effectiveness, and ease of deployment, making the system suitable for small- to medium-scale fish farms.

3.2 Internet of Things Definition

The Internet of Things (IoT) refers to a network of physical devices, vehicles, appliances, and other physical objects that are embedded with sensors, software, and network connectivity, allowing them to collect and share data. [23]

3.3 Hardware and Software Tools

3.3.1 Raspberry Pi

The Raspberry Pi is a compact, credit card-sized single-board computer developed by the Raspberry Pi Foundation in the United Kingdom. Despite its small size, it offers impressive computing power and has become a popular platform for a wide range of innovative projects. With models priced between 650 da and 4550 da, the Raspberry Pi is an affordable solution for performing general computing tasks and interfacing with a variety of devices through its General Purpose Input/output (GPIO) pins.

The board is built around a Broadcom-based ARM processor and integrates essential components such as a graphics chip, RAM, and multiple connectors for external peripherals. Operating a Raspberry Pi closely resembles using a traditional personal computer; it requires a keyboard, mouse, display, power supply, and a microSD card preloaded with an operating system (serving as its storage).

In addition to standard USB ports and an Ethernet connection for networking and internet access, the Raspberry Pi supports a wide range of open-source, Linux-based operating systems. To date, over 30 different Linux distributions customized for the Raspberry Pi have been released. The Foundation has also introduced various accessories, including a camera module, the Gertboard, and the Compute Module Kit, enabling users to expand the board's capabilities for more advanced hardware development. [24]

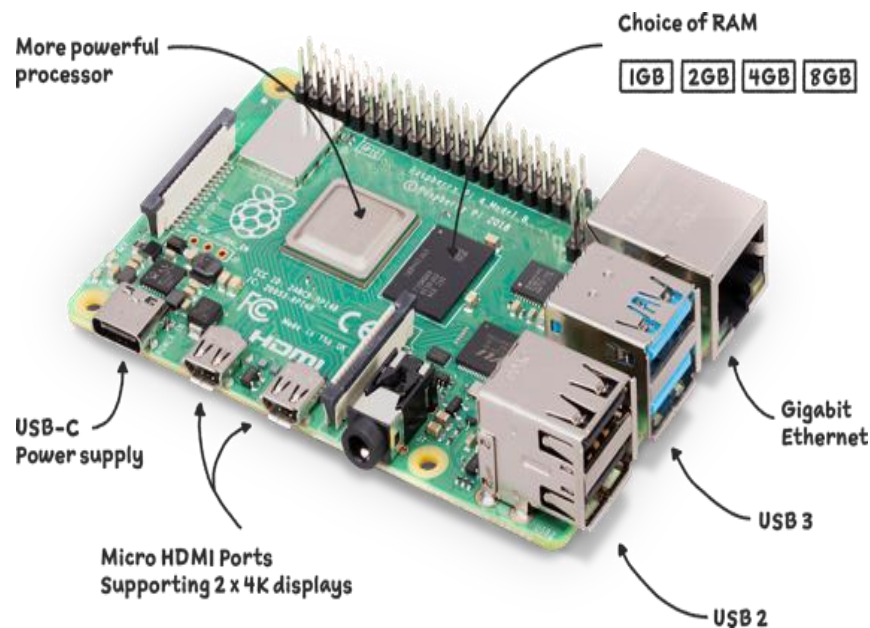


Figure 15: Raspberry pi 4 Model B.

3.3.2 IoT Sensor

An IoT sensor is a device that can detect and measure changes in the physical world, such as temperature, motion, light, sound, air quality, and others. It can

communicate with other devices or systems through wired or wireless networks, including LoRaWAN, and send data to the network server for analysis or action.

IoT sensors are used for various applications, such as smart homes, health monitoring, retail, manufacturing, transportation, environmental management, and more. [25]

3.3.3 Temperature Sensor: DS18B20

The DS18B20 digital thermometer provides 9-bit to 12-bit Celsius temperature measurements and has an alarm function with nonvolatile user-programmable upper and lower trigger points. The DS18B20 communicates over a 1-Wire bus that by definition requires only one data line (and ground) for communication with a central microprocessor. In addition, the DS18B20 can derive power directly from the data line (“parasite power”), eliminating the need for an external power supply.

Each DS18B20 has a unique 64-bit serial code, which allows multiple DS18B20s to function on the same 1-Wire bus. Thus, it is simple to use one microprocessor to control many DS18B20s distributed over a large area.

Applications that can benefit from this feature include HVAC environmental controls, temperature monitoring systems inside buildings, equipment, or machinery, and process monitoring and control systems. [26]



Figure 16: Temperature Sensor: DS18B20.

3.3.4 pH Sensor: PH-4502C

The PH4502C pH Sensor Module, developed by DIYMORE, is an electronic component used to measure the pH level of a liquid; indicating whether it is acidic or alkaline. It is commonly applied in areas such as environmental monitoring, aquaculture systems, hydroponic setups, and laboratory experiments. The sensor is typically integrated with microcontrollers like the Arduino UNO to enable real-time pH data acquisition and analysis. [27]



Figure 17: pH Sensor: ph-4502c.

3.3.5 TDS Meter V1.0: KS0429

The KS0429 TDS Meter V1.0 by Keyestudio is a sensor module designed to measure the Total Dissolved Solids (TDS) present in water, which reflects the concentration of dissolved substances such as salts, minerals, and organic matter. It provides a quantitative indication of water quality and is particularly useful in applications such as aquaculture, hydroponics, and environmental monitoring. This module is compatible with Arduino microcontrollers and outputs an analog signal that can be processed for real-time monitoring and analysis. [28]



Figure 18: TDS Meter V1.0.

3.3.6 Breadboard

A breadboard (protoboard), serves as the base platform for constructing and prototyping electronic circuits. It is widely used for assembling temporary circuits or conducting experimental designs without soldering. The breadboard is composed of a grid of perforated holes aligned in rows and columns, beneath which are internally connected spring clips. These clips securely hold wires and components in place, enabling straightforward and flexible connections. The grid's holes are uniformly spaced at 0.1 inches (2.54 mm) in both horizontal and vertical directions. [29]

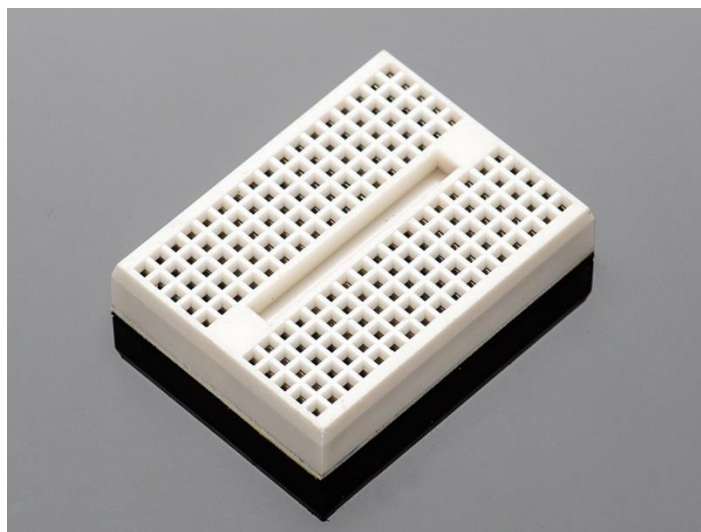


Figure 19: Breadboard.

3.3.7 Resistors

A resistor is a passive electronic component that opposes the flow of electric current by introducing resistance into a circuit. It is a fundamental element found in nearly all electrical and electronic systems. The unit of resistance is the ohm (Ω), defined as the resistance that produces a voltage drop of one volt when one ampere of current flows through it. This relationship is governed by Ohm's Law, expressed as: $R = \frac{V}{I}$; where R is resistance, V is voltage, and I is current.

Resistors serve various functions including limiting current, dividing voltage, dissipating energy as heat, matching and loading circuits, controlling gain, setting time constants. Their resistance values can span over nine orders of magnitude, making them suitable for diverse applications; from compact resistors in microelectronic circuits to large power resistors used as dynamic brakes in railway systems. [30]



Figure 20: Resistor.

3.3.8 Jumperwires

A jumper wire is a short electrical conductor used to establish connections between components on a breadboard or other prototyping platform without soldering. These wires are essential for building and testing circuits quickly and efficiently. Jumper wires typically come with male or female connector ends (or both), allowing them to connect directly into the headers or sockets of microcontrollers, sensors, and other electronic modules. They are widely used in prototyping environments due to their

reusability, flexibility, and compatibility with standard breadboards and development boards. [31]

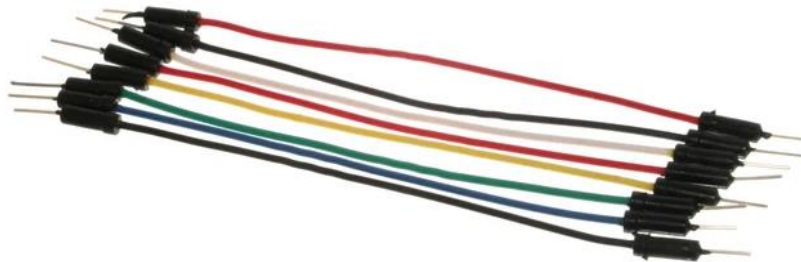


Figure 21: Jumper Wires.

3.3.9 ADC Converter

3.3.9.1 ADS1015 Definition

The ADS1015 is a 12-bit precision analog-to-digital converter (ADC) with an I²C interface, designed for applications requiring accurate analog signal measurement. This module features four single-ended input channels or two differential channels, allowing it to measure multiple analog voltages. It includes a programmable gain amplifier (PGA), enabling it to handle signals with a wide range of input voltages. The ADS1015 is ideal for interfacing analog sensors with microcontrollers such as the Arduino or Raspberry Pi, and is well-suited for use in environmental monitoring, instrumentation, and other data acquisition tasks. [32]

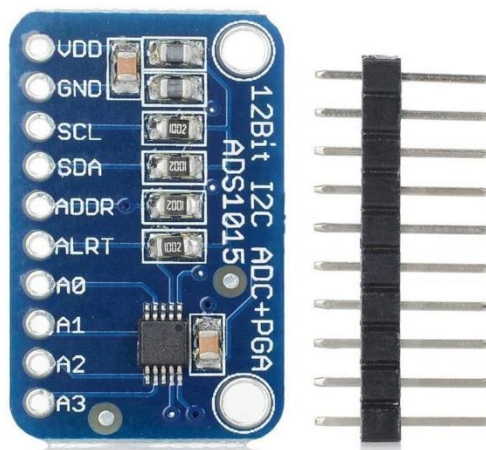


Figure 22: ADS1015 Converter.

3.3.9.2 ADS1015 Pin Configuration

The ADS1015 features a series of pins that enable its integration with microcontroller-based systems. It includes a VDD pin for power supply input and a GND pin as the ground reference. Communication is handled through the SCL (serial clock) and SDA (serial data) lines, which operate over the I²C protocol. The ADDR pin allows users to configure the I²C address, making it possible to connect multiple ADS1015 devices on the same bus. The ALERT/RDY pin serves dual purposes: it can function as an interrupt output for a comparator or signal the end of a conversion process. The device also provides four analog input pins AIN0 through AIN3 which can be used as either single-ended inputs or paired as differential input channels. This versatile pin layout makes the ADS1015 a robust choice for high-resolution, multi-channel analog signal acquisition in embedded and IoT applications. [33]

3.3.10 Putty Emulator

Putty is a free, open-source terminal emulator and network communication tool primarily used on Windows platforms. It enables users to establish remote connections with other computers or devices using various protocols, including Secure Shell (SSH), Telnet, and more. Widely adopted by system administrators, network engineers, and developers, Putty serves as a reliable utility for accessing and managing remote systems efficiently.

One of Putty's key strengths lies in its lightweight nature and ease of installation, making it accessible for users across various levels of expertise. It supports secure encrypted connections, helping maintain confidentiality during remote sessions. Moreover, Putty offers a high degree of customization, allowing users to adjust its interface and configuration to fit specific preferences or needs. As an open-source project, it benefits from ongoing enhancements and community-driven development, ensuring it remains current and robust. [34]



Figure 23: PuTTY Emulator Logo. [35]

3.3.11 Visual Code Studio

Visual Studio Code (commonly referred to as VS Code) is a free, open-source source-code editor developed by Microsoft. Available across major operating systems including Windows, macOS, and Linux, it provides a lightweight yet powerful platform for software development. Despite its minimal footprint, VS Code offers advanced functionality that has made it one of the most widely adopted development tools in recent years.

VS Code supports a broad range of programming languages such as Java, C++, Python, CSS, Go, and Dockerfile. Its rich extension ecosystem allows users to enhance its functionality by integrating linters, debuggers, version control tools, and support for cloud and web development workflows. The editor's customizability and performance efficiency make it a preferred choice for developers working on diverse projects across multiple environments. [36]

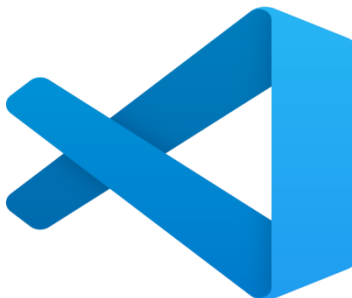


Figure 24: VS Code Logo. [37]

3.4 The Process

3.4.1 Raspberry Pi Setup

This section documents the process of setting up the Raspberry Pi, including installing the operating system and preparing the environment for data acquisition and sensor communication. [38]

3.4.1.1 Hardware Used

Raspberry Pi 4 Model B (or your model)

16/32 GB MicroSD card

Power supply (5V, 3A)

HDMI cable and monitor (for initial setup)

USB keyboard and mouse

MicroSD card reader (for flashing OS)

3.4.1.2 Installing Raspberry Pi OS

The operating system used in this project is Raspberry Pi OS (Lite or Full version depending on need). The following steps outline the process:

Step 1: Downloading Raspberry Pi Imager

The official Raspberry Pi Imager tool was downloaded from the Raspberry Pi Foundation website: <https://www.raspberrypi.com/software/>

Step 2: Preparing the SD Card

Insert the microSD card into a card reader connected to a PC.

Launch Raspberry Pi Imager and choose:

OS: Raspberry Pi OS (32-bit) or Raspberry Pi OS Lite for minimal installs.

Storage: Select the microSD card.

Optional: Click the settings icon to preconfigure:

Enable SSH

Set hostname

Configure Wi-Fi credentials (SSID and password)

Set locale and keyboard layout

Click Write and wait for the process to complete.

3.4.2 Remote Development Setup

Remote access to the Raspberry Pi was configured using PuTTY and Visual Studio Code (VS Code) with the Remote SSH extension. This setup allowed writing, editing, and executing Python scripts directly on the Raspberry Pi from a host machine. A structured project folder was created, containing sensor modules, the main execution script, and a CSV log file for recorded data.

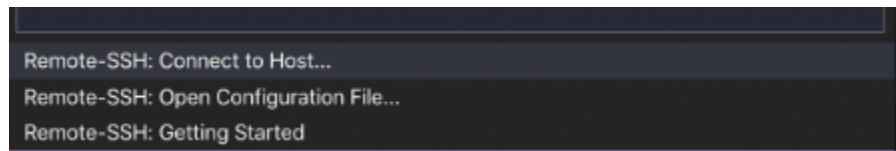


Figure 25: SSH Command to use IP Address of RPi.

3.4.3 Sensor Integration and Wiring

This section describes the physical integration of all sensors used in the aquaculture monitoring system. Four sensors were connected to a Raspberry Pi 4 Model B, with analog sensors interfaced via an ADS1015 analog-to-digital converter. Each sensor was assigned a specific function and connected through appropriate GPIO or I2C channels. [39] [40] [41]

3.4.3.1 DS18B20 – Water Temperature Sensor (Digital, 1-Wire)

- Type: Digital temperature sensor (waterproof version)
- Signal Protocol: 1-Wire
- Connections:

VCC to 3.3V pin on the Raspberry Pi

GND to any GND pin

DATA to GPIO4 (physical pin 7)

Note: A 4.7kΩ pull-up resistor is typically connected between the data and VCC.

3.4.3.2 pH-4502C – pH Sensor (Analog via ADS1015)

Type: Analog pH sensor module

Signal Output: Voltage signal (0–5V range typically scaled)

Connected via: ADS1015 analog-to-digital converter

Wiring:

pH Sensor VCC to 5V on Raspberry Pi

pH Sensor GND to GND

Analog output (AOUT) to ADS1015 A0

ADS1015 I2C wiring:

VDD to 3.3V

GND to GND

SCL to GPIO3 (physical pin 5)

SDA to GPIO2 (physical pin 3)

3.4.3.3 TDS Meter V1.0 – Conductivity Sensor (Analog via ADS1015)

Type: Analog total dissolved solids sensor

Signal Output: Voltage signal proportional to conductivity

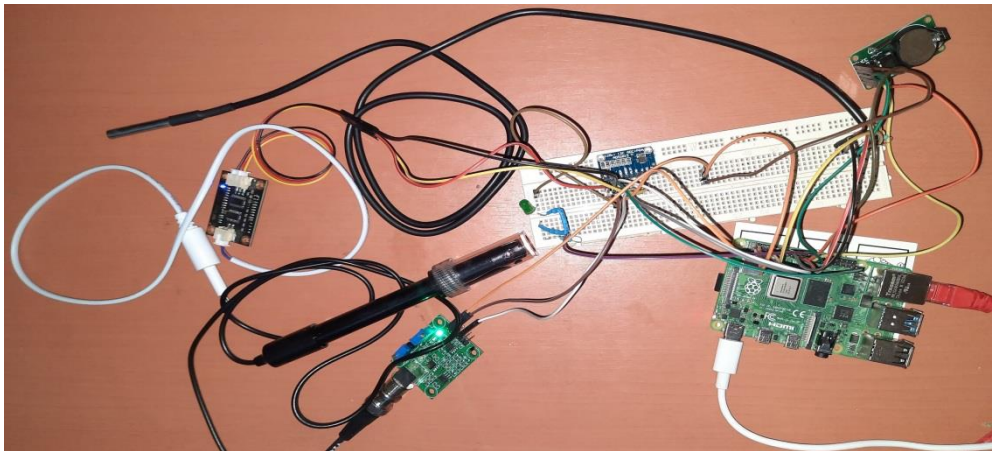
Connected via: ADS1015 (same as pH sensor)

Wiring:

TDS Sensor VCC to 5V

GND to GND

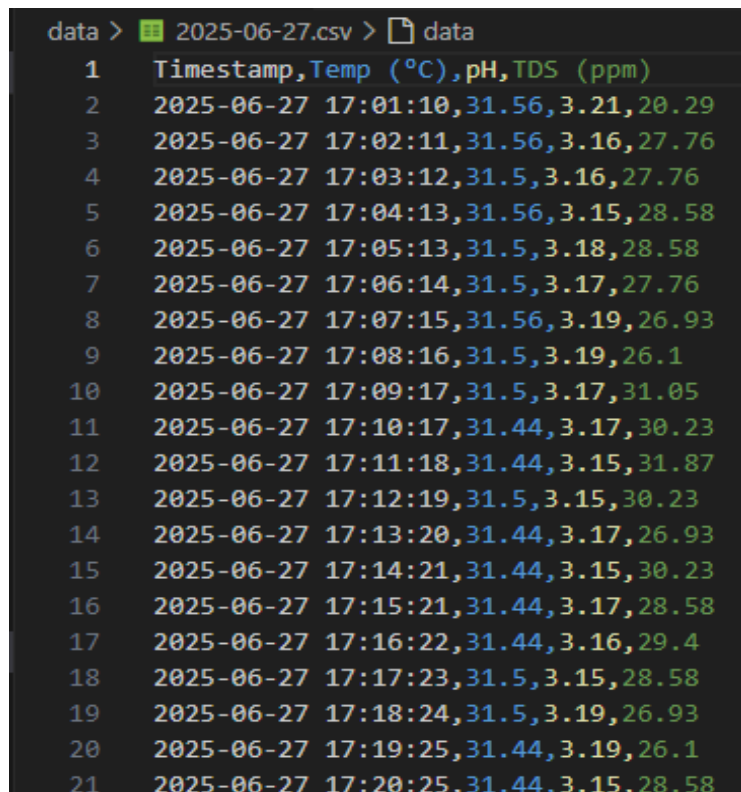
Analog output (AOUT) to ADS1015 A1



Figures 26: Aquaculture Intelligent System.

3.4.4 Development and Data Logging

Individual Python modules were written for each sensor, and a central script (main.py) was created to read data sequentially and save it in CSV format. Required libraries such as adafruit-circuitpython-ads1x15 and adafruit_dht were installed to interface with the sensors. The system was tested to verify real-time data logging and sensor response.



1	Timestamp,Temp (°C),pH,TDS (ppm)
2	2025-06-27 17:01:10,31.56,3.21,20.29
3	2025-06-27 17:02:11,31.56,3.16,27.76
4	2025-06-27 17:03:12,31.5,3.16,27.76
5	2025-06-27 17:04:13,31.56,3.15,28.58
6	2025-06-27 17:05:13,31.5,3.18,28.58
7	2025-06-27 17:06:14,31.5,3.17,27.76
8	2025-06-27 17:07:15,31.56,3.19,26.93
9	2025-06-27 17:08:16,31.5,3.19,26.1
10	2025-06-27 17:09:17,31.5,3.17,31.05
11	2025-06-27 17:10:17,31.44,3.17,30.23
12	2025-06-27 17:11:18,31.44,3.15,31.87
13	2025-06-27 17:12:19,31.5,3.15,30.23
14	2025-06-27 17:13:20,31.44,3.17,26.93
15	2025-06-27 17:14:21,31.44,3.15,30.23
16	2025-06-27 17:15:21,31.44,3.17,28.58
17	2025-06-27 17:16:22,31.44,3.16,29.4
18	2025-06-27 17:17:23,31.5,3.15,28.58
19	2025-06-27 17:18:24,31.5,3.19,26.93
20	2025-06-27 17:19:25,31.44,3.19,26.1
21	2025-06-27 17:20:25,31.44,3.15,28.58

Figures 27: Example of a Dataset from our Intelligent System.

3.5 Conclusion

This chapter detailed the design and implementation of the data acquisition system developed for aquaculture monitoring. By combining low-cost sensors, a Raspberry Pi microcontroller, and simple data logging scripts, the system successfully enables real-time collection and storage of key environmental parameters, including temperature, pH, and total dissolved solids. The process covered the setup of hardware and software components, sensor integration, wiring, and initial testing in a controlled water environment. The choices made in this design focused on affordability, scalability, and adaptability, providing a practical foundation for small- to medium-scale fish farms seeking to transition from manual monitoring to a digital, IoT-based approach. This data acquisition platform also lays the groundwork for future enhancements involving intelligent analysis and automated decision-making.

Chapter 4 : Model Implementation

4.1 Introduction

This chapter outlines the theoretical and technological basis for the future expansion of the monitoring system toward intelligent aquaculture management. It also includes an overview of machine learning (ML) and deep learning (DL) as potential tools for future data analysis, anomaly detection, and decision support. This theoretical foundation establishes a clear path for developing smarter, more autonomous aquaculture systems that respond dynamically to real-time environmental data.

4.2 Artificial Intelligence Definition

Artificial Intelligence (AI) refers to a scientific discipline focused on developing computer systems and machines capable of performing tasks that typically require human intelligence. These tasks include reasoning, learning from data, and making decisions, often applied to datasets of such scale and complexity that human analysis alone would be insufficient.

AI is inherently multidisciplinary, drawing on fields such as computer science, statistics, data analytics, hardware and software engineering, linguistics, neuroscience, psychology, and philosophy. In business contexts, AI commonly refers to a suite of technologies, mainly machine learning and deep learning that enable advanced applications such as predictive analytics, natural language processing, intelligent recommendations, data classification, and automated decision-making. [42]

4.3 Artificial Intelligence for Internet of Things

Artificial Intelligence of Things (AIoT) refers to the integration of artificial intelligence (AI) capabilities into Internet of Things (IoT) ecosystems. This convergence aims to optimize the performance of IoT systems, enhance interactions between humans and machines, and enable more advanced and efficient data processing and analytics across connected devices. [43]

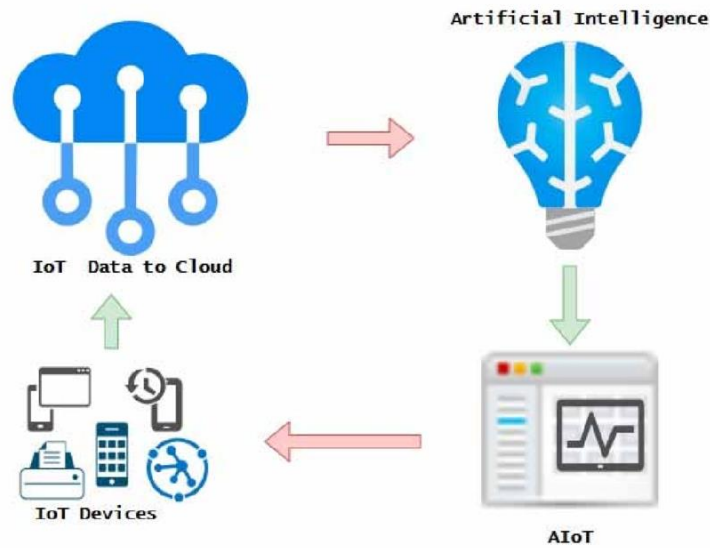


Figure 28: Conceptual integration of IoT with AI for intelligent systems. [44]

4.4 Machine Learning

4.4.1 Definition

Machine learning (ML) is a subfield of artificial intelligence that empowers systems to autonomously learn from historical data and experiences without being explicitly programmed. By employing algorithms that iteratively process and analyze data, ML systems can identify patterns, make predictions, and continuously adapt to new information. Unlike traditional programming, which relies on predefined rules, machine learning enables computers to improve their performance over time based on the data they encounter. [45]

HOW DOES MACHINE LEARNING WORK?

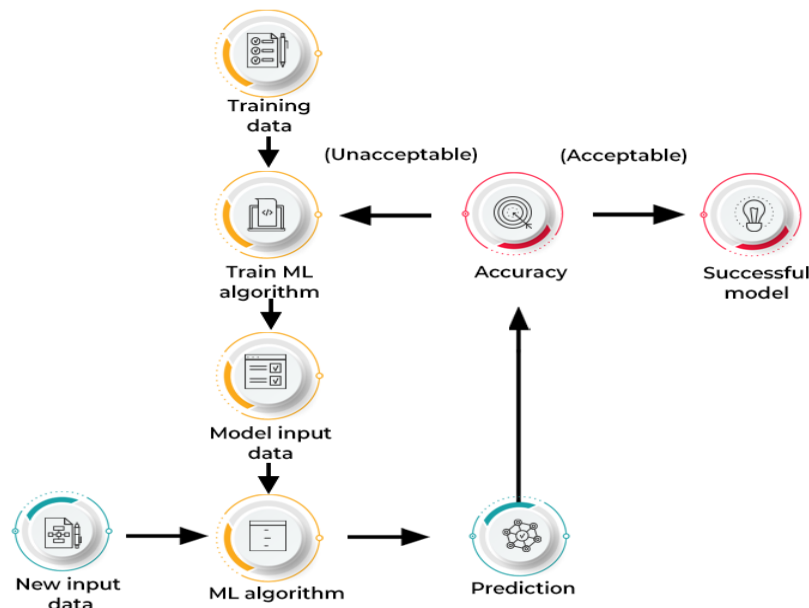


Figure 29: Machine Learning Process. [45]

4.4.2 ML Types

Machine learning (ML) is a subset of artificial intelligence that enables systems to learn from data and improve their performance over time without being explicitly programmed. It can be broadly categorized into four main types: supervised learning, unsupervised learning, semi-supervised learning, and reinforcement learning. [45]

4.4.2.1 Supervised Learning

Supervised learning involves training a model on a labeled dataset, where both input and output variables are known. The model learns to predict outcomes based on input data by recognizing patterns. For example, if a system is trained with labeled images of parrots and crows, it can later identify new images by comparing features such as shape and color. This technique is commonly used in applications like spam filtering and email categorization.

4.4.2.1.1 Classification: Predicts categorical outcomes (e.g., true/false, spam/not spam). Common algorithms include Random Forest, Decision Tree, Logistic Regression, and Support Vector Machine.

4.4.2.1.2 **Regression:** Predicts continuous outcomes (e.g., stock prices, temperature). Examples include Simple Linear Regression and Lasso Regression.

4.4.2.2 Unsupervised Learning

In unsupervised learning, the model is trained on an unlabeled dataset and aims to uncover hidden patterns or groupings within the data. For example, when analyzing images of various fruits, the model might group them based on color or shape without prior labeling.

4.4.2.2.1 Clustering: Groups similar items together (e.g., customer segmentation). Algorithms include K-Means, DBSCAN, and Principal Component Analysis (PCA).

4.4.2.2.2 Association: Discovers relationships between variables in large datasets (e.g., market basket analysis). Common algorithms are Apriori and FP-Growth.

4.4.2.3 Semi-Supervised Learning

Semi-supervised learning combines a small amount of labeled data with a large amount of unlabeled data during training. This hybrid approach helps improve learning accuracy while reducing the manual effort required to label data. For instance, a student who learns in class (labeled data) and reviews at home (unlabeled data) is an analogy of this method.

4.4.2.4 Reinforcement Learning

Reinforcement learning (RL) is a feedback-driven technique where an agent learns by interacting with an environment. The agent receives rewards for desirable actions and penalties for poor choices, refining its decisions to maximize cumulative reward. A practical example is an AI playing a video game and improving its performance over time based on scores.

RL methods include:

Positive Reinforcement: Strengthens behavior by rewarding desirable outcomes.

Negative Reinforcement: Encourages behaviors that avoid negative results.

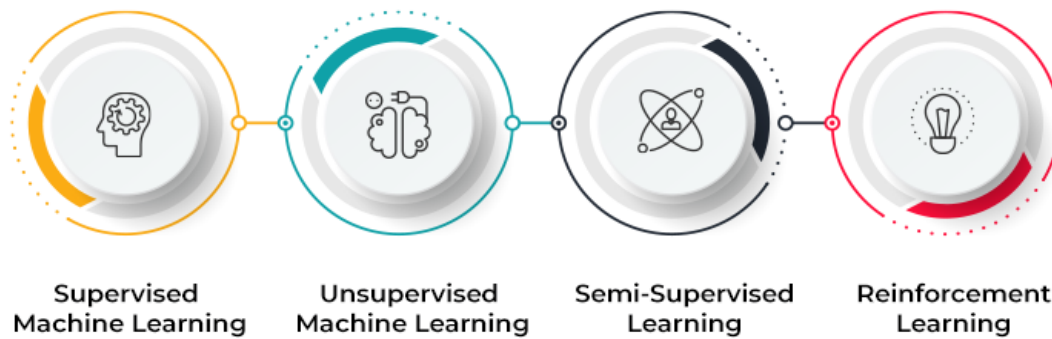


Figure 30: Types of Machine Learning. [45]

4.4.3 Machine Learning Models

A machine learning model is a trained computational entity designed to recognize specific patterns within data. During the training phase, the model learns from a dataset using a particular algorithm, enabling it to make predictions or decisions based on new, unseen data. For example, a model trained on images labeled with various emotions can later identify emotions in new facial images. These models are integral to applications such as image recognition, recommendation systems, and real-time analytics. [46]

4.5 Deep Learning

4.5.1 Definition

Deep learning (DL) is revolutionizing how machines interpret, process, and derive insights from complex, high-volume datasets. Inspired by the structure and function of the human brain, deep learning uses artificial neural networks (ANNs) to autonomously extract patterns and perform intelligent tasks. These networks are composed of multiple layers of interconnected neurons, where each neuron performs non-linear computations to enable hierarchical data representation. In a typical deep neural network, input data flows through an input layer, several hidden layers, and an output layer, which together generate the final prediction. [47]

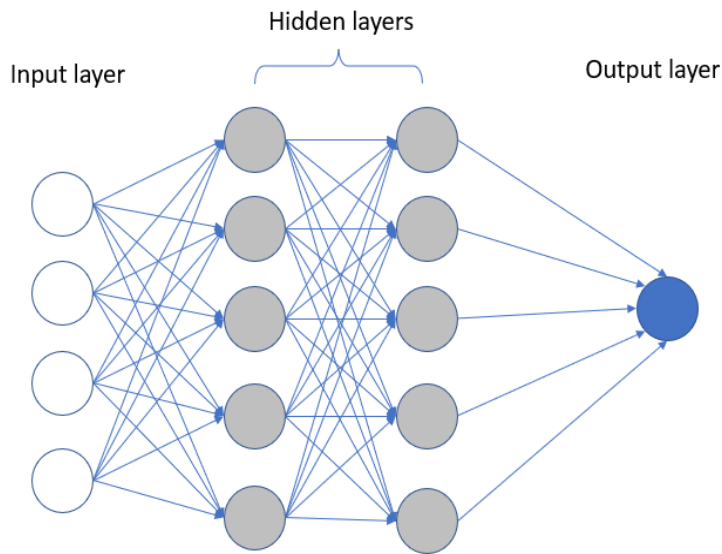


Figure 31: Deep Learning's Layers. [48]

4.5.2 DL in ML Paradigms

Deep learning spans multiple machine learning paradigms [47], enhancing their performance and applicability:

4.5.2.1 Supervised Learning: Deep learning models like Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs) are trained on labeled data to perform classification, regression, or sequence prediction tasks such as image recognition and language translation.

4.5.2.2 Unsupervised Learning: In the absence of labeled data, deep learning models like autoencoders and generative networks identify hidden patterns, group similar data points, and perform tasks such as clustering and anomaly detection.

4.5.2.3 Reinforcement Learning: Deep learning models also support decision-making through trial-and-error by optimizing cumulative rewards. Techniques such as Deep Q-Networks (DQN) and Deep Deterministic Policy Gradient (DDPG) are commonly used in fields like autonomous robotics and strategic game playing.

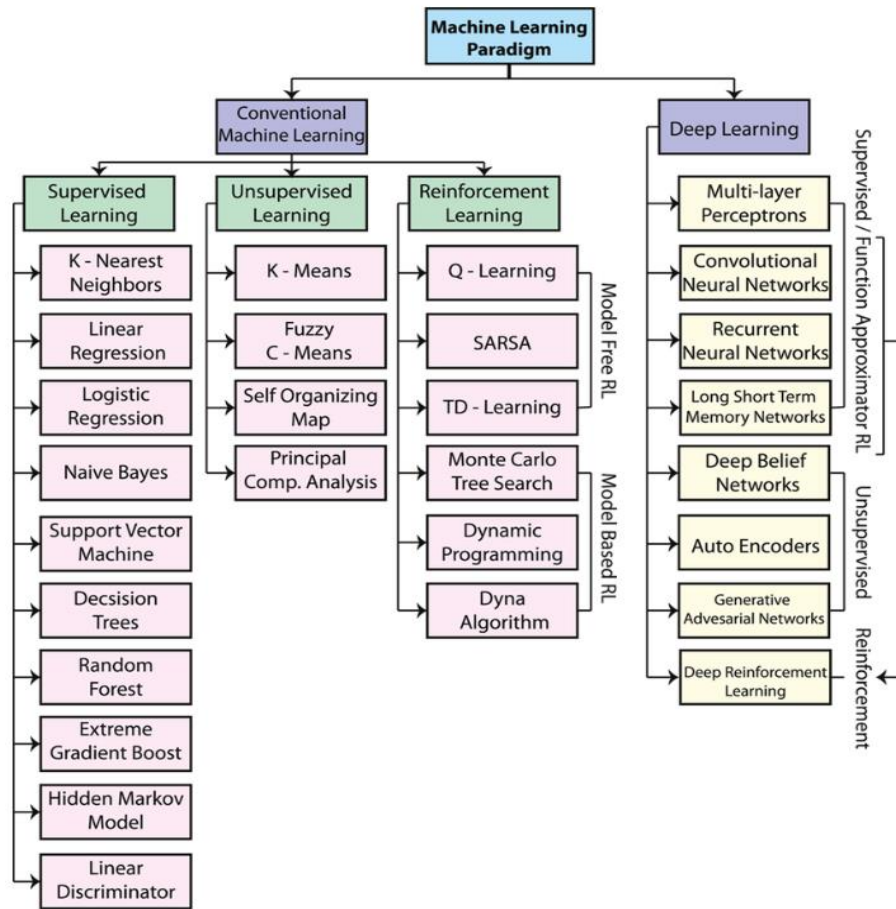


Figure 32: Machine Learning Paradigms. [49]

4.5.3 Comparison Between Machine Learning and Deep Learning

Aspect	Machine Learning	Deep Learning
Approach to Learning	Utilizes statistical algorithms to identify patterns and relationships within the dataset.	Employs artificial neural networks to automatically learn complex data representations.
Data Requirements	Performs effectively with smaller datasets.	Requires large volumes of data to achieve high accuracy.
Task Suitability	Well-suited for simpler tasks with limited labeled data.	Ideal for complex tasks such as image recognition, speech processing, and natural language understanding.
Training Time	Typically faster to train due to simpler models.	Training time is longer due to the depth and complexity of neural networks.
Feature Extraction	Manual extraction of relevant features is necessary, especially for tasks like object detection in images.	Automatically extracts features during the learning process, offering an end-to-end solution.
Model Complexity and Interpretability	Models are generally less complex and easier to interpret.	Models tend to act as "black boxes" making interpretation more difficult.
Computational Resources	Can run efficiently on CPUs with modest hardware requirements	Demands high-performance computing resources, often requiring GPUs for training.

Table 2: ML and DL comparison.

4.6 Chosen Dataset

4.6.1 Justification for Dataset Selection

Due to time constraints and the unavailability of certain hardware components during the implementation phase, it was not feasible to collect sufficient real-time data using the designed IoT monitoring system. As a result, a reliable and publicly available dataset was selected to support the application and evaluation of AI models in this study.

4.6.2 Dataset Used

The dataset used in this study is sourced from an IoT-based monitoring system for freshwater tilapia (*Oreochromis niloticus*) aquaculture. It is publicly available on Mendeley Data [50] and was designed to support research in aquaculture water quality monitoring, intelligent management strategies, and predictive modeling.

The data was collected using IoT sensors strategically placed in aquaculture ponds. Readings were taken every hour over a six-month period. This dataset provides valuable insights into the relationship between water quality parameters and the health and growth of tilapia in controlled aquaculture conditions. Missing data due to sensor or communication failures were addressed using interpolation methods, and regular sensor calibrations ensured the accuracy of the collected data.

The dataset integrates environmental measurements, intervention records, and fish health indicators. The main columns include:

Datetime: The timestamp of each sensor reading.

Month: The month during which data was collected (January to June).

Temperature (°C): Water temperature readings.

Dissolved Oxygen (mg/L): Levels of dissolved oxygen in the water.

pH: Water pH values.

Turbidity (NTU): Measure of water clarity (Nephelometric Turbidity Units).

Average Fish Weight (g): The average weight of tilapia in the pond at each timestamp.

Survival Rate (%): The percentage of fish survival during the monitoring period.

Disease Occurrence (Cases): The number of disease cases observed at each time point.

Oxygenation Automatic: Indicates if automatic oxygenation was applied (Yes/No).

Oxygenation Interventions: Whether oxygenation interventions were applied (Yes/No).

Corrective Interventions: The number of corrective measures taken.

Thermal Risk Index: Categorical indicator of thermal risk ("High" or "Normal").

Low Oxygen Alert: Status indicator showing "Critical" if dissolved oxygen levels fell below 5 mg/L, otherwise "Safe."

Health Status: Fish health status derived from thermal and oxygenation risk indicators, categorized as "Stable" or "At Risk."

The target variables for machine learning tasks in this study were:

Survival Rate (%): Continuous variable representing the percentage of fish survival.

Average Fish Weight (g): Continuous variable representing fish growth performance.

Health Status: Categorical variable representing fish health, categorized as Stable or At Risk.

4.7 ML and DL Algorithms Choices

4.7.1 Random Forest Regressor

The Random Forest Regressor is an ensemble learning algorithm that builds multiple decision trees during training and predicts output by averaging the results of individual trees, improving accuracy and controlling overfitting. [51]

Justification: This model was applied to predict continuous outcomes such as survival rate (%) and average fish weight (g). It was chosen for its ability to handle non-linear relationships, reduce overfitting through ensemble averaging, and provide high predictive performance even with limited data preprocessing.

Features Used: Temperature (°C), pH, Dissolved Oxygen (mg/L), Turbidity (NTU)

4.7.2 Random Forest Classifier

The Random Forest Classifier is an ensemble machine learning algorithm that creates multiple decision trees during training and predicts class labels based on a majority vote from the trees, enhancing classification accuracy and reducing overfitting. [52]

Justification: This model was used to classify pond health status into Stable or At Risk. It was selected for its robustness to noise, strong classification accuracy, and ability to identify the most important features contributing to predictions.

Features Used: Temperature (°C), pH, Dissolved Oxygen (mg/L), Turbidity (NTU)

4.7.3 XGBoost

XGBoost is a powerful and efficient gradient boosting algorithm designed for speed and performance. It builds an ensemble of decision trees sequentially, where each new tree aims to correct the errors of the previous ones, making it highly effective for structured data tasks. [53]

Justification: XGBoost was used as an alternative classifier for pond health status. It was chosen for its high accuracy, speed, built-in regularization, and native handling of missing values. Its performance was compared against the Random Forest and SVM classifiers.

Features Used: Temperature (°C), pH, Dissolved Oxygen (mg/L), Turbidity (NTU)

4.7.4 Support Vector Machine (SVM)

Support Vector Machine is a supervised learning algorithm that finds the optimal hyperplane to separate data points of different classes, maximizing the margin between them for better classification or regression performance. [54]

Justification: The SVM classifier was used to provide a baseline comparison for health status classification. It is effective in high-dimensional spaces and suitable for smaller datasets.

Features Used: Temperature (°C), pH, Dissolved Oxygen (mg/L), Turbidity (NTU)

4.7.5 Isolation Forest

Isolation Forest is an unsupervised machine learning algorithm used for anomaly detection. It works by randomly selecting features and splitting values to isolate data points, with anomalies requiring fewer splits to be separated from the rest. [55]

Justification: This model was used to detect abnormal water quality conditions that could pose risks to fish health. It was selected for its efficiency and strong performance on multidimensional sensor data without requiring labeled samples.

Features Used: Standardized values of Temperature, pH, Dissolved Oxygen, and Turbidity (StandardScaler)

4.7.6 Autoregressive Integrated Moving Average (ARIMA)

ARIMA, which stands for Autoregressive Integrated Moving Average, is a statistical model used for analyzing and forecasting time series data. It combines

autoregression, differencing to make data stationary, and moving averages to capture patterns over time. [56]

Justification: ARIMA was applied to univariate time series forecasting, such as predicting future values of temperature. It was selected as a baseline time series model due to its interpretability and strong performance on short-term forecasts with limited features.

Features Used: Historical time series of individual parameters (e.g., temperature)

4.7.7 Long Short Term Memory (LSTM)

LSTM is a type of recurrent neural network (RNN) architecture designed to learn and retain information over long sequences, making it highly effective for tasks involving time series, speech, and natural language data. [57]

Justification: LSTM was employed for multivariate time series forecasting, such as predicting future Dissolved Oxygen levels. It was selected for its ability to model long-term temporal dependencies and complex nonlinear relationships in sequential data.

Features Used: Sequences of scaled values for Temperature, pH, Dissolved Oxygen, and Turbidity (MinMaxScaler)

4.8 Preprocessing

4.8.1 Python

Python is a high-level, versatile programming language widely known for its clear syntax and readability. It supports multiple programming paradigms, including object-oriented, procedural, and functional programming. Python's extensive standard libraries and large ecosystem of third-party modules make it ideal for a wide range of applications, such as web development, automation, data analysis, artificial intelligence, and scientific computing. Its simplicity and flexibility have contributed to its widespread use among beginners and professionals alike. [58]

4.8.2 Packages Used

NumPy: Numerical operations, array manipulation, and data structuring for sequence modeling.

Pandas: Loading datasets, handling missing values, and performing tabular data manipulation.

Scikit-learn (sklearn): Data preprocessing (e.g., scaling and encoding), model training, classification, regression, evaluation metrics, class balancing, and anomaly detection.

XGBoost: Gradient boosting classifier with support for class imbalance handling (scale_pos_weight) and robust structured data learning.

TensorFlow/Keras: Building and training the LSTM model for multivariate time series forecasting.

Matplotlib: Visualizing regression and classification results, forecasts, and feature importances.

Seaborn: Enhanced statistical plots including confusion matrices and feature importance heatmaps.

Statsmodels: Time series forecasting using ARIMA and statistical modeling for univariate predictions.

```
# Load and Preprocess Data

import pandas as pd
import numpy as np
from sklearn.model_selection import train_test_split
from sklearn.preprocessing import LabelEncoder, StandardScaler, MinMaxScaler

# Load dataset
data = pd.read_excel("Data_Model_IoTMLCQ_2024.xlsx")

# Drop rows with missing values
data_clean = data.dropna()

# Select input features
features = ['Temperature (°C)', 'pH', 'Dissolved Oxygen (mg/L)', 'Turbidity (NTU)']
X = data_clean[features]

# Define targets
y_survival = data_clean['Survival Rate (%)']
y_weight = data_clean['Average Fish Weight (g)']
y_health = data_clean['Health Status']

# Encode health status labels (Stable / At Risk) into integers
le = LabelEncoder()
y_health_encoded = le.fit_transform(y_health)

# Scale features
standard_scaler = StandardScaler()
X_scaled = standard_scaler.fit_transform(X) # Used for Isolation Forest

minmax_scaler = MinMaxScaler()
X_lstm_scaled = minmax_scaler.fit_transform(X) # Used for LSTM
```

Figure 33: Data Loading and Preprocessing.

4.9 Results and Plots

This section presents the results obtained from the machine learning and deep learning models applied in this study. The plots illustrate the performance of each model on the respective predictive tasks.

4.9.1 Survival Rate Prediction (Random Forest Regressor)

Figure 34 presents the comparison between the actual survival rates and the predicted values generated by the Random Forest Regressor. The x-axis represents the sample index (i.e., individual timestamps), and the y-axis shows the survival rate as a percentage. The prediction curve appears mostly flat and fails to track fluctuations in the actual survival data, which may suggest the model was unable to generalize from the input features. This limitation is likely due to the narrow variance in the survival rate variable, which remains relatively stable across the dataset. Although Random Forest is generally a robust model for regression, its performance here was limited due to the lack of strong signal in the target variable and possibly the influence of correlated features with low predictive power.

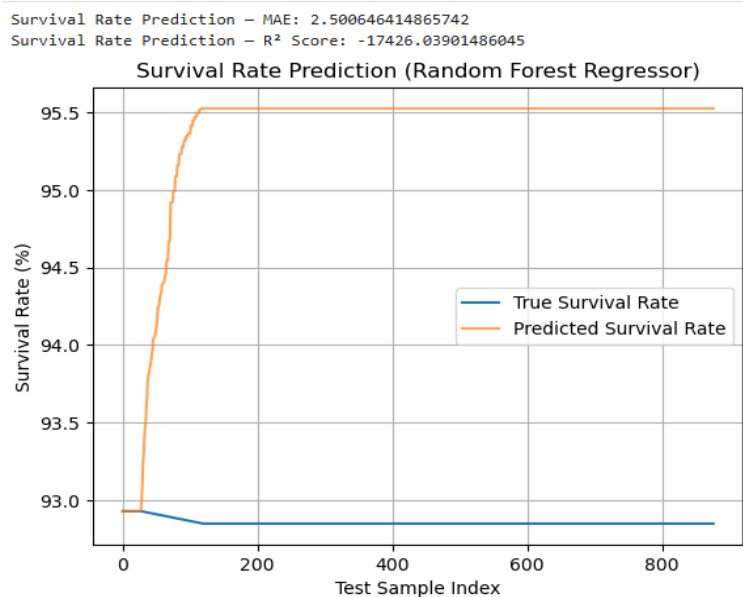


Figure 34: True vs Predicted Survival Rate using Random Forest Regressor.

4.9.2 Average Fish Weight Prediction (Random Forest Regressor)

In Figure 35, the Random Forest Regressor is applied to predict the average fish weight. The plot displays the actual fish weights and the model's predictions over time. Similar to the previous plot, the predicted values remain relatively static and do not align with the observed changes in weight. This could be attributed to weak correlations between water quality parameters and fish growth within the dataset or insufficient variability in the target output. The model may also have overfit the training data due to the small number of impactful features. Nevertheless, the attempt was useful for confirming the limitations of ensemble regressors on low-variance biological metrics when contextual variables are missing.

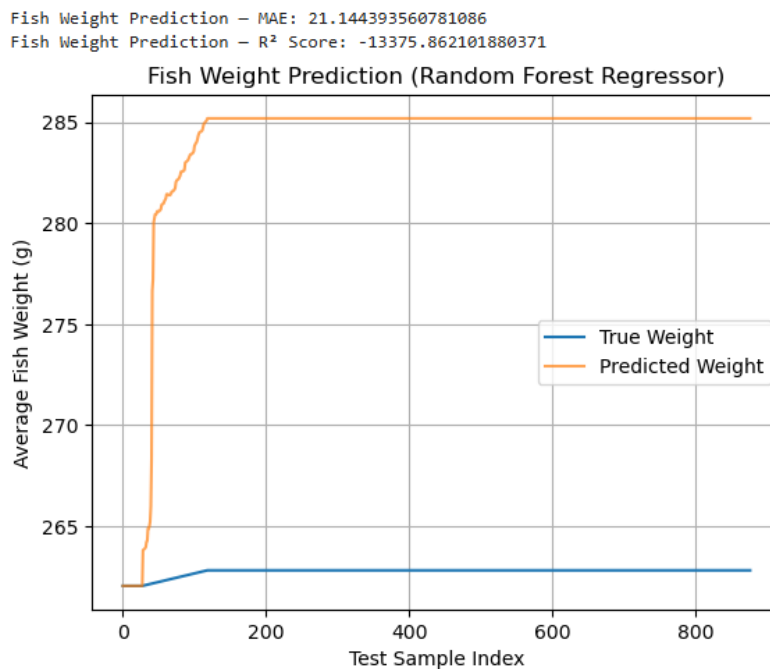


Figure 35: True vs Predicted Average Fish Weight using Random Forest Regressor.

4.9.3 Health Status Classification (Random Forest Classifier)

4.9.3.1 Confusion Matrix

Figure 36 shows the confusion matrix of the Random Forest Classifier used for predicting fish health status (Stable vs At Risk). The classifier performed well in

predicting the majority class but showed some confusion in identifying the minority “At Risk” class, resulting in a few false negatives. This could be a consequence of class imbalance, where the “Stable” condition dominates the dataset. Still, the model was able to provide a reasonable decision boundary, making it a solid baseline for binary classification tasks.

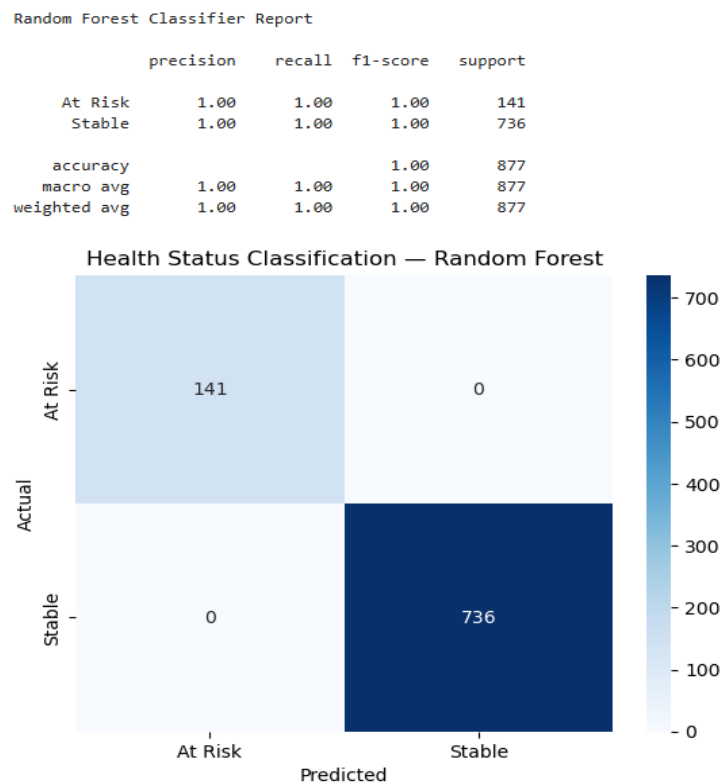


Figure 36: Confusion Matrix of Random Forest Classifier for Health Status Classification.

4.9.3.2 Feature Importance

In Figure 37, a bar chart illustrates the feature importance scores as determined by the Random Forest Classifier. Temperature was identified as the most influential variable in determining health status, followed by dissolved oxygen and pH. Turbidity had the lowest impact among the selected features. This ranking aligns with biological expectations, as temperature plays a critical role in fish metabolism and oxygen solubility, while dissolved oxygen and pH directly affect respiration and physiological balance. Turbidity, though relevant, appeared to exert less predictive influence in this

specific dataset. This visualization enhances the interpretability of the Random Forest model and highlights the most relevant water quality parameters for future monitoring and decision-making systems.

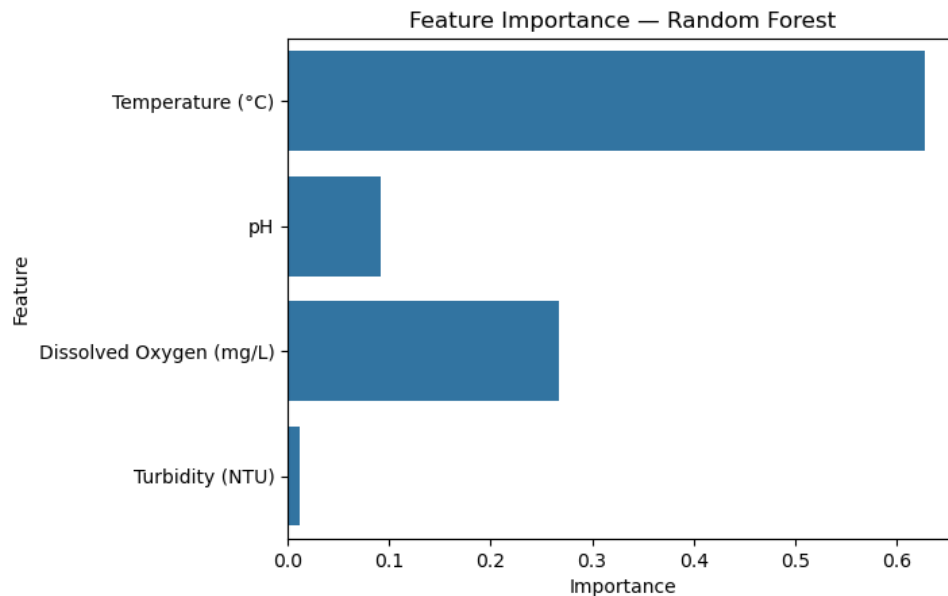


Figure 37: Feature Importance of Random Forest Classifier in Health Status Prediction.

4.9.4 Health Status Classification (XGBoost Classifier)

Figure 38 displays the confusion matrix for the XGBoost classifier applied to health status classification. Given the clear imbalance in the dataset, where “Stable” samples significantly outnumber “At Risk” samples, the `scale_pos_weight` parameter was adjusted to penalize the model for misclassifying the minority class. This adjustment aimed to improve sensitivity to the minority category.

Despite this effort, the results showed minimal improvement compared to the unbalanced version. The model maintained perfect or near-perfect precision and recall for both classes (F1-score = 1.00). This may suggest that the features used are linearly separable enough for the classifier to perform well, or it may reflect an overfitting issue due to the limited variability in the dataset’s features and labels. While balancing did not harm performance, it did not significantly enhance the model's ability to detect “At Risk” cases beyond what was already achieved by the original configuration.

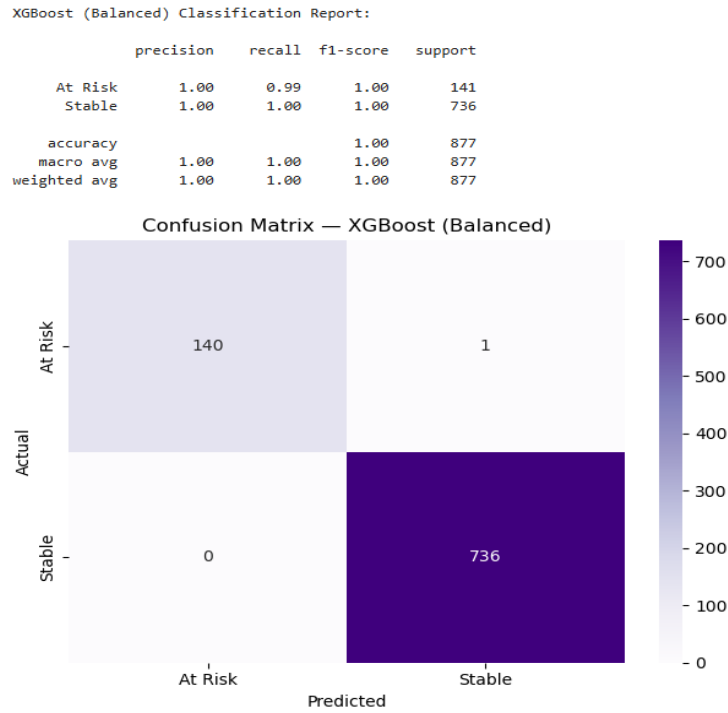


Figure 38: Confusion Matrix of XGBoost Classifier for Health Status Classification.

4.9.5 Health Status Classification (SVM Classifier)

Figure 39 shows the confusion matrix of the SVM classifier with `class_weight='balanced'`. This technique automatically adjusts weights inversely proportional to class frequencies in the input data to address imbalance.

The balanced SVM model achieved very high scores: a precision and recall of 0.99 for “At Risk” and 1.00 for “Stable,” with an overall accuracy of 100%. However, the performance increase was marginal compared to the baseline SVM. This result indicates that while class balancing helped slightly mitigate the imbalance, the SVM algorithm remained constrained by the linear separability of the input features and the underlying structure of the dataset. Thus, class balancing had limited impact on improving model generalization or robustness in this case.

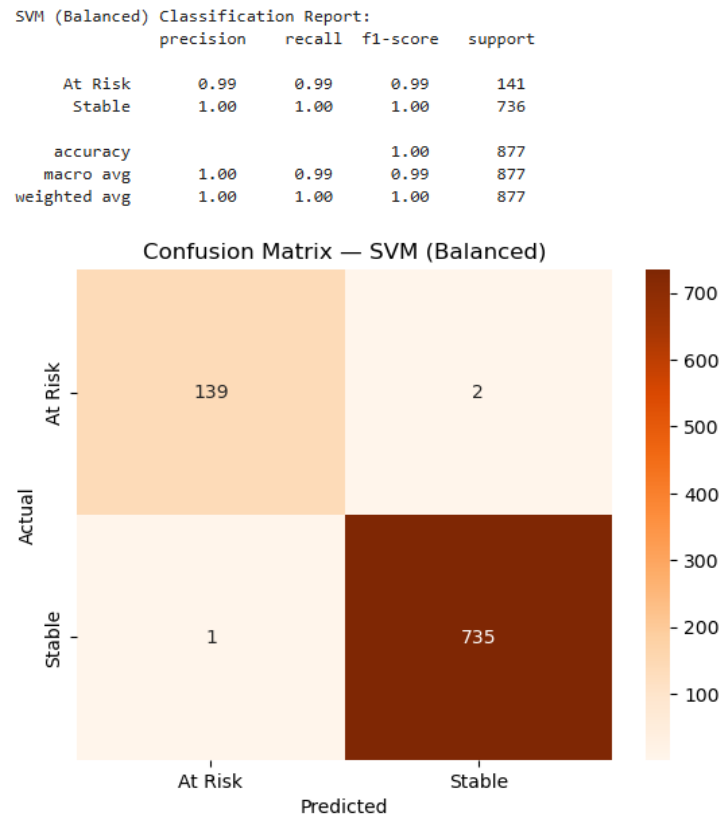


Figure 39: Confusion Matrix of SVM Classifier for Health Status Classification.

4.9.6 Anomaly Detection (Isolation Forest)

In Figure 40, a scatter plot presents the output of the Isolation Forest algorithm used to detect anomalies in the water quality data. Temperature and dissolved oxygen are plotted along the x- and y-axes respectively, while color-coding distinguishes between normal observations (blue) and anomalies (red). The model successfully flagged a number of outliers, which could correspond to sudden environmental changes or sensor errors. This demonstrates that Isolation Forest is effective for highlighting unusual measurements that may pose biological risks. However, without expert domain validation, it remains difficult to confirm whether the flagged anomalies are biologically relevant events or false positives due to natural fluctuations.

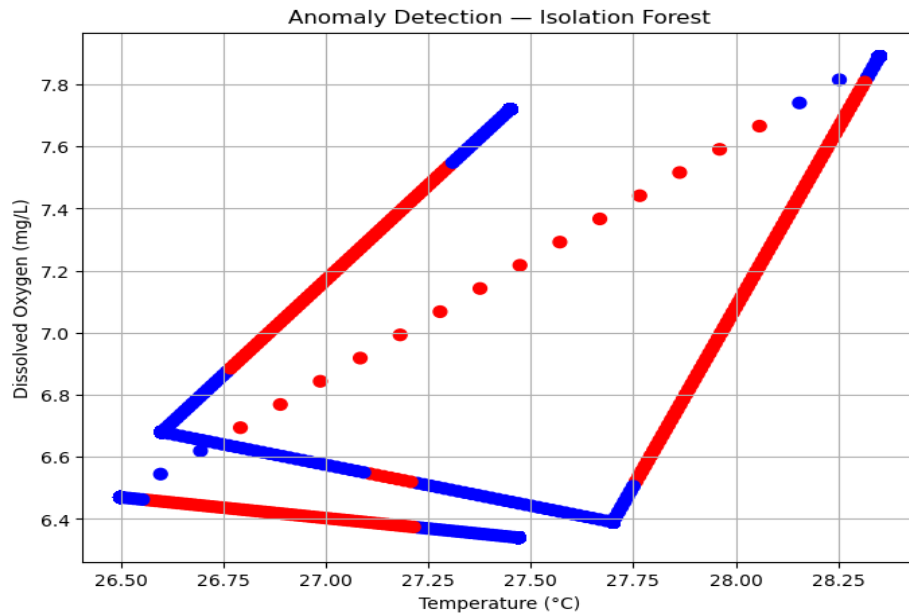


Figure 40: Anomaly Detection in Temperature and Dissolved Oxygen using Isolation Forest.

4.9.7 Temperature Forecasting (ARIMA)

Figure 41 shows the ARIMA model's forecasting performance for water temperature. The blue line represents observed values, while the red line extends from the last observed point, depicting the forecast for the next 10 time steps. The model captured the underlying trend of the series with reasonable smoothness, suggesting that ARIMA is effective for short-term forecasting of stable environmental variables like temperature. However, its forecast range remains narrow and linear, unable to accommodate sudden shifts or nonlinear patterns, which limits its adaptability in dynamic aquaculture conditions.

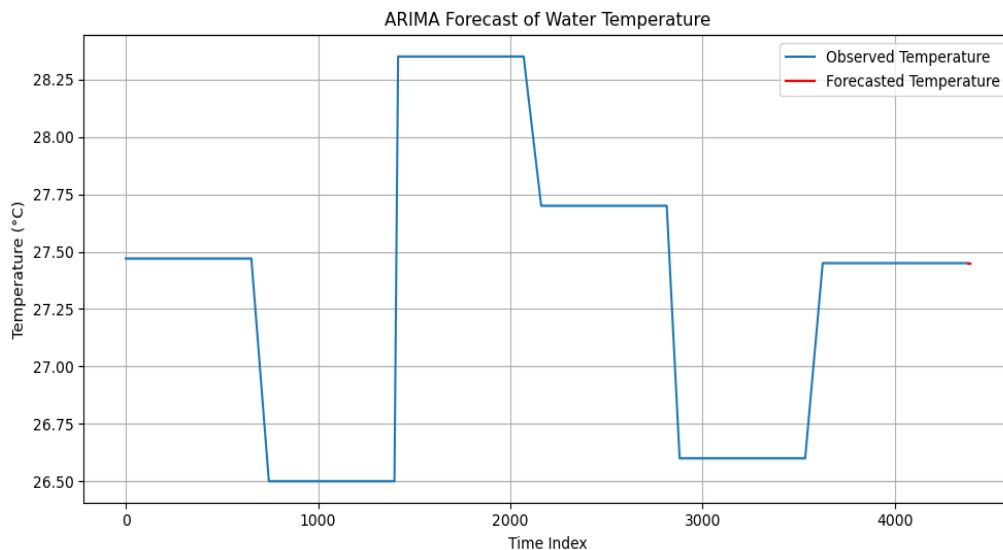


Figure 41: Observed vs Forecasted Water Temperature using ARIMA.

4.9.8 Multivariate Time Series Forecasting (LSTM)

Figure 42 presents the LSTM model's prediction for dissolved oxygen (DO) based on multivariate input features. The plot compares the true DO values (blue) with predicted values (orange) for a portion of the test set. The LSTM network demonstrated a closer fit to actual values than traditional time series models, especially in learning the trend and periodic fluctuations. Despite some deviations during sharp transitions, the LSTM model proved its capacity to capture complex temporal dependencies and multivariate relationships. This suggests that with sufficient training data and tuning, LSTM-based models can be highly effective for real-time environmental forecasting in aquaculture systems.

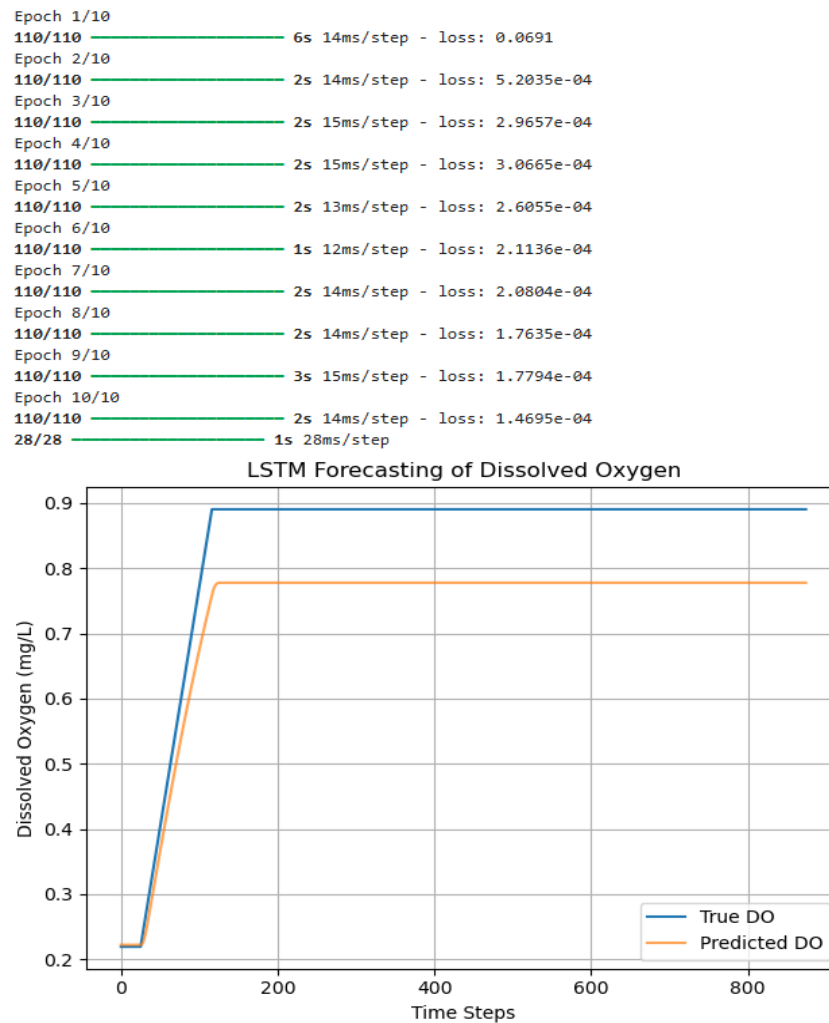


Figure 42: True vs Predicted Dissolved Oxygen using LSTM.

4.10 Summary of ML and DL Models Applied to the Tilapia Dataset

Model	Task	Parameters Used	Performance & Insights	Limitations
Random Forest Regressor	Survival Rate Prediction	Temperature, pH, DO, Turbidity	Low variance prediction curve. Model failed to reflect real survival rate variations, likely due to stable target variable.	Limited sensitivity to small fluctuations. Weak correlation with selected features.
Random Forest Regressor	Fish Weight Prediction	Temperature, pH, DO, Turbidity	Output curve was flat and unaligned with actual weights. Results suggest insufficient feature strength to model growth patterns.	Dataset lacked variables directly influencing fish weight.
Random Forest Classifier	Health Status Classification	Temperature, pH, DO, Turbidity	Performed well in predicting “Stable” class and gave interpretable feature importance.	Struggled with minority “At Risk” class due to class imbalance.
XGBoost Classifier	Health Status Classification	Temperature, pH, DO, Turbidity	Applied class balancing. Reported excellent recall and precision, but the improvement was not significantly higher than Random Forest.	Real gain over RF was marginal due to data structure and class overlap.
SVM Classifier	Health Status Classification	Temperature, pH, DO, Turbidity	SVM with class weighting improved minority class recognition slightly. Still achieved nearly identical scores to other models.	Sensitive to noisy and nonlinearly separable data.
Isolation Forest	Anomaly Detection	Scaled Temperature and DO	Identified extreme values effectively. Helpful for real-time alerting.	May misclassify normal fluctuations as anomalies.
ARIMA	Temperature Forecasting	Historical Temperature	Modeled trend well for short-range forecasting. Suitable for stable seasonal changes.	Inflexible to short-term variations. Limited to univariate forecasting.
LSTM	Dissolved Oxygen Forecasting	Scaled multivariate inputs	Closest approximation to real DO values. Captured temporal patterns better than traditional models.	Requires longer training time and more data. Slight loss of accuracy in sudden value shifts.

Table 3: Comparative Summary of Machine Learning and Deep Learning Model Performance on the Tilapia Aquaculture Dataset.

4.11 Conclusion

This chapter provided the theoretical foundation for integrating artificial intelligence into aquaculture monitoring systems. It introduced key concepts such as machine learning, deep learning, and the convergence of AI with IoT in the form of AIoT, highlighting their relevance for advancing toward intelligent, autonomous aquaculture management. The chapter demonstrated how these technologies can be applied to environmental data for tasks such as predictive analysis, anomaly detection, and decision support, contributing to more responsive and efficient aquaculture practices.

General Conclusion

This thesis presented the design and implementation of an IoT-based data acquisition system for freshwater aquaculture, combined with machine learning and deep learning models for environmental data analysis. The proposed system enables continuous monitoring of key water quality parameters using affordable hardware, providing fish farmers with timely information to support better management decisions.

Through the application of various AI models including Random Forest, XGBoost, SVM, Isolation Forest, ARIMA, and LSTM, the study demonstrated the potential of data-driven methods for predicting survival rate, fish weight, health status, and detecting anomalies. The results highlight the feasibility and value of integrating IoT with AI to support intelligent aquaculture systems.

This work lays a foundation for future developments that could include automated control and more advanced decision support, contributing to more sustainable and efficient aquaculture practices.

Perspectives and Future Enhancements

To build upon the foundations laid by this work, several future enhancements could be explored to further advance intelligent aquaculture systems:

Real-world data acquisition from the developed intelligent system: A key future step involves deploying the proposed IoT-based monitoring system in an operational aquaculture environment to collect real-time data directly from the physical setup. This self-collected dataset, tailored to the specific sensors and conditions of the implemented system, will enable the training and evaluation of machine learning models under real-world conditions, leading to more accurate, context-specific predictions and decision-making.

Integration of blockchain technology: The incorporation of blockchain could ensure secure, immutable, and transparent recording of sensor data and system interventions, enhancing trust and traceability within aquaculture operations.

AI-driven disease detection: Future work could focus on developing artificial intelligence models capable of early detection and prediction of disease outbreaks, enabling proactive management and reducing the risk of large-scale losses.

Automated feeding systems: The integration of intelligent, sensor-guided automated feeding mechanisms could reduce manual effort, optimize feed usage, and contribute to improved fish health and nutrition management.

These enhancements would contribute to the development of more resilient, efficient, and sustainable aquaculture systems, aligning with modern precision farming and smart aquaculture practices.

Bibliography

- [1] Department of Primary Industries and Regional Development, “Aquaculture” *Government of Western Australia*, [Online]. Available : <https://www.fish.wa.gov.au/Fishing-and-Aquaculture/Aquaculture/Pages/default.aspx> [Accessed: 16-Mar-2025].
- [2] “Aquaculture” *ScienceDirect*, [Online]. Available: <https://www.sciencedirect.com/topics/social-sciences/aquaculture#definition>. [Accessed: 16-Mar-2025].
- [3] Colin E. Nash, “Prepared for the Aquaculture Development and Coordination Programme.”, [Online]. Available: <https://www.fao.org/4/s2761e/s2761e01.htm> [Accessed: 16-Mar-2025].
- [4] M. Duyari, P. Sahoo, P. Payra, and S. Samanta, “Basic Comparison between Traditional Farming and Modern Farming in Aquaculture” *ResearchGate*, Sept. 2023. [Online]. Available: <https://www.researchgate.net/publication/374144426> [Accessed: 16-Mar-2025].
- [5] DHI, “The Secret to Supporting Sustainable Marine Aquaculture,” *DHI Blog*, [Online]. Available: <https://blog.dhigroup.com/the-secret-to-supporting-sustainable-marine-aquaculture/> [Accessed: 16-Mar-2025].
- [6] T. J. Alam, A. A. S. B. Hayder, A. F. Apu, M. H. A. Banna, and M. S. Rahman, “IoT Based Biofloc Aquaculture Monitoring System,” in *Proc. 4th Int. Conf. Trends Comput. Cogn. Eng.*, M. S. Kaiser, S. Waheed, A. Bandyopadhyay, M. Mahmud, and K. Ray, Eds. Singapore: Springer, 2023, vol. 618, *Lecture Notes in Networks and Systems*, pp. 85–94. [Online]. Available: https://doi.org/10.1007/978-981-19-9483-8_9 [Accessed: 02-Apr-2025]
- [7] T. Lin, S. Li, D. Zhang, X. Liu, and Y. Ren, “Effects of stocking density on the survival, growth, and stress levels of the juvenile lined seahorse (*Hippocampus erectus*) in recirculating aquaculture systems,” *Biology*, vol. 13, no. 10, p. 807, Oct. 2024. [Online]. Available: <https://www.mdpi.com/2079-7737/13/10/807> [Accessed: 10-Apr-2025]
- [8] Y. Alfiko, D. Xie, R. T. Astuti, J. Wong, and L. Wang, “Insects as a feed ingredient for fish culture: Status and trends” *Aquaculture and Fisheries*, vol. 7, no. 2, pp. 166–178, Mar. 2022. [Online]. Available:

<https://www.sciencedirect.com/science/article/pii/S2468550X21001465> [Accessed: 10-Apr-2025]

[9] R. Subramani, “Implications of quorum sensing and quorum quenching in aquaculture health management” in *Recent Developments in Applied Microbiology and Biochemistry*, N. S. Puvanakrishnan and S. R. Rajagopal, Eds., Singapore: Springer, 2019, pp. 239–254. doi: 10.1007/978-981-32-9409-7_18 [Accessed: 25-Apr-2025]

[10] Innovasea, “Recirculating Aquaculture System (RAS) Design,” Innovasea, [Online]. Available: <https://www.innovasea.com/land-based-aquaculture/ras-design/> [Accessed: 25-Apr-2025]

[11] J. Zhang, S. Zhang, D. Kitazawa, J. Zhou, S. Park, S. Gao, and Y. Shen, “Bio-mitigation based on integrated multi-trophic aquaculture in temperate coastal waters: practice, assessment, and challenges” , May 2019. [Online]. Available: <https://www.researchgate.net/publication/333002843>. [Accessed: 25-Apr-2025]

[12] Endo, M. Aquaponics in plant factory. In *Plant Factory Using Artificial Light*; Elsevier, 2019; pp. 339–352. Available online: <https://www.sciencedirect.com/science/article/abs/pii/B9780128139738000324?via%3Dihub> [Accessed: 28-Apr-2025]

[13] E. A. Frimpong and G. Anane-Taabeah, “Social and economic performance of tilapia farming in Ghana”. [Online]. Available: https://www.researchgate.net/publication/301220437_The_Social_and_Economic_Performance_of_Tilapia_Farming_in_Ghana [Accessed: 28-Apr-2025]

[14] Planex Aquaculture, “Harvesting and post-harvest handling of fish at a fish farm,” Planex Aquaculture, [Online]. Available: <https://www.planexaquaculture.com/harvesting-and-post-harvest-handling-of-fish-at-a-fish-farm/> [Accessed: 01-May-2025]

[15] C. R. Engle, J. van Senten, C. Clark, and N. Boldt, “Has the regulatory compliance burden reduced competitiveness of the U.S. tilapia industry?” *Fishes*, vol. 8, no. 3, p. 151, Mar. 2023. [Online]. Available: <https://www.mdpi.com/2410-3888/8/3/151> [Accessed: 01-May-2025]

[16] F. A. Cerda Dubó, “Technology and Innovation in Aquaculture” LinkedIn, Jul. 14, 2024. [Online]. Available: <https://www.linkedin.com/pulse/technology-innovation-aquaculture-franco-alfredo-cerda-dub%C3%B3-yeh1e/> [Accessed: 05-Apr-2025]

[17] D. Dhinakaran, S. Gopalakrishnan, M. D. Manigandan, and T. P. Anish, “IoT-Based Environmental Control System for Fish Farms with Sensor Integration and Machine Learning Decision Support”, Oct. 2023, doi: 10.17762/ijritcc.v11i10.8482. [Accessed: 14-Feb-2025]

- [18] K. K. Gayam, A. Jain, R. Singh, A. Gehlot, and S. V. Akram, "Smart Aquaponics with Integration of AI and IoT for Yield Enhancement through Real-Time Monitoring and Decision Support", Nov. 2023. doi: 10.17762/ijritcc.v11i10.8887 [Accessed: 14-Feb-2025]
- [19] C. M. Amrita, D. Babiola, S. Nivetha, and P. Sivanesan, "Development of IoT Based Automated Smart Water Quality Control System in Tilapia Culture Tank", Jan. 2023, doi: 10.1063/5.0149416. [Accessed: 14-Feb-2025]
- [20] Tolentino, Lean Karlo S., Celline P. De Pedro, Jatt D. Icamina, John Benjamin E. Navarro, Luigi James D. Salvacion, Gian Carlo D. Sobrevilla, Apolo A. Villanueva, Timothy M. Amado, Maria Victoria C. Padilla, and Gilfred Allen M. Madrigal. "Development of an IoT-based intensive aquaculture monitoring system with automatic water correction" Dec. 2021. [Accessed: 14-Feb-2025]
- [21] K. Preetham, B. C. Mallikarjun, K. Umesha, F. M Mahesh., and S. Neethan, "Aquaculture Monitoring and Control System: An IoT-Based Approach", 2019. [Accessed: 14-Feb-2025]
- [22] N. Rosaline and S. Sathyalakshimi, "IoT-Based Aquaculture Monitoring and Control System", 2019, doi: 10.1088/1742-6596/1362/1/012071. [Accessed: 14-Feb-2025]
- [23] IBM, "Internet of Things (IoT)," IBM Think, [Online]. Available: <https://www.ibm.com/think/topics/internet-of-things>. [Accessed: 20-Apr-2025]
- [24] A. Nayyar, D. Singh, and R. R. Shukla, "Raspberry Pi—A Small, Powerful, Cost Effective and Efficient Form Factor Computer: A Review" Int. J. Adv. Res. Comput. Sci. Softw. Eng., vol. 5, no. 12, pp. 720–737, Dec. 2015. [Online]. Available: https://www.researchgate.net/publication/305668622_Raspberry_Pi-A_Small_Powerful_Cost_Effective_and_Efficient_Form_Factor_Computer_A_Review [Accessed: 18-Apr-2025].
- [25] TEKTELIC Communications, "What is an IoT Sensor?", [Online]. Available: <https://tektelic.com/what-it-is/iot-sensor/>. [Accessed: 20-Apr-2025]
- [26] DS18B20: Programmable Resolution 1-Wire Digital Thermometer, Rev. 6, July 2019. [Online]. Available: <https://www.analog.com/media/en/technical-documentation/data-sheets/ds18b20.pdf>. [Accessed: 20-Apr-2025]
- [27] Cirkuit Designer, "PH4502C pH Sensor Module" Cirkuit Designer Docs, [Online]. Available: <https://docs.cirkitdesigner.com/component/607ee68e-1b94-4135-add3-ebb6e47d72a5/ph4502c>. [Accessed: 13-May-2025]

[28] Keyestudio, "KS0429 keyestudio TDS Meter V1.0" Keyestudio Wiki, [Online].

Available: https://wiki.keyestudio.com/KS0429_keyestudio_TDS_Meter_V1.0.

[Accessed: 15-May-2025]

[29] "Breadboard," ScienceDirect Topics, [Online]. Available:

<https://www.sciencedirect.com/topics/engineering/breadboard> [Accessed: 17-May-2025]

[30] "What is a Resistor?" EEP – Electrical Engineering Portal, [Online]. Available:

<https://eepower.com/resistor-guide/resistor-fundamentals/what-is-a-resistor/> [Accessed:

17-May-2025]

[31] SparkFun Education, "What is a jumper wire?" SparkFun Blog, [Online]. Available:

<https://blog.sparkfuneducation.com/what-is-jumper-wire> [Accessed: 17-May-2025]

[32] Adafruit, "ADS1015 12-Bit ADC - 4 Channel with Programmable Gain Amplifier"

Adafruit, [Online]. Available: <https://www.adafruit.com/product/1083> [Accessed: 26-May-2025]

[33] Texas Instruments, ADS1013, ADS1014, and ADS1015 12-Bit Analog-to-Digital Converters With Internal Reference, Oscillator, and Programmable Comparator,

Adafruit, 2012. [Online]. Available: <https://cdn-shop.adafruit.com/datasheets/ads1015.pdf>. [Accessed: 27-May-2025]

[34] "What is PuTTY?" Lenovo Glossary, Lenovo, [Online]. Available:

<https://www.lenovo.com/us/en/glossary/putty/> [Accessed: 10-Apr-2025]

[35] Louisiana State University, "IT Knowledge Base: How to Download and Install PuTTY" LSU Grok Knowledge Base, [Online]. Available:

<https://networking.grok.lsu.edu/Article.aspx?articleid=17237> [Accessed: 10-Apr-2025]

[36] "What is Visual Studio Code?" How.dev, [Online]. Available:

<https://how.dev/answers/what-is-visual-studio-code> [Accessed: 10-Apr-2025]

[37] Wikimedia Commons, "File:Visual Studio Code 1.35 icon.svg", Wikimedia Commons, Jun. 3, 2019. [Online]. Available:

https://commons.wikimedia.org/wiki/File:Visual_Studio_Code_1.35_icon.svg [Accessed: 10-Apr-2025]

[38] Raspberry Pi Foundation, "Getting started with your Raspberry Pi," Raspberry Pi Documentation, [Online]. Available:

<https://www.raspberrypi.com/documentation/computers/getting-started.html> [Accessed: 10-Apr-2025]

[39] Maxim Integrated, DS18B20: Programmable Resolution 1-Wire Digital Thermometer, [Online]. Available: <https://www.circuitbasics.com/wp-content/uploads/2016/03/DS18B20-Datasheet.pdf> [Accessed: 19-May-2025]

[40] DIYMore, PH-4502C Analog pH Sensor Kit, [Online]. Available: <https://cdn.awsli.com.br/969/969921/arquivos/ph-sensor-ph-4502c.pdf> [Accessed: 19-May-2025]

[41] "Data Sheet – Farnell 3161977," Farnell, [Online]. Available: <https://www.farnell.com/datasheets/3161977.pdf> [Accessed: 19-May-2025]

[42] Google Cloud, "What is artificial intelligence (AI)?", Google Cloud, [Online]. Available: <https://cloud.google.com/learn/what-is-artificial-intelligence> [Accessed: 24-Apr-2025]

[43] TechTarget, "Artificial Intelligence of Things (AIoT)," IoT Agenda, [Online]. Available: <https://www.techtarget.com/iotagenda/definition/Artificial-Intelligence-of-Things-AIoT> [Accessed: 24-Apr-2025]

[44] V. D. Gowda, M. Kaur, D. Srinivas, K. D. V. Prasad, and R. Shekhar, "AIoT Integration: Advancements and Challenges in Smart Sensing Technologies for Smart Devices" in Impact of Artificial Intelligence on IoT: A Survey, ResearchGate, Jan. 2024. [Online]. Available: https://www.researchgate.net/figure/Integration-of-IoT-to-AI-technique_fig2_378049431 [Accessed: 24-Apr-2025]

[45] V. Kanade, "What Is Machine Learning? Definition, Types, Applications, and Trends" Spiceworks, Apr. 4, 2022. [Online]. Available: <https://www.spiceworks.com/tech/artificial-intelligence/articles/what-is-ml/> [Accessed: 24-Apr-2025]

[46] Microsoft, "What is a machine learning model?" Microsoft Learn, Feb. 13, 2025. [Online]. Available: <https://learn.microsoft.com/en-us/windows/ai/windows-ml/what-is-a-machine-learning-model> [Accessed: 24-Apr-2025]

[47] "Introduction to Deep Learning" GeeksforGeeks, Apr. 9, 2025. [Online]. Available: <https://www.geeksforgeeks.org/introduction-deep-learning/> [Accessed: 24-Apr-2025]

[48] H. Patel, "How to Visualize the Layers of a Deep Learning Model" Medium, Sep. 3, 2020. [Online]. Available: <https://medium.com/@hrushi669/how-to-visualize-the-layers-of-a-deep-learning-model-30d3fd7de7e5> [Accessed: 24-Apr-2025]

- [49] U. S. Hashmi, S. Akbar, R. Adve, P. W. Moo, and J. Ding, “Artificial intelligence meets radar resource management: A comprehensive background and literature review” *IET Radar, Sonar & Navigation*, vol. 17, no. 1, pp. 153–178, Jan. 2023, doi: 10.1049/rsn2.12337 [Accessed: 24-Apr-2025]
- [50] R. Baena-Navarro, Y. Carriazo-Regino, and F. Torres-Hoyos, “IoT Monitoring Dataset of Water Quality and Tilapia (*Oreochromis niloticus*) Health in Aquaculture Ponds in Montería, Colombia (2024)” *Mendeley Data*, vol. 1, 2024. [Online]. Available: <https://doi.org/10.17632/3g2b4sh65m.1> [Accessed: 16-Jun-2025].
- [51] Scikit-learn, “sklearn.ensemble.RandomForestRegressor” Scikit-learn documentation, [Online]. Available: <https://scikit-learn.org/stable/modules/generated/sklearn.ensemble.RandomForestRegressor.html> [Accessed: 19-Jun-2025].
- [52] GeeksforGeeks, “Random Forest Classifier using Scikit-learn” GeeksforGeeks, [Online]. Available: <https://www.geeksforgeeks.org/random-forest-classifier-using-scikit-learn/> [Accessed: 20-Jun-2025].
- [53] GeeksforGeeks, “XGBoost in Machine Learning” GeeksforGeeks, [Online]. Available: <https://www.geeksforgeeks.org/machine-learning/xgboost/> [Accessed: 20-Jun-2025].
- [54] GeeksforGeeks, “Support Vector Machine (SVM) Algorithm” GeeksforGeeks, [Online]. Available: <https://www.geeksforgeeks.org/machine-learning/support-vector-machine-algorithm/> [Accessed: 20-Jun-2025].
- [55] GeeksforGeeks, “What is Isolation Forest?” GeeksforGeeks, [Online]. Available: <https://www.geeksforgeeks.org/machine-learning/what-is-isolation-forest/> [Accessed: 21-Jun-2025].
- [56] Investopedia, “Autoregressive Integrated Moving Average (ARIMA)” Investopedia, [Online]. Available: <https://www.investopedia.com/terms/a/autoregressive-integrated-moving-average-arima.asp> [Accessed: 22-Jun-2025].
- [57] GeeksforGeeks, “Deep Learning: Introduction to Long Short-Term Memory” GeeksforGeeks, [Online]. Available: <https://www.geeksforgeeks.org/deep-learning-introduction-to-long-short-term-memory/> [Accessed: 22-Jun-2025].
- [58] GeeksforGeeks, “What is Python?” GeeksforGeeks, [Online]. Available: <https://www.geeksforgeeks.org/what-is-python/> [Accessed: 04-Jun-2025].

الجمهورية الجزائرية الديمقراطية الشعبية
وزارة التعليم العالي والبحث العلمي

جامعة عبد الحميد بن باديس-مستغانم-
كلية العلوم الدقيقة والإعلام الآلي

تصريح شرفي خاص بالالتزام بقواعد النزاهة العلمية
لإنجاز البحث

أنا الممضي أدناه،
الطالب(ة): بلحميتي ندى الزهراء رقم التسجيل الجامعي: 202037035199
الحامل لبطاقة التعريف الوطنية رقم: 110020977012080000 والصادرة بتاريخ:
13/11/2021

عن بلدية مزغران- مستغانم
المسجل ب: كلية العلوم الدقيقة والإعلام الآلي / قسم : إعلام آلي ورياضيات
شعبة : إعلام آلي / التخصص : ذكاء اصطناعي لإنترنت الأشياء
والمكلف بإنجاز مذكرة ماستر بعنوان:

“IOT-Based Intelligent Monitoring System for Aquaculture”

أصرح بشرفي أنني ألتزم بمراعاة المعايير العلمية والمنهجية ومعايير الأخلاقيات العلمية والذ
زاهة الأكاديمية المطلوبة في إنجاز البحث ، وأتحمل المسؤولية الشخصية عن كل المحت
وى المتضمن في البحث المذكور أعلاه .

التاريخ: 22/06/2025

إمضاء المعني

* ملحق القرار الوزاري رقم 933 المؤرخ في 28 جويلية 2016 الذي يحدد القواعد المتعلقة بالوقاية من السرقة العلمية
ومكافحتها.