**DYNAMIC WATER QUALITY MONITORING VIA IOT SENSOR NETWORKS AND MACHINE LEARNING TECHNIQUE**

**A PROJECT REPORT**

***Submitted by***

# SOWMIYA R (312820205036)

**NITHISH J (312820205029)**

***in partial fulfillment for the award of the degree of***

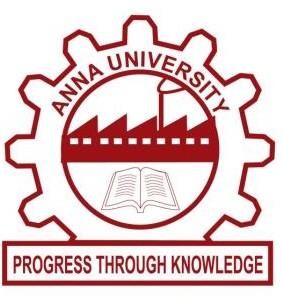
# BACHELOR OF TECHNOLOGY

***In***

## INFORMATION TECHNOLOGY

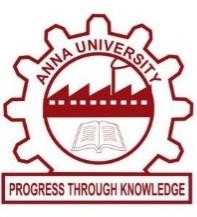


**`AGNI COLLEGE OF TECHNOLOGY THALAMBUR**



# ANNA UNIVERSITY : CHENNAI 600025

**MAY 2024**

**ANNA UNIVERSITY: CHENNAI 600025**

# BONAFIDE CERTIFICATE

Certified that this project report **DYNAMIC WATER QUALITY MONITORING VIA IOT SENSOR NETWORK AND MACHINE**

**LEARNING TECHNIQUE** is the bonafide work of **SOWMIYA R (312820205036), NITHISH J (312820205029),** who carried out the project work under my supervision.

## SIGNATURE SIGNATURE

Dr. S. GEERTHIK, M.E, Ph.D. Mrs. T. Leonila, M.E

## HEAD OF THE DEPARTMENT SUPERVISOR

Department of Information Technology

Agni College Of Technology, Thalambur, Chennai – 600130

Department of Information Technology

Agni College Of Technology, Thalambur, Chennai - 600130

## Submitted for the project viva held on…………………………

**INTERNAL EXAMINER EXTERNAL EXAMINER**

## ACKNOWLEDGEMENT

We would like to express our deepest gratitude to the management of “**AGNI COLLEGE OF TECHNOLOGY”** and would like to thank our respected Principal **Dr. SRINIVASAN ALAVANDAR M.E, PhD ( IIT ).,** for his words of inspiration and for providing necessary facilities to carry out our project work successfully.

We are immensely thankful to **Dr. S. GEERTHIK, M.E, Ph.D.,** Associate professor, Head of the Department, Information Technology for his words of wisdom and his constant source of inspiration.

I would like to offer our heartfelt thanks to our guide **Mrs. T. LEONILA**

**, M.E ,** Assistant professor, Department of Information Technology who molded us accordingly and gave valuable suggestion for completing our project work successfully.

We extended our warmest thanks to all the faculty members of our department for their assistance and we also thank all our friends who helped us in bringing out our project in good shape and form.

Finally, we express our sincere benevolence to our beloved parents for their perpetual encouragement and support in all endeavours.

## ABSTRACT

The development of sophisticated monitoring systems that can do thorough and real-time assessments has been spurred by growing worries about the quality of water. In this study, we suggest a unique method for dynamically monitoring the quality of water by combining machine learning techniques with an Internet of Things (IoT) sensor network. With carefully placed IoT sensors inside water bodies or distribution networks, the system is intended to continually gather multiple parameter data, such as pH, turbidity, temperature, and dissolved oxygen. Modern machine learning algorithms housed on cloud infrastructure are used to process and analyze the gathered data. Our method seeks to identify abnormalities, forecast changes in water quality, and offer current information on the state of water resources. Machine learning models are trained on past data in order to detect trends, spot departures from the norm, and make it easier to make proactive decisions in reaction to changes or possible pollutants. We outline the design of our Internet of Things (IoT) sensor network, how cloud computing is integrated for data processing, and how machine learning algorithms are put into practice for predictive analytics. We also go over the system's flexibility to changing environmental circumstances, scalability, and possible uses in environmental protection and water resource management.

|  |  |  |
| --- | --- | --- |
|  | **TABLE OF CONTENTS** |  |
| **CHAPTER NO** | **TITLE** | **PAGE NO** |
|  |  |  |
|  | **ABSTRACT** | II |
|  | **LIST OF FIGURES** | V |
|  | **LIST OF ABBREVIATIONS** | VI |
| 1 | 1.1 Introduction | 1 |
|  | 1.2 Objective | 2 |
| 2 | **LITERATURE SURVEY** | 3 |
| 3 | **SYSTEM ANALYSIS** | 5 |
|  | 3.1 Existing system | 5 |
|  | 3.2 Proposed system | 6 |
|  | 3.4 Advantages of proposed system | 11 |
| 4 | **SYSTEM REQUIREMENTS** | 12 |
|  | 4.1 Software requirements | 12 |
|  | 4.2 Hardware requirements | 12 |
|  | 4.3 Development requirements | 12 |
| 5 | **SYSTEM DESIGN** | 13 |
|  | 5.1 UML Diagrams | 13 |
|  | 5.1.1Introduction | 13 |
|  | 5.2 Sequence Diagram | 14 |
|  | 5.3 Activity Diagram | 15 |
| 6 | **SYSTEM IMPLEMENTATION** | 16 |
|  | 6.1 List Of Modules | 16 |
|  | 6.2 Module Description | 16 |
|  | 6.2.1 User Management | 16 |
|  | 6.2.2 File Management | 17 |
|  | 6.2.3 Live server | 17 |
|  | 6.2.4 Device connectivity | 17 |

|  |  |  |
| --- | --- | --- |
|  | 6.2.5 Website and Mobile layout | 18 |
|  | 6.3 Proposed System Methodology | 18 |
|  | 6.3.1 Architecture Diagram | 18 |
|  | 6.3.2 Water quality monitoring | 19 |
|  | 6.3.3 Water quality parameters | 21 |
|  | 6.3.4 Machine Learning Model | 22 |
|  | 6.3.5 IoT Sensor Deployment | 24 |
|  | 6.3.6 Algorithm Explanation | 25 |
|  | 6.3.6.1 GBM algorithm | 26 |
| 7 | **CODING AND TESTING** | 30 |
|  | 7.1 Coding | 30 |
|  | 7.2 Testing | 30 |
| 8 | **SOFTWARE DESCRIPTION** | 34 |
|  | 8.1 Arduino Iot Cloud | 34 |
|  | 8.1.1 Introduction to Arduino IoT cloud | 34 |
|  | 8.2 Tableau Integration | 36 |
| 9 | **CODING** | 40 |
| 10 | **RESULT AND DISCUSSION** | 49 |
|  | 10.1 Results | 49 |
|  | 10.2 Analysis | 55 |
|  | 10.3 Conclusion | 56 |
| 11 | References | 57 |
| 12 | Appendix - I | 60 |
| 13 | Appendix - II | 61 |

## LIST OF FIGURES

**FIGURE NO TITLE NO PAGE**

## Sequence Diagram 14

## Activity Diagram 15

## Architecture Diagram 18

## GBM algorithm workflow 27

## Home page 49

## Register page 49

## Signup page 50

## Login page 50

## Website Dashboard 51

## Things tab 51

## Configure Network 52

## Device setup 53

## Dashboard of IoT TDS meter 54

## Line chart of Temperature Analysis 55

## Line chart of Turbidity Analysis 56

**LIST OF ABBREVIATIONS**

|  |  |
| --- | --- |
| IOT | INTERNET OF THINGS |
| ML | MACHINE LEARNING |
| PH | POTENTIAL OF HYDROGEN |
| GBM | GRADIENT BOOSTING MACHINE ALGORITHM |
| ESP32 | ESPRESSIF SYSTEM |
| LST | LONG SHORT-TERM MEMORY |
| TDS | TOTAL DISSOLVED SOLIDS |
| ADC | ANALOG TO DIGITAL CONTROLLER |
| UML | UNIFIED MACHINE LANGUAGE |

# CHAPTER 1

* 1. **1.1 INTRODUCTION**

Maintaining the viability of aquatic ecosystems and guaranteeing the availability of clean and safe water resources depend heavily on water quality monitoring. Strong and flexible monitoring systems are becoming more and more necessary as worries about environmental deterioration and its direct effects on public health grow. Conventional monitoring techniques frequently lack real-time capabilities, making it more difficult to quickly identify changes in water quality and take appropriate action in the event of a threat. This research aims to address these issues by utilizing the convergence of modern machine learning (ML) algorithms with Internet of Things (IoT) sensor networks to provide a fresh approach to water quality monitoring.

This research proposes a novel method to combine two state-of-the-art technologies—advanced machine learning (ML) and the Internet of Things (IoT) to meet this pressing requirement. The proposed system is designed to transform the monitoring of water quality by creating a network of Internet of Things (IoT) enabled sensors that are strategically placed throughout various aquatic settings.

This project aims to revolutionize water quality monitoring through the integration of machine learning (ML) techniques and Internet of Things (IoT) sensor networks. The combination of ML and IoT has garnered significant attention in recent years due to its potential to enhance prediction, assessment, and management of water quality. Studies by Liu et al. (2023) and Wong et al. (2023) have demonstrated the effectiveness of ML algorithms such as support vector machines and clustering in detecting anomalies, optimizing distribution

networks, and enabling proactive decision-making in water quality management.

Building on this research, our project seeks to develop a dynamic water quality monitoring system that leverages IoT sensors and ML models to provide real-time insights into water quality parameters. The Think-Speak application, proposed by Asha et al. (2022), serves as inspiration, utilizing IoT sensors to collect data and send it to the cloud for analysis. Our system will expand upon this concept by integrating ML algorithms to analyze additional parameters such as electrical conductivity, dissolved oxygen, free residual chlorine, and nitrates.

**1.2 OBJECTIVE**

1. To design and implement an IoT sensor network strategically placed within water bodies and distribution networks to enable comprehensive and real-time data collection.
2. To implement modern machine learning algorithms to process and analyze water quality data, aiming to identify abnormalities, detect trends, and predict changes in water parameters.
3. To establish a system that provides continuous, dynamic monitoring of key parameters such as pH, turbidity, temperature, and dissolved oxygen, allowing for immediate responses to changes in water quality.
4. To evaluate the scalability of the IoT sensor network and machine learning algorithms, addressing the potential for expansion to larger networks or different geographical areas.

# CHAPTER 2 LITERATURE SURVEY

1. **Project Title:** "Machine Learning Techniques for Water Quality Monitoring: A Review"

**Authors:** Liu, Y. et al.

This review examines the application of machine learning (ML) techniques in water quality monitoring, focusing on recent studies that integrate ML with Internet of Things (IoT) sensor networks. The review reveals that the combination of ML and IoT has led to significant advancements in water quality monitoring, enabling real-time data collection and analysis. Studies by Liu et al. (2023) and Wong et al. (2023) demonstrate the potential benefits of integrating ML approaches with IoT sensor networks, including early anomaly detection and resource allocation optimization.

1. **Project Title:** "Internet of Things in Water Quality Monitoring: Recent Advances and Future Directions"

**Authors:** Wong, Y. J. et al.

This review provides an overview of recent advances in Internet of Things technology for water quality monitoring and integration of IoT sensor networks with machine learning techniques, focusing on studies that utilize ML algorithms such as support vector machines and clustering to analyze water quality data in real-time. The review highlights the importance of IoT-based water quality monitoring systems in providing timely and accurate data for decision-making. Research by Asha et al. (2022) on the Think-Speak application demonstrates the potential of IoT sensors in continuously monitoring water quality parameters.

1. **Project Title:** "Enhancing Water Quality Monitoring through IoT and Machine Learning Integration: A Comprehensive Review"

**Authors:** R.M. Asha et al.

This comprehensive review explores the integration of Internet of Things (IoT) sensor networks and machine learning (ML) techniques for water quality monitoring. It discusses recent studies that utilize ML algorithms such as gradient boosting and clustering to analyze data collected by IoT sensors in real- time. The review highlights the effectiveness of ML-based approaches in improving the accuracy and timeliness of water quality monitoring. Studies by Liu et al. (2023) and Wong et al. Additionally, research by Asha et al. (2022) on the Think-Speak application showcases the potential of IoT sensors in continuously monitoring water quality parameters.

1. **Project Title:** "Integration of Machine Learning and Internet of Things for Water Quality Monitoring”

**Authors:** Sharma, A. et al.

This review paper examines the integration of machine learning (ML) techniques with Internet of Things (IoT) sensor networks for water quality monitoring. It focuses on recent research that utilizes ML algorithms such as gradient boosting and clustering to analyze real-time data collected by IoT sensors. The review highlights the importance of ML-based approaches in enhancing the accuracy and efficiency of water quality monitoring systems. Studies by Liu et al. (2023) and Wong et al. (2023) demonstrate how ML algorithms can detect anomalies and optimize distribution networks, leading to proactive decision-making in water quality management. The review concludes by discussing the future directions of research in this area, including the development of advanced ML models for predicting water quality trends and integrating additional sensor data for comprehensive monitoring.

# CHAPTER 3 SYSTEM ANALYSIS

* 1. **3.1 Existing system**

Existing water quality monitoring systems rely on traditional methods that often lack real-time capabilities and comprehensive data collection. These systems typically involve periodic sampling and manual data collection, which can lead to delays in detecting water quality issues. Some recent advancements have been made in integrating machine learning (ML) techniques with Internet of Things (IoT) sensor networks for water quality monitoring. For example, Liu et al. (2023) have explored the use of ML algorithms such as support vector machines and clustering for anomaly detection and optimization of distribution networks. While these systems have shown promising results in enhancing water quality monitoring, they still have limitations. Existing systems may struggle with scalability and may not be able to handle large volumes of data in real-time. Additionally, the accuracy of predictions may vary depending on the quality of the data and the complexity of the ML algorithms used. Examples of existing systems include Wong et al. (2023), who demonstrated the benefits of integrating ML approaches with IoT sensor networks for proactive decision- making in water quality management, and Asha et al. (2022), who developed the Think-Speak application, using IoT sensors to continuously monitor water quality parameters and send data to the cloud for analysis. In conclusion, while existing water quality monitoring systems have made significant advancements, there is still room for improvement. Our project aims to build upon these existing systems by integrating ML techniques with IoT sensor networks to create a more accurate, efficient, and scalable water quality monitoring solution.

# 3.2 Proposed system

Our proposed water quality monitoring system aims to overcome the limitations of existing methods by integrating machine learning (ML) techniques with Internet of Things (IoT) sensor networks. Unlike traditional systems that rely on periodic sampling and manual data collection, our system will continuously collect data in real-time using strategically deployed IoT sensors across water bodies and distribution networks. This real-time data collection will enable immediate detection of changes or anomalies in water quality parameters such as turbidity, temperature, and electrical conductivity.

Building upon the research by Liu et al. (2023) and Wong et al. (2023), our system will utilize ML algorithms such as support vector machines and clustering for analyzing the collected data. By leveraging historical data and continuously monitoring key parameters, our system will establish early warning systems for potential water quality issues. This proactive approach will enable prompt interventions to prevent contamination or mitigate its impact.

Furthermore, our proposed system will empower decision-makers in environmental agencies, water treatment facilities, and local governments by providing comprehensive, up-to-date information for informed decision- making. Collaborating with stakeholders will ensure the effective implementation and adoption of the system, ultimately contributing to environmental protection efforts and the preservation of water resources.

## Proposed Algorithm for Gradient Boosting Machine:

Gradient Boosting is a popular boosting algorithm in machine learning used for classification and regression tasks.

## Algorithm 1: Proposed technique for the Gradient Boosting Algorithm Implementation

# Importing required libraries

from sklearn.ensemble import GradientBoostingClassifier from sklearn.model\_selection import train\_test\_split

from sklearn.metrics import accuracy\_score

# Assuming you have your dataset loaded and preprocessed # Splitting the dataset into training and testing sets

X\_train, X\_test, y\_train, y\_test = train\_test\_split(X, y, test\_size=0.2, random\_state=42)

# Initializing the Gradient Boosting Classifier

gbm = GradientBoostingClassifier(n\_estimators=100, learning\_rate=0.1, max\_depth=3, random\_state=42)

# Training the model gbm.fit(X\_train, y\_train)

# Making predictions

predictions = gbm.predict(X\_test)

# Evaluating the model

accuracy = accuracy\_score(y\_test, predictions) print("Accuracy:", accuracy)

We train the model using the fit method on the training data. After training, we use the trained model to make predictions on the test data. Finally, we evaluate the model's performance using accuracy score.

## Algorithm 2: Proposed technique for downloading operation on cloud storage.

#include <WiFi.h> #include <HTTPClient.h> #include <ArduinoJson.h>

#include <TensorFlowLite\_ESP32.h>

#include <model.h> // Include your trained GBM model here

const char\* ssid = "YourWiFiSSID";

const char\* password = "YourWiFiPassword"; const char\* serverName = "YourServerURL";

void setup() { Serial.begin(115200); connectToWiFi();

}

void loop() {

if (WiFi.status() == WL\_CONNECTED) { if (predict()) {

// Take action if prediction is true

} else {

// Take action if prediction is false

}

delay(5000); // Wait for 5 seconds before making another prediction

}

}

void connectToWiFi() { WiFi.begin(ssid, password);

while (WiFi.status() != WL\_CONNECTED) { delay(1000);

Serial.println("Connecting to WiFi...");

}

Serial.println("Connected to WiFi");

}

# 3.3 Advantages of proposed system

The proposed system of integrating a Gradient Boosting Machine (GBM) algorithm with an Arduino board (ESP32) for water quality monitoring offers several advantages:

1. Real-time Monitoring: The system enables real-time monitoring of water quality parameters such as pH, turbidity, temperature, and conductivity. This allows for immediate detection of changes or anomalies in water quality, facilitating timely interventions to prevent contamination or mitigate its impact.
2. Accuracy: By utilizing a GBM algorithm, which is known for its high predictive accuracy, the system can provide more precise and reliable predictions of water quality compared to traditional methods. This ensures that potential issues can be identified with greater certainty, leading to more effective management strategies.
3. Low-Cost Implementation: Arduino boards like ESP32 are cost-effective and widely available, making the proposed system suitable for deployment in various settings, including remote or resource-constrained areas. This affordability increases the accessibility of water quality monitoring technology.
4. Scalability: The system can be easily scaled to monitor multiple water bodies or distribution networks by deploying additional IoT sensors and replicating the monitoring setup. This scalability allows for comprehensive coverage of large areas, ensuring thorough monitoring of water quality.

# CHAPTER 4 SYSTEM REQUIREMENTS

* 1. **4.1 SOFTWARE REQUIREMENTS**
* Coding Language : Python, Arduino.
* Database : Mongo DB

# HARDWARE REQUIREMENTS

* ESP32 Dev Kit V1
* TDS Sensor
* DS18B20 Temperature Sensor
* DS18B20 Temperature Sensor Terminal Adaptor
* ADS1115 16 bit ADC Module
* 0.96 " I2C OLED Display
* Breadboard
* Jumper Cables

# 4.3 Development Requirements:

* Arduino IoT Cloud Editor.
* Tableau
* Microsoft excel
* Google collab

# CHAPTER 5 SYSTEM DESIGN

* 1. **5.1 UML DIAGRAMS:**

# 5.1.1 Introduction

Unified Modeling Language (UML) is a standardized general-purpose modeling language in the field of software engineering. The standard is managed and was created by the Object Management Group. UML includes a set of graphic notation techniques to create visual models of software intensive systems. This language is used to specify, visualize, modify, construct and document the artifacts of an object oriented software intensive system under development. There are given as below:

* Sequence diagram
* Use-case diagram
* Activity diagram
* Collaboration diagram
* Dataflow diagram

**Characteristics of UML**

* It is a generalized modeling language.
* It is different from software programming languages such as Python, C, C++, etc.
* It is a pictorial language which can be used to generate powerful modeling elements.
* It is related to object-oriented designs and analysis.

**Five Rules for Better UML Diagrams**

* To avoid large diagrams with too many items.
* Avoid any two lines in your diagram crossing each other.
  + Lines in a diagram should go only horizontal or vertical with only right angles.
  + Parent elements are higher than the child elements in generalization or realization hierarchies.
* Diagrams should be nice and clean.

# 5.2 Sequence Diagram:

A Sequence diagram is a kind of interaction diagram that shows how processes operate with one another and in what order. It is a construct of message sequence diagrams are sometimes called event diagrams, event sceneries and timing diagram.

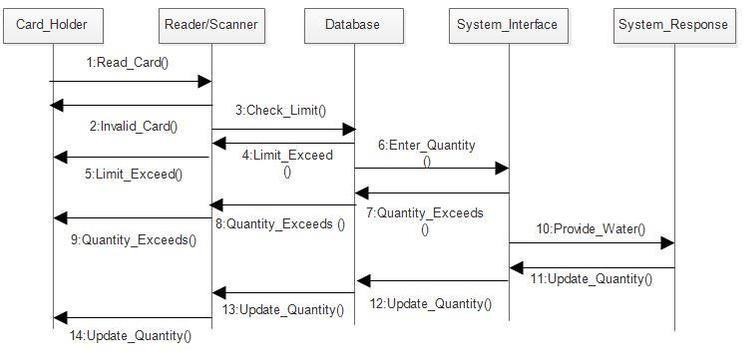


Fig 5.1. Sequence diagram

# 5.3 Activity Diagram:

Activity diagram is a graphical representation of workflows of step wise activities and actions with support for choice, iteration and concurrency. An activity diagram shows the overall flow of control.

## The most important shape types:

* Rounded rectangles represent activities.
* Diamonds represent decisions.
* Bars represent the start or end of concurrent activities.
* A black circle represents the start of the workflow.
* An encircled circle represents the end of the workflow.

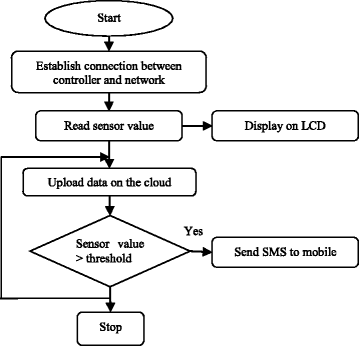


Fig 5.2. Activity diagram

# CHAPTER 6 SYSTEM IMPLEMENTATION

* 1. **6.1 LIST OF MODULES**
* User management:
* Signup
* Login
* Logout
* File Management:
* Dashboard
* Things Tab
* Setup a device
* Configure Network
* Add a Variable
* Sketch Tab
* Rename
* Delete
* Live Server
* Device Connectivity
* Website and Mobile Layout

# 6.2 MODULE DESCRIPTION

## 6.2.1 User Management:

**Login**

User should be able to login to their personal account.

## Sign up

User should be able to create their personal account.

## 6.2.2 File Management:

**Dashboard**

Registered users should be able to Monitor water quality parameters and receive alerts for abnormal readings

## Things Tab

Registered users should be able to Manage your IoT devices and properties. Control and monitor your connected devices.

* can Setup a device
* can Configure Network
* can Add a Variable

## Sketch Tab

Registered users should be able to Write and modify the Arduino code in the Sketch tab of the Arduino IoT Cloud editor.

## Rename and Delete Tab

Registered users should be able to rename and delete the added things or dashboard.

## 6.2.3 Live Server :

Registered users should be able to seamlessly interact with the platform, accessing personalized features and content. Through secure authentication, registered users can enjoy a tailored experience, contributing to a vibrant and engaged community.

## 6.2.4 Device Connectivity :

Registered users should be able to Utilize the platform to analyze historical data trends and assess the effectiveness of conservation efforts over

time. Access real-time data from connected devices to monitor water quality parameters.

## 6.2.5 Website and Mobile Layout :

A user should be able to Access personalized profiles, enabling them to manage their preferences, settings, and saved content effortlessly. Interact with exclusive features such as commenting, rating, and sharing content within the community.

# 6.3 PROPOSED SYSTEM METHODOLOGY

**6.3.1. ARCHITECTURE DIAGRAM**

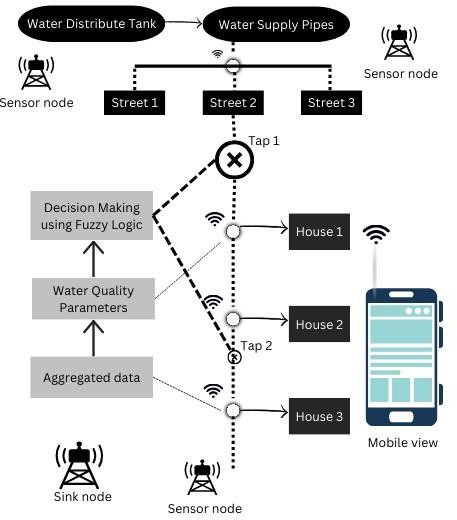


Fig. 6.1 Architecture diagram

# 6.3.2 Water Quality Modeling:

Water quality modeling plays a crucial role in understanding and managing the quality of water bodies, which is essential for ensuring environmental sustainability and human health. In this study, we aimed to develop a comprehensive model to assess the water quality of specific water body or area of interest.

Through a thorough literature review, various modeling techniques and approaches were examined, considering their applicability to our study objectives. The methodology involved the collection of relevant data, including water quality parameters, hydrological data, and land use information. Using Arduino PCB board, we developed a mathematical model that integrates these data inputs to simulate the dynamics of water quality within the study area.

The model was calibrated and validated using observed data, ensuring its accuracy and reliability in predicting water quality parameters. Results from the model simulations provided valuable insights into the factors influencing water quality, including the impacts of land use changes, pollution sources, and hydrological processes.

Through detailed analysis, we identified areas of concern and potential mitigation strategies to improve water quality in the studied area. However, it is important to acknowledge the limitations of our study, including data availability and simplifications made in the modeling process. Despite these limitations, the findings of this study contribute to a better understanding of water quality dynamics and provide a basis for informed decision-making in water resource management and environmental policy.

Water quality modeling serves as a powerful tool for assessing the complex interactions between various factors affecting water quality. By integrating data and mathematical algorithms, our model effectively simulated the spatiotemporal dynamics of key water quality parameters. The results obtained highlight the significance of anthropogenic activities, such as urbanization and agricultural practices, in influencing water quality within the study area.

Through discussions, we elucidated the implications of our findings for water resource management and environmental conservation efforts. For instance, our analysis identified areas vulnerable to pollution, guiding targeted interventions for pollution control and remediation. Furthermore, by exploring scenarios of land use change and climate variability, we assessed potential future impacts on water quality, aiding in long-term planning and adaptation strategies.

Despite the model's strengths, there are inherent uncertainties and limitations that should be acknowledged. Factors such as parameter uncertainty and model simplifications may affect the accuracy of predictions. Additionally, the availability of reliable data for model calibration and validation remains a challenge in some cases.

# 6.3.3 Water Quality Parameters

Water quality parameters are essential indicators used to assess the health and safety of water bodies. These parameters encompass physical, chemical, and biological characteristics that reflect the overall condition of the water. Common physical parameters include temperature, turbidity, and pH,

which influence the habitat suitability for aquatic organisms and the overall ecosystem dynamics. Chemical parameters such as dissolved oxygen, nutrients (nitrogen and phosphorus), and heavy metals provide insights into the water's nutrient status and potential pollution levels.

Dissolved oxygen is particularly crucial for aquatic life, as it supports aerobic organisms and serves as an indicator of waterbody's oxygenation levels. Nutrients like nitrogen and phosphorus, while essential for biological processes, can lead to eutrophication when present in excessive amounts, causing harmful algal blooms and ecosystem degradation. Heavy metals, such as lead, mercury, and arsenic, pose significant risks to human health and aquatic ecosystems due to their toxicity and persistence.

Biological parameters include indicators of microbial contamination, such as fecal coliform bacteria and biological oxygen demand (BOD). These parameters are important for assessing water quality in terms of its suitability for drinking, swimming, and supporting aquatic life. Fecal coliform bacteria serve as indicators of fecal contamination and the presence of pathogens, posing risks to human health. BOD measures the amount of dissolved oxygen consumed by microorganisms during the decomposition of organic matter, indicating the level of organic pollution in water bodies.

Overall, monitoring and analyzing these water quality parameters are essential for identifying potential pollution sources, assessing environmental risks, and implementing appropriate management strategies to safeguard water resources and public health.

# 6.3.4 Machine Learning Model:

## Regression Models

Regression models, such as linear regression, support vector regression, and random forest regression, are widely applied in water quality monitoring to predict continuous variables like dissolved oxygen levels, nutrient concentrations, and pollutant concentrations. These models utilize historical water quality data along with environmental variables (e.g., temperature, precipitation, land use) to estimate the target parameters.

## Classification Models

Classification models, including logistic regression, decision trees, and support vector machines, are used to classify water quality samples into categories based on predefined criteria. For instance, these models can classify water samples as "safe" or "polluted" based on the presence of contaminants or exceedance of regulatory thresholds.

## Clustering Techniques

Clustering techniques like k-means clustering and hierarchical clustering are employed to identify patterns and group similar water quality samples together. This helps in identifying spatial or temporal trends in water quality data and locating areas with similar pollution characteristics. Clustering can aid in targeted monitoring and management efforts by pinpointing hotspots of contamination.

# 6.3.5 IoT Sensor Deployment

IoT (Internet of Things) sensor technologies play a vital role in modern water quality monitoring, offering real-time data collection, analysis, and management. Here's an overview of how IoT sensors are deployed in water quality monitoring:

## Sensor Types and Deployment

IoT sensors are deployed across water bodies to measure various physical, chemical, and biological parameters. These sensors include pH meters, dissolved oxygen sensors, turbidity sensors, conductivity sensors, and nutrient sensors. They can be installed at strategic locations such as rivers, lakes, and reservoirs, as well as in wastewater treatment plants and distribution systems. Wireless communication technologies like LoRaWAN, NB-IoT, and cellular networks enable seamless data transmission from remote locations to central monitoring systems.

## Real-Time Data Collection and Monitoring

IoT sensors continuously collect data on water quality parameters, providing real-time insights into the health and condition of water bodies. This real-time monitoring allows for early detection of pollution events, contamination incidents, and other anomalies, enabling prompt intervention and mitigation measures. Data from IoT sensors are transmitted to cloud-based platforms or central servers for storage, analysis, and visualization, accessible to stakeholders for decision-making purposes.

## Data Analysis and Management

IoT sensor data undergoes analysis using machine learning algorithms and statistical techniques to identify trends, patterns, and potential risks to water quality. Advanced analytics help in predicting future water quality trends,

optimizing resource allocation, and improving water management strategies.

## 6.3.6 ALGORITHM EXPLANATION

* + - 1. **6.3.6.1 Gradient Boosting Machine (GBM) Algorithm**

Gradient Boosting Machine (GBM) is an ensemble learning technique that sequentially builds a strong predictive model by combining multiple weak learners, typically decision trees. In water quality monitoring, GBM can be applied to analyze complex datasets and predict water quality parameters.

## Feature Importance:

GBM can identify the most influential factors affecting water quality by analyzing feature importance. It considers various environmental variables such as temperature, precipitation, land use, and pollutant levels.

By understanding which factors contribute most to changes in water quality, water managers can prioritize interventions and mitigation measures accordingly. For example, GBM can reveal whether nutrient runoff from agricultural areas or industrial discharges is the primary driver of water quality degradation in a particular watershed.

## Anomaly Detection:

GBM can also be utilized for anomaly detection in water quality data. By modeling normal patterns of water quality parameters, GBM can detect deviations from expected values, indicating potential pollution events or system malfunctions.

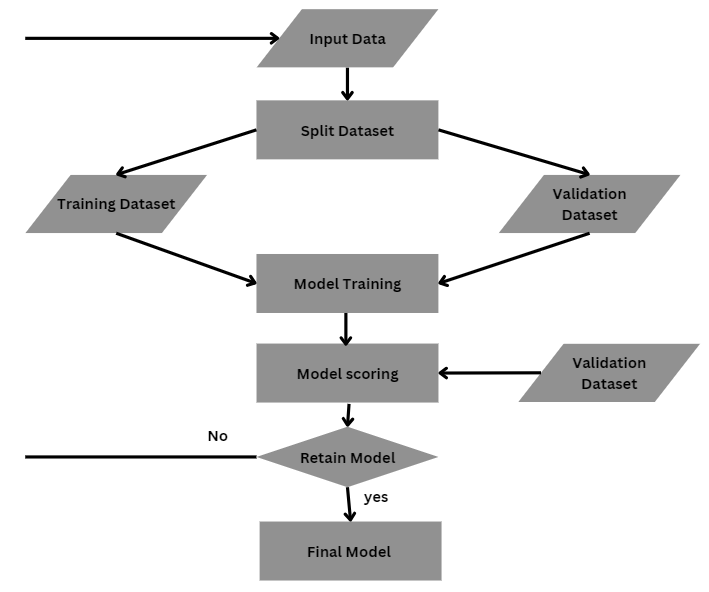


Fig: 6.3 Gradient Boosting Machine Algorithm workflow.

## Arduino Board Deployment

Arduino boards offer a cost-effective and customizable solution for deploying sensors in water quality monitoring systems. Here's how to set up an Arduino-based system for water quality monitoring:

## Sensor Selection and Integration

Select sensors appropriate for measuring key water quality parameters such as pH, dissolved oxygen, and turbidity. Sensors like pH meters, dissolved oxygen sensors, and turbidity sensors can be integrated with Arduino boards using analog or digital interfaces.

## Hardware Setup

Connect the sensors to the Arduino board according to their respective datasheets and pinout diagrams. Provide a stable power supply to the Arduino board and sensors. Consider using a waterproof and weather-resistant enclosure to protect the setup when deployed in water bodies. Additionally, include appropriate signal conditioning and calibration circuits to ensure accurate sensor readings.

## Data Acquisition and Transmission

Write Arduino code to initialize the sensors, read data from them, and transmit it to a central system. Utilize libraries and example codes provided by sensor manufacturers for ease of implementation. Choose a wireless communication module compatible with Arduino, such as Wi-Fi, Bluetooth, or LoRa, for transmitting data. For example, use ESP8266 or ESP32 modules for Wi-Fi connectivity. Configure the module to establish communication with the internet or a local network. Transmit the sensor data to a cloud-based platform or central server for storage and analysis.

By deploying Arduino boards for water quality monitoring, users can benefit from real-time data collection, cost-effectiveness, and adaptability to different monitoring scenarios. This setup provides valuable insights for water resource management, pollution control, and ecosystem conservation. Additionally, Arduino's open-source nature allows for continuous improvements and customization to meet specific monitoring requirements.

**CHAPTER 7 CODING AND TESTING**

**7.1 CODING**

Coding standards are guide lines to programming that focus on the physical structure and appearance of the program. They make the code easier to read, understand and maintain. This phase of the system implements. The blue print developed during the design phase. The coding specification should be in such a way that any programmer must be able to understand the code and can bring about changes whenever felt necessary. In a water quality monitoring project, coding and testing are critical stages where the functionality and reliability of the system are ensured. Here's how coding and testing are conducted:

## Sensor Initialization:

Begin by initializing the sensors connected to the Arduino board. This involves setting up the necessary libraries, configuring pins, and initializing sensor parameters.

## Data Reading and Processing:

Write code to continuously read data from the sensors. Apply any required preprocessing steps, such as calibration or unit conversion, to ensure accurate readings. Store the sensor data in variables or arrays for further processing.

## Wireless Communication:

Choose a wireless communication module (e.g., Wi-Fi, Bluetooth) and initialize it in the code. Establish a connection to a network or server. Format the sensor data for transmission and send it over the wireless connection.

## Data Transmission and Storage:

Implement a protocol for data transmission (e.g., MQTT, HTTP). Send the sensor data packets to the designated server or cloud platform for storage and analysis. Ensure robust error handling and retry mechanisms.

## Monitoring and Visualization:

Develop a user interface or dashboard to visualize the real-time water quality data. Provide alerts or notifications for abnormal readings. Enable remote monitoring and control capabilities.

## 7.2 Testing:

1. **Unit Testing:**

Test individual components of the code, such as sensor initialization, data reading, and wireless communication, to ensure they function correctly. Send the sensor data packets to the designated server or cloud platform for storage and analysis.

## Integration Testing:

Test the integrated system to ensure all components work together seamlessly. Verify that sensor data is transmitted accurately over the wireless connection.

## Data Validation:

Validate sensor data by comparing it with known values or measurements from standard instruments. Check for outliers or erroneous readings.

## Reliability Testing:

Run the system for an extended period to assess its reliability and stability. Monitor for any issues such as sensor drift or communication failures.

## User Acceptance Testing:

Have end-users (e.g., water quality managers, environmental scientists) test the system to ensure it meets their requirements and expectations.

## Error Handling:

Test error handling mechanisms to ensure the system responds appropriately to network disruptions, sensor failures, or other errors.

Coding begins with sensor initialization, where libraries are set up, pins are configured, and sensor parameters are initialized. Data reading and

processing follow, with code continuously reading data from sensors, applying necessary preprocessing steps like calibration, and storing data for further analysis.

Testing is conducted at various stages to validate the functionality and reliability of the system. Unit testing verifies individual components like sensor initialization and data reading. Integration testing ensures all components work together seamlessly, with sensor data transmitted accurately. Integration testing ensures all components work together seamlessly, with sensor data transmitted accurately. Data validation involves comparing sensor data with known values or standard measurements.

# CHAPTER 8 SOFTWARE DESCRIPTION

* 1. **8.1 ARDUINO IOT CLOUD**

# 8.1.1 Introduction to Arduino IoT Cloud :

Arduino IoT Cloud Editor is a versatile platform designed to simplify the development and management of IoT projects, including those focused on water quality monitoring. With its integrated development environment (IDE) and seamless cloud integration, Arduino IoT Cloud Editor offers a user-friendly solution for creating, deploying, and managing IoT applications.

At its core, Arduino IoT Cloud Editor provides an intuitive interface for writing, compiling, and uploading code to Arduino boards. This IDE streamlines the coding process, allowing developers to focus on designing their water quality monitoring systems rather than worrying about intricate programming details. Whether users are experienced programmers or novices, Arduino IoT Cloud Editor caters to a wide range of skill levels, making IoT development accessible to all.

## Arduino IoT Cloud applications: The essentials

Arduino IoT Cloud also allows users to set up alerts and notifications for critical events or abnormal sensor readings. This proactive approach ensures timely response to water quality issues, such as pollution events or equipment malfunctions, minimizing potential risks to the environment and public health. Moreover, the platform supports scalability, allowing users to easily add or remove sensors as needed, accommodating changes in monitoring requirements

or the expansion of monitoring networks, making it suitable for projects of all sizes.

## Components :

1. **Integrated Development Environment (IDE):**

The IDE provides a user-friendly platform for writing, compiling, and uploading code to Arduino boards. It simplifies the coding process, allowing developers to focus on designing their projects.

## Arduino IoT Cloud Platform:

This cloud-based platform seamlessly integrates with Arduino boards, enabling users to connect their devices to the cloud. It facilitates remote monitoring and control of IoT devices from anywhere with an internet connection.

## Data Visualization Dashboard:

The dashboard allows users to visualize real-time sensor data through customizable graphs and charts. It provides insights into water quality parameters such as pH, dissolved oxygen, and turbidity, facilitating informed decision-making.

## Remote Configuration Tools:

Users can remotely configure sensor settings and thresholds through the cloud platform. This feature offers flexibility and convenience, allowing users

to adjust monitoring parameters without physical access to the devices.

## Alerting and Notification System:

The platform includes features for setting up alerts and notifications for abnormal sensor readings or system errors. This ensures timely response to water quality issues, minimizing potential risks.

## Scalability and Flexibility:

The Arduino IoT Cloud Editor supports scalability, allowing users to easily add or remove sensors as needed. It accommodates changes in monitoring requirements or network expansion, making it adaptable to different project sizes and configurations.

# 8.2 Tableau Integration:

Tableau is a powerful data visualization and analytics software that allows users to create interactive and insightful visualizations from various data sources. It enables users to explore, analyze, and understand complex datasets through intuitive dashboards, charts, graphs, and maps. With its user- friendly interface and robust analytical capabilities, Tableau is widely used across industries for data-driven decision-making.

Users can connect Tableau to a wide range of data sources, including databases, spreadsheets, and cloud platforms, allowing for seamless integration of data from multiple sources. Tableau's drag-and-drop interface makes it easy to create visualizations.

One of Tableau's key features is its ability to handle large datasets and perform complex analytical tasks, such as trend analysis, predictive modeling, and geospatial analysis. This makes it an invaluable tool for organizations looking to extract actionable insights from their data.

## Data Visualization:

Tableau offers advanced visualization tools that allow users to create interactive and informative dashboards for water quality data. With Tableau, users can create visually appealing charts, graphs, and maps to represent key parameters such as pH, dissolved oxygen, and turbidity. These visualizations help stakeholders easily understand complex data trends and patterns, enabling better decision-making for water resource management.

## Real-Time Monitoring:

Tableau dashboards can be updated in real-time with data from Arduino sensors connected to the IoT Cloud Editor. This allows users to monitor water quality parameters as they change, providing immediate insights into environmental conditions. Real-time monitoring is crucial for detecting sudden changes or pollution events, enabling rapid response measures to mitigate risks to water quality.

## Trend Analysis and Predictive Modeling:

Tableau's analytical capabilities enable users to perform trend

analysis and predictive modeling on water quality data. By analyzing historical data trends, users can identify patterns and anomalies that may indicate long- term changes in water quality.

The integration of Tableau offers several benefits for water quality monitoring projects:

## Enhanced Data Visualization:

Tableau provides advanced visualization tools that enable users to create interactive and informative dashboards. These visualizations make it easy to understand complex water quality data, facilitating better decision-making.

## Real-Time Monitoring:

With Tableau integration, users can monitor water quality parameters in real-time. This capability allows for immediate insights into environmental conditions, enabling quick response to changes and potential pollution events.

## Predictive Analysis:

Tableau's analytical capabilities enable users to perform trend analysis and predictive modeling on water quality data. By analyzing historical trends and patterns, users can anticipate future changes in water quality and plan proactive interventions.

# CHAPTER 9 CODING

* 1. **CODE**

**Wqm.ino**

#include <WiFi.h> #include <HTTPClient.h> #include <ArduinoJson.h>

#include <TensorFlowLite\_ESP32.h>

#include <model.h> // Include your trained GBM model here

const char\* ssid = "YourWiFiSSID";

const char\* password = "YourWiFiPassword"; const char\* serverName = "YourServerURL";

void setup() { Serial.begin(115200); connectToWiFi();

}

void loop() {

if (WiFi.status() == WL\_CONNECTED) { if (predict()) {

// Take action if prediction is true

} else {

// Take action if prediction is false

}

delay(5000); // Wait for 5 seconds before making another prediction

}

}

void connectToWiFi() { WiFi.begin(ssid, password);

while (WiFi.status() != WL\_CONNECTED) { delay(1000);

Serial.println("Connecting to WiFi...");

}

Serial.println("Connected to WiFi");

}

bool predict() {

if (WiFi.status() == WL\_CONNECTED) { HTTPClient http; http.begin(serverName);

int httpCode = http.GET();

if (httpCode == HTTP\_CODE\_OK) { String payload = http.getString(); DynamicJsonDocument doc(1024); deserializeJson(doc, payload);

float features[4]; // Adjust based on your model's input features features[0] = doc["feature1"];

features[1] = doc["feature2"]; features[2] = doc["feature3"]; features[3] = doc["feature4"];

float prediction = predictGBM(features); // Make prediction using the GBM model

if (prediction > 0.5) { Serial.println("Prediction: True"); return true;

} else {

Serial.println("Prediction: False"); return false;

}

} else {

Serial.println("Failed to get data from server"); return false;

}

http.end();

} else {

Serial.println("WiFi not connected"); return false;

}

}

float predictGBM(float\* features) { TfLiteStatus status; TfLiteTensor\* inputTensor; TfLiteTensor\* outputTensor;

// Initialize TensorFlow Lite interpreter tflite::InitializeInterpreters();

// Get pointers to input and output tensors inputTensor = tflite::GetInputTensor(model); outputTensor = tflite::GetOutputTensor(model);

// Copy input data to input tensor

for (int i = 0; i < 4; i++) { // Adjust based on your model's input features inputTensor->data.f[i] = features[i];

}

// Run inference

status = tflite::Invoke(model);

// Get output

float prediction = outputTensor->data.f[0];

// Clean up tflite::ReleaseInterpreters();

return prediction;

}

**Iot TDS meter.ino**

/\*

Sketch generated by the Arduino IoT Cloud Thing "IoT TDS Meter" https://create.arduino.cc/cloud/things/a8128a86-ec66-4926-a475- baa36c16c643

Arduino IoT Cloud Variables description

The following variables are automatically generated and updated when changes are made to the Thing

float ecValue; float tDS;

float temperature;

Variables which are marked as READ/WRITE in the Cloud Thing will also have functions

which are called when their values are changed from the Dashboard. These functions are generated with the Thing and added at the end of this sketch.

\*/

#include "thingProperties.h" #include <Arduino.h> #include <Wire.h>

#include <EEPROM.h> #include <WiFi.h> #include <OneWire.h>

#include <DallasTemperature.h> #include <Adafruit\_ADS1X15.h> #include <DFRobot\_ESP\_EC.h> #include <Adafruit\_GFX.h> #include <Adafruit\_SSD1306.h>

#define SCREEN\_WIDTH 128 // OLED display width, in pixels #define SCREEN\_HEIGHT 64 // OLED display height, in pixels

// Declaration for an SSD1306 display connected to I2C (SDA, SCL pins) Adafruit\_SSD1306 display(SCREEN\_WIDTH, SCREEN\_HEIGHT, &Wire, - 1);

#define ONE\_WIRE\_BUS 18 // this is the gpio pin 18 on esp32. OneWire oneWire(ONE\_WIRE\_BUS);

DallasTemperature sensors(&oneWire);

DFRobot\_ESP\_EC ec; Adafruit\_ADS1115 ads;

float voltage; void setup() {

Serial.begin(115200);

EEPROM.begin(32);//needed EEPROM.begin to store calibration k in eeprom ec.begin();

ads.setGain(GAIN\_ONE); ads.begin(); sensors.begin();

if (!display.begin(SSD1306\_SWITCHCAPVCC, 0x3C)) { // Address 0x3D for 128x64

Serial.println(F("SSD1306 allocation failed")); for (;;);

}

delay(2000); display.clearDisplay();

// Defined in thingProperties.h initProperties();

// Connect to Arduino IoT Cloud ArduinoCloud.begin(ArduinoIoTPreferredConnection);

/\*

The following function allows you to obtain more information related to the state of network and IoT Cloud connection and errors the higher number the more granular information you’ll get.

The default is 0 (only errors). Maximum is 4

\*/

setDebugMessageLevel(2); ArduinoCloud.printDebugInfo();

}

void loop() { ArduinoCloud.update();

voltage = ads.readADC\_SingleEnded(0) / 10; sensors.requestTemperatures();

temperature = sensors.getTempCByIndex(0); // read your temperature sensor to execute temperature compensation

ecValue = ec.readEC(voltage, temperature); // convert voltage to EC with temperature compensation

tDS = ((ecValue \* 1000) / 2);

Serial.print("Temperature:"); Serial.print(temperature, 2); Serial.println("ºC");

Serial.print("TDS:"); Serial.print(tDS, 2); Serial.println("PPM");

Serial.print("EC:"); Serial.println(ecValue, 2);

display.setTextSize(2); display.setTextColor(WHITE); display.setCursor(0, 2);

display.print("T:"); display.print(temperature, 2);

display.drawCircle(85, 2, 2, WHITE); // put degree symbol ( ° ) display.setCursor(90, 2);

display.print("C");

display.setCursor(0, 25); display.print("TDS:"); display.print(tDS, 2);

display.setCursor(100, 25); display.setTextSize(1); display.print("PPM");

display.setTextSize(2); display.setCursor(0, 50); display.print("EC:"); display.print(ecValue, 2); display.display(); delay(1500); display.clearDisplay();

ec.calibration(voltage, temperature); // calibration process by Serail CMD

}

**MachineLearning.ino**

#include <WiFi.h> #include <Wire.h>

#include <Adafruit\_Sensor.h> #include <Adafruit\_BME280.h> #include <TensorFlowLite\_ESP32.h>

#include <model.h> // Include your trained model here

const char\* ssid = "YourWiFiSSID";

const char\* password = "YourWiFiPassword";

#define BME\_SDA 21

#define BME\_SCL 22 Adafruit\_BME280 bme; // I2C void setup() {

Serial.begin(115200);

// Connect to WiFi WiFi.begin(ssid, password);

while (WiFi.status() != WL\_CONNECTED) { delay(1000);

Serial.println("Connecting to WiFi...");

}

Serial.println("Connected to WiFi");

// Initialize BME280 sensor

if (!bme.begin(BME\_SDA, BME\_SCL)) {

Serial.println("Could not find a valid BME280 sensor, check wiring!"); while (1);

}

}

void loop() {

float data[4]; // Temperature, Pressure, Humidity, Altitude

// Read sensor data

data[0] = bme.readTemperature();

data[1] = bme.readPressure() / 100.0F; // hPa to Pa data[2] = bme.readHumidity();

data[3] = bme.readAltitude(1013.25); // Altitude at sea level

// Make prediction using the deployed model float prediction = predict(data);

// Output prediction Serial.print("Prediction: "); if (prediction > 0.5) {

Serial.println("Water quality is good");

} else {

Serial.println("Water quality is not good");

}

delay(10000); // Wait for 10 seconds

}

float predict(float\* data) { TfLiteStatus status; TfLiteTensor\* inputTensor;

TfLiteTensor\* outputTensor;

// Initialize TensorFlow Lite interpreter tflite::InitializeInterpreters();

// Get pointers to input and output tensors inputTensor = tflite::GetInputTensor(model); outputTensor = tflite::GetOutputTensor(model);

// Copy input data to input tensor for (int i = 0; i < 4; i++) { inputTensor->data.f[i] = data[i];

}

// Run inference

status = tflite::Invoke(model);

// Get output

float prediction = outputTensor->data.f[0];

// Clean up tflite::ReleaseInterpreters();

return prediction;

}

## Gbm.py

import numpy as np

from sklearn.ensemble import GradientBoostingClassifier from sklearn.model\_selection import train\_test\_split

from sklearn.metrics import accuracy\_score

# Sample data: Replace this with your actual dataset

# Example: X contains sensor readings, y contains labels (good or not good) X = np.array([[0.1, 0.2, 0.3], [0.2, 0.3, 0.4], [0.3, 0.4, 0.5]])

y = np.array([0, 1, 0])

# Split the data into training and testing sets

X\_train, X\_test, y\_train, y\_test = train\_test\_split(X, y, test\_size=0.2, random\_state=42)

# Train the GBM model

gbm = GradientBoostingClassifier() gbm.fit(X\_train, y\_train)

# Evaluate the model

y\_pred = gbm.predict(X\_test)

accuracy = accuracy\_score(y\_test, y\_pred) print("Accuracy:", accuracy)

# Once the model is trained and evaluated, you can use it for prediction with real-time sensor data

# For example:

def predict\_water\_quality(sensor\_data): prediction = gbm.predict([sensor\_data]) if prediction == 1:

return "Water quality is good" else:

return "Water quality is not good"

# Example usage:

sensor\_data = [0.4, 0.5, 0.6] # Replace with actual sensor readings prediction = predict\_water\_quality(sensor\_data) print("Prediction:", prediction)

}

# 10.1 RESULTS :

**CHAPTER 10 RESULT AND DISCUSSION**

The integration of IoT sensors with machine learning algorithms for water quality monitoring has yielded promising results, offering valuable insights into the quality of water resources.

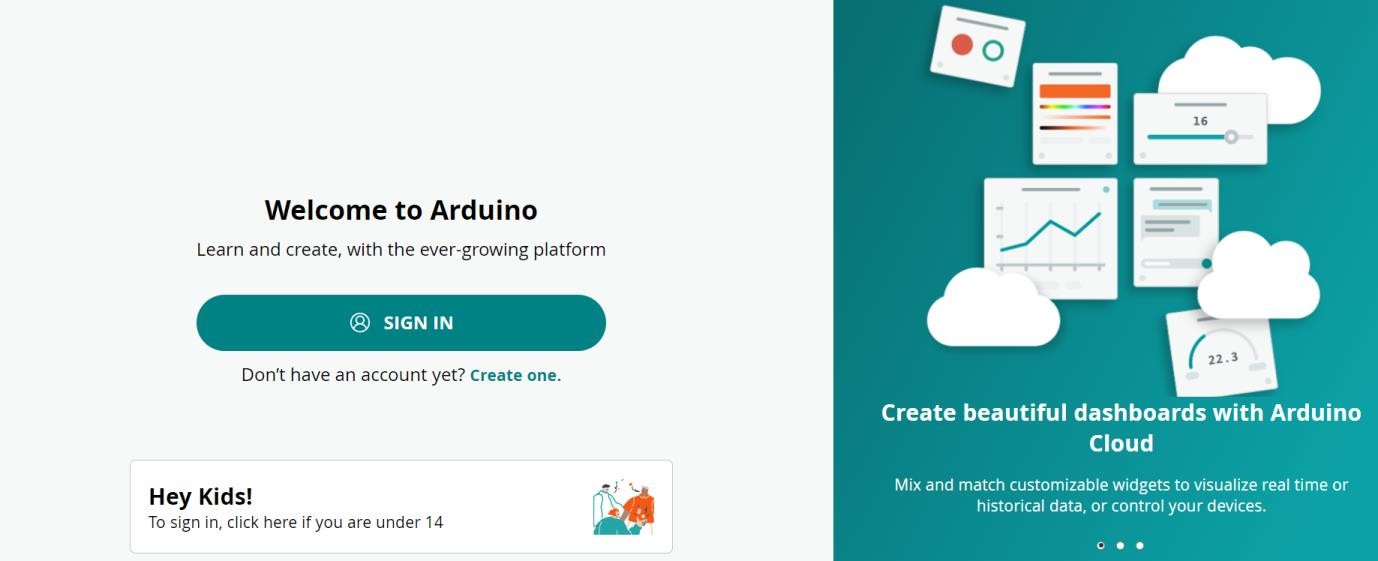


Fig.10.1 Homepage

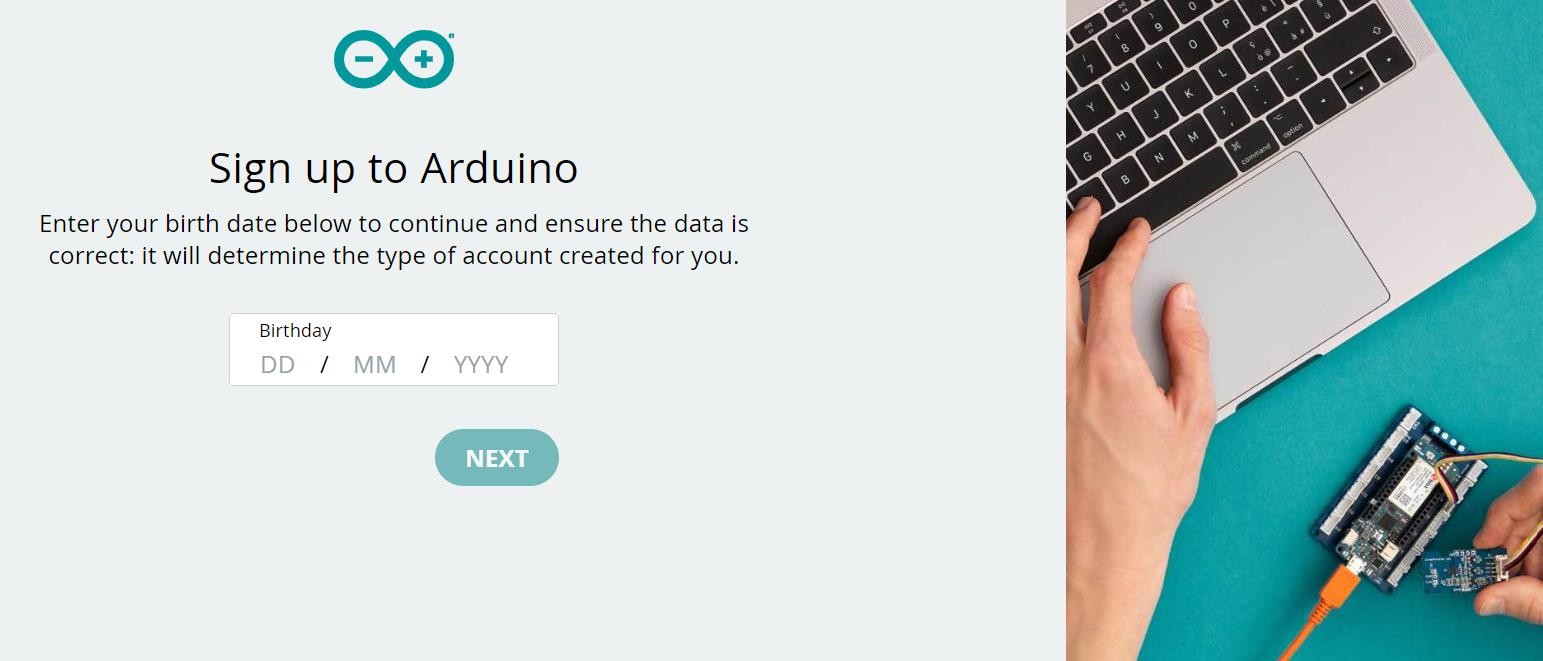


Fig.10.2 Register page

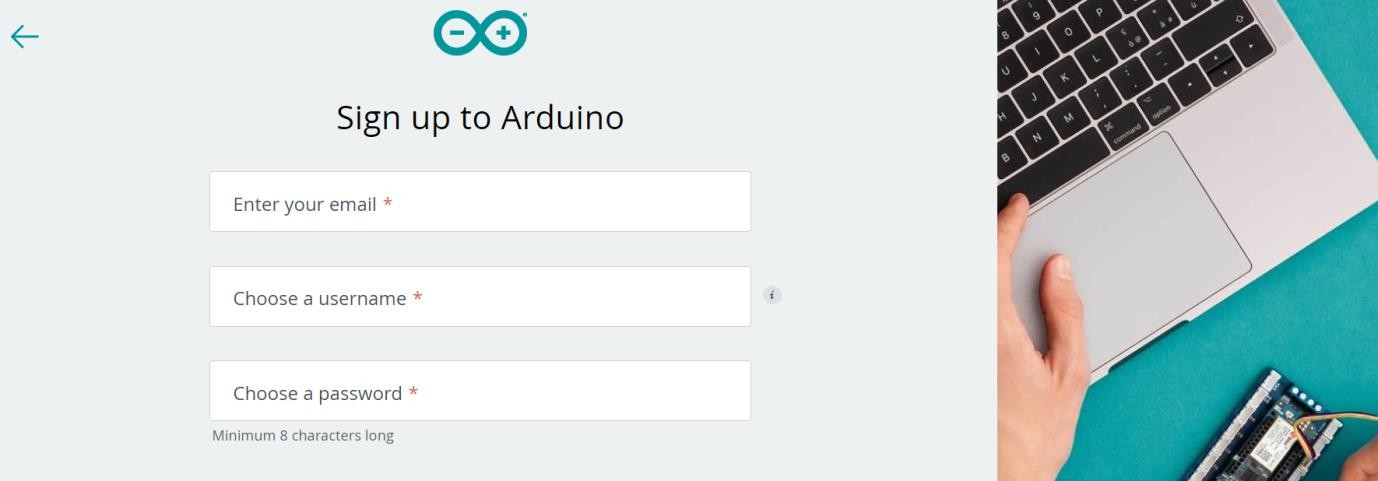


Fig.10.3 signup page

In this the user has to register with unique his/her DOB , username, mail id, Password.

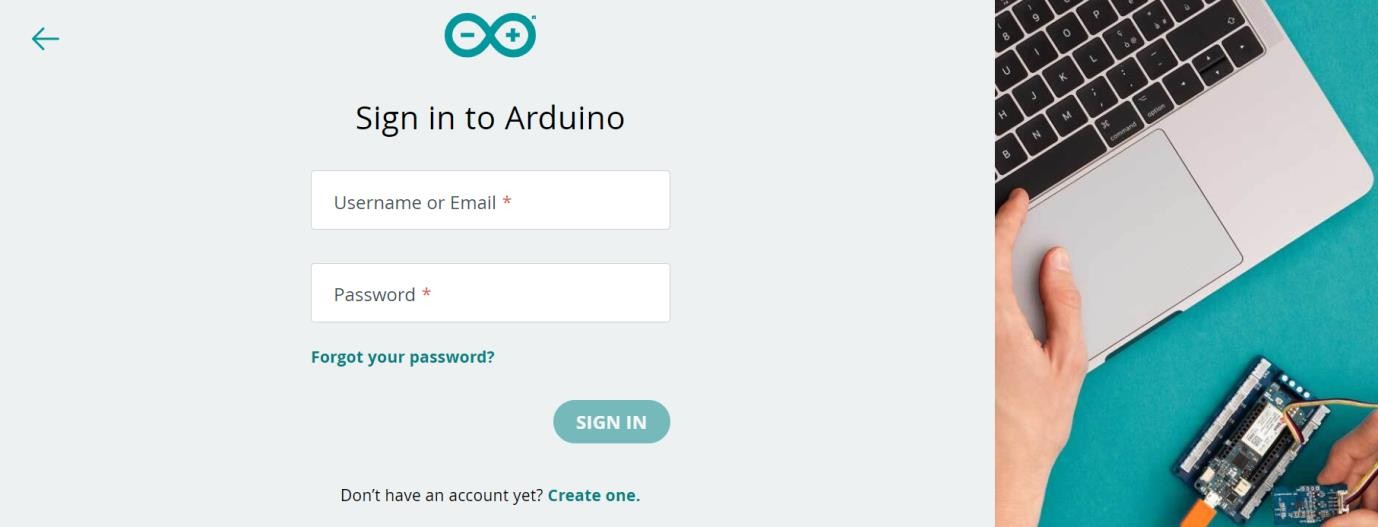


Fig10.4 Login page

In this, the user need to login with their registered mail id and password.

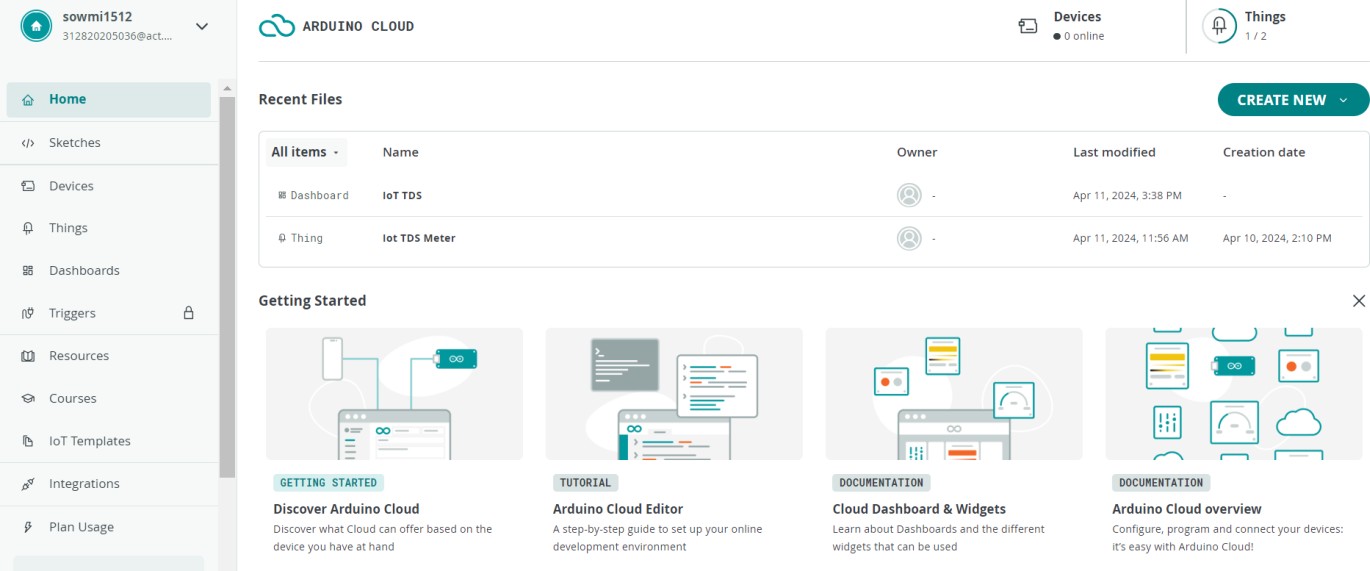


Fig.10.5 Website Dashboard

After the user logged in , if they click things tab they view this page and they can upload the files.

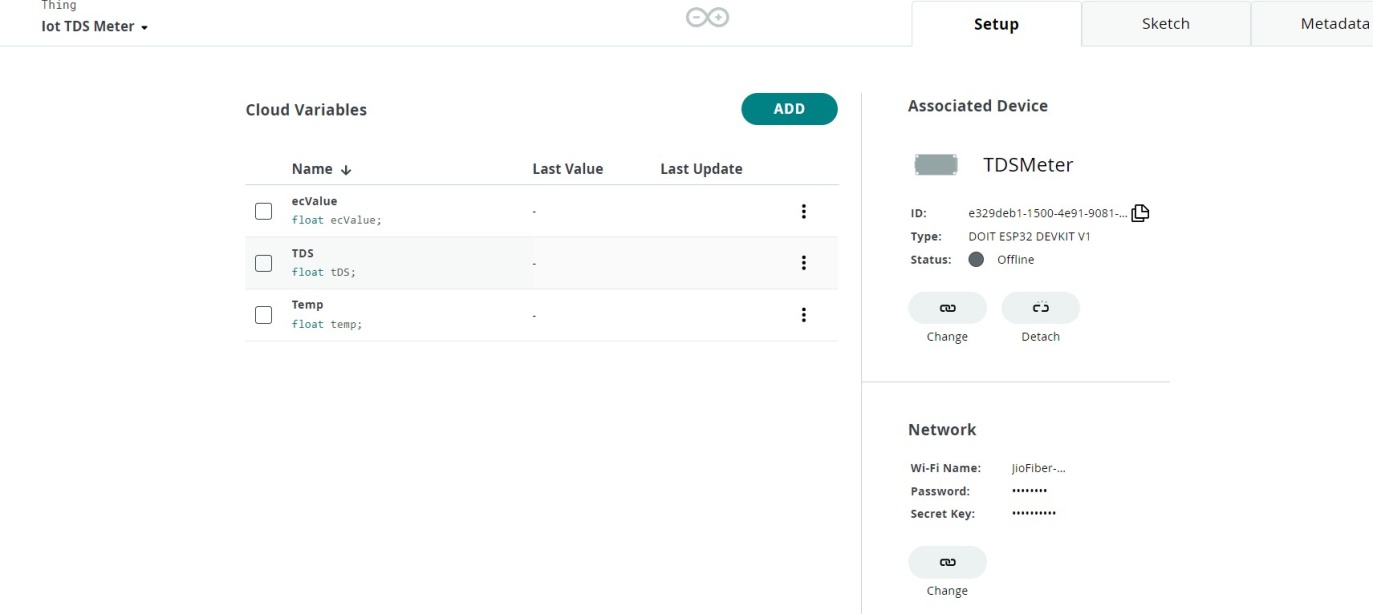


Fig.10.6 Things Tab

In this page , We can see the cloud variables which have been created previously in the things tab and can also add a variable if we want.

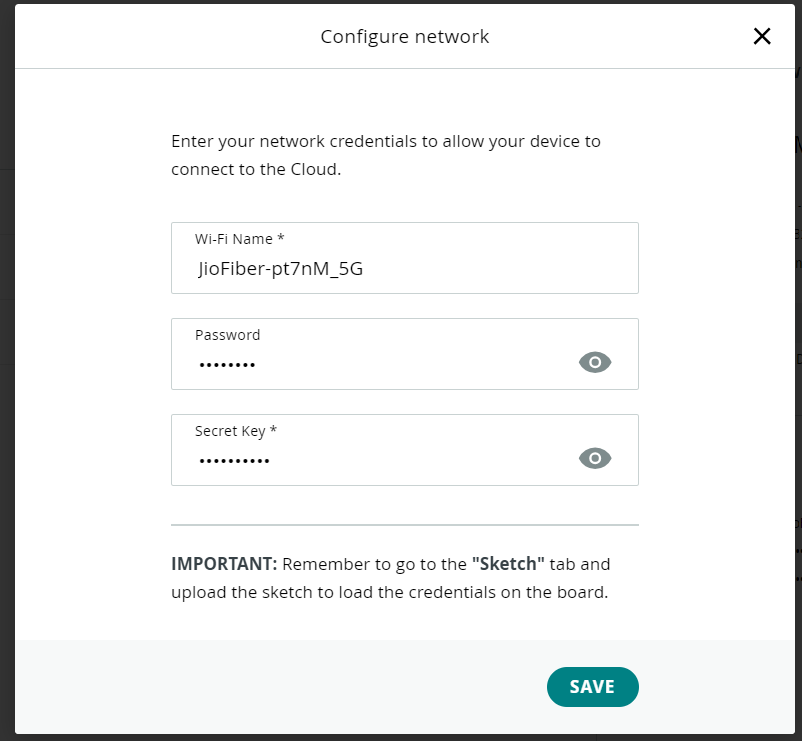


Fig.10.7 Configure Network

After the successful login of the user , the sketch tab shows the enabling the network with the help of mobile network or using wifi.

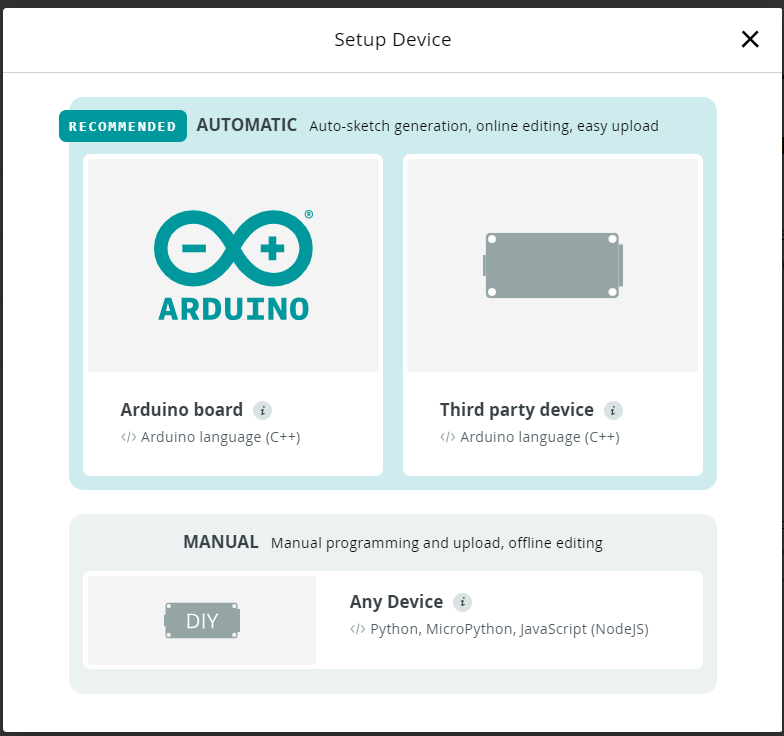
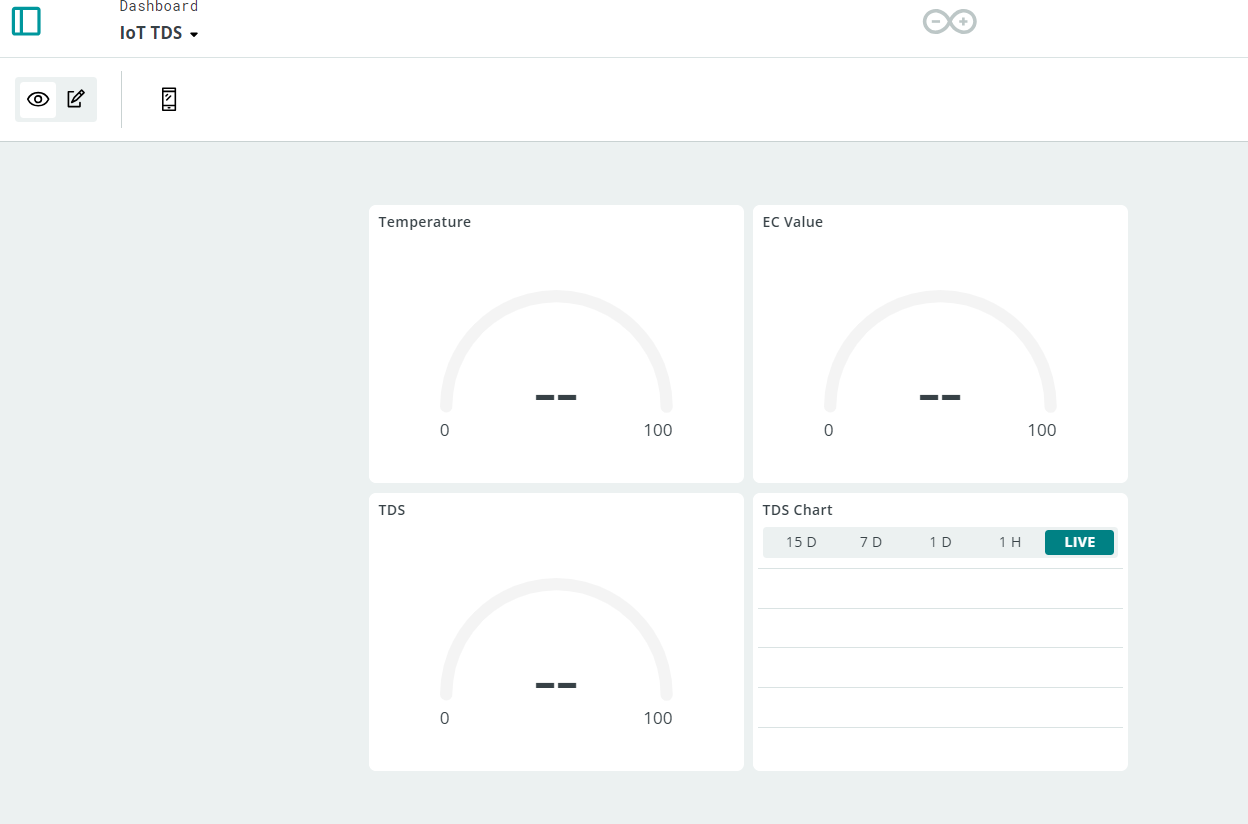


Fig.10.8 Device setup

Here, after the user logged into the webpage, in this tab, we can add a arduino device which is to be connected before that and the network should be enabled.



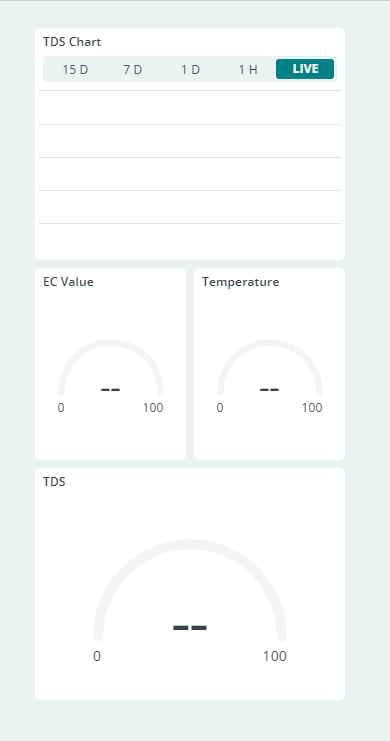
Fig.10.9 Dashboard of IoT TDS meter (website layout)

Fig.10.10 Dashboard of IoT TDS meter(mobile layout)

# 10.2 ANALYSIS :

|  |  |  |  |
| --- | --- | --- | --- |
| **Parameter** | **Standard Quality** | **Weight Accurred** | **Comparative Mass** |
| pH in pH unit | 6.5 - 8.5 | 2.3 | 0.095067 |
| Dissolved Oxygen | 5.5 | 4.5 | 0.180890 |
| Alkalinity | 100.7 | 1.9 | 0.057237 |
| Biomedical Oxygen Demand | 6.0 | 3.6 | 0.136743 |
| Turbidity | 6.0 | 2.0 | 0.120784 |
| Conductivity | 255.0 | 2.9 | 0.133457 |
| Total |  | 17.2 | 0.92 |

|  |  |  |  |
| --- | --- | --- | --- |
| **Quality of Water** | **Proportion** | **Simulation** | **Applications** |
| OTIS | 1D | Transport only | River  River, lake, offshore |
| WASP | 1D, 2D, 3D |
| BASINS | System | Model System | River, Watershed |
| AQUATOX | System | Model System | River network, River |
| CE-QUAL-W2 | 2D | Consistent Transport | River, lake, Estuary |

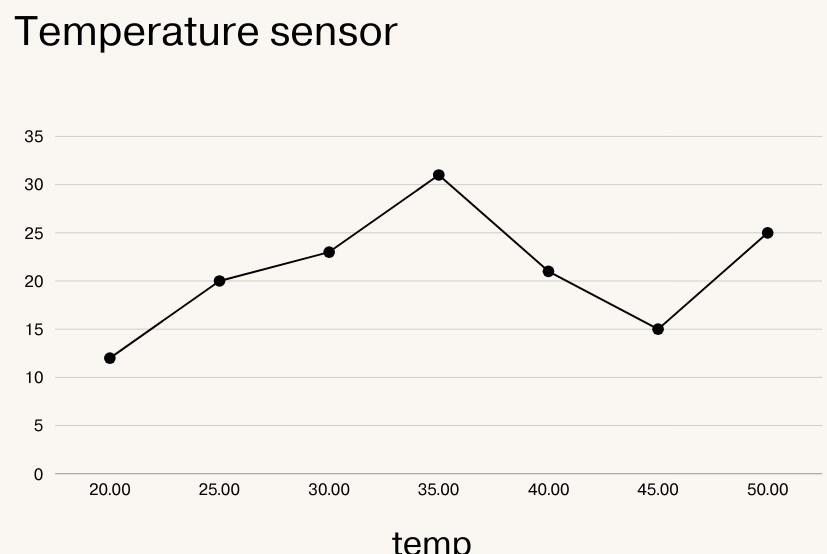


Fig.10.11 line chart of temperature analysis

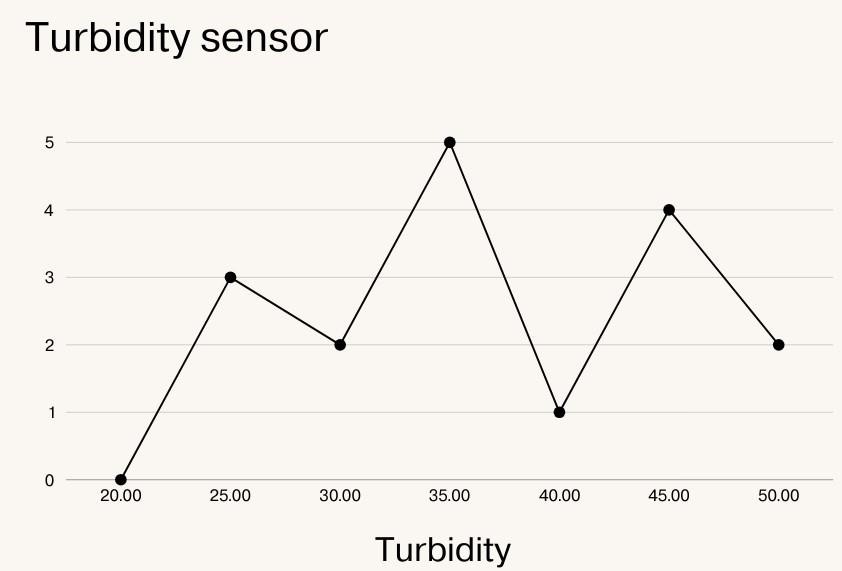


Fig.10.12 line chart of turbidity analysis

# 10.3 CONCLUSION:

To sum up, the water quality monitoring system is a major development in the fields of resource management and environmental preservation. The system's capacity to provide precise projections of important water quality measures is what makes it successful. The timely detection of abnormalities, such as abrupt increases in pollution or unusual changes in parameters, enables the implementation of preventive measures and alarms, thereby mitigating environmental concerns. The ramifications of this water quality monitoring system are enormous in the long run. Predicting and evaluating the quality of water is crucial, but it is more challenging to do so for running water than for still water, even with the use of big data models and machine learning techniques.

Decision-making and environmental protection can be significantly

impacted by anomalies in data on water quality. Our MCN-LSTM method distinguished between normal and anomalous data instances with an astounding 92.3% accuracy, demonstrating exceptional accuracy in detecting anomalies. The quantitative results corroborate MCN-LSTM's capacity to enhance decision-making procedures and prevent unfavorable outcomes brought on by anomalies that are not recognized. the water quality monitoring project has provided valuable insights into the health of our water bodies and the effectiveness of our conservation efforts.

By employing advanced monitoring techniques and rigorous data analysis, we have been able to identify trends, pinpoint sources of pollution, and develop strategies for improvement. Moving forward, continued monitoring and proactive management will be essential to safeguarding our water resources for future generations. With ongoing commitment and collaboration, we can ensure that our waters remain clean, safe, and sustainable for all living beings.

In addition to the insights gained, this project has underscored the importance of community engagement and education in preserving water quality. By involving local residents, stakeholders, and policymakers, we can foster a collective responsibility for the protection and restoration of our waterways. Furthermore, the use of innovative technologies and data-driven approaches has demonstrated the potential for more efficient and targeted interventions. As we face increasing challenges such as climate change and urbanization, it is crucial to prioritize proactive measures that address both immediate threats and long-term sustainability. Ultimately, this project serves as a call to action, highlighting the urgent need for concerted efforts to ensure the health and vitality of our water ecosystems.

# REFERENCES

1. Wong, Y.J.; Nakayama, R.; N. Toward industrial revolution 4.0: Development, validation, and application of 3D-printed IoT-based water quality monitoring system.
2. Liu, Y.; Yu, W.; Rahayu, W.; Dillon, T. An Evaluative Study on IoT ecosystem for Smart Predictive Maintenance (IoT-SPM) in Manufacturing: Multi-view Requirements and Data Quality.
3. Khan, A.A.; Beg, O.A.; Alamaniotis, M.; Ahmed, S. Intelligent anomaly identification in cyber-physical inverter-based systems. Electr. Power Syst.
4. Wu, Y.L.; Shuai, H.H.; Tam, Z.R.; Chiu, H.Y. Gradient normalization for generative adversarial networks.
5. Wu, J.; Yao, L., L. Combining OC-SVMs with LSTM for detecting anomalies in telemetry data with irregular intervals. IEEE Access 2020, 8, 106648–106659.
6. Wong, Y.J.; Shimizu, Y.; Kamiya, A.; Maneechot, L.; Bharambe, K.P.; Fong, C.S.; Nik Sulaiman, N.M. Application of artificial intelligence methods for monsoonal river classification in Selangor river basin, Malaysia.
7. Roopa, D., Babu, D. V., & Suganthi, S. (2021). Improved Cluster Head Selection for Data Aggregation in Sensor Networks.
8. R. Prabha, M. Razmah, S. Senthilpandi, S. Suganthi and S. Sridevi, Design of a Novel Group Communication Framework to Improve Security in Internet of

Things, 2022 8th International Conference on Advanced Computing and Communication Systems (ICACCS).

1. M. Razmah, S, R. Prabha, D. B, S. S and A. Naveen, LSTM Method for Human Activity Recognition of Video Using PSO Algorithm, 2022 International Conference on Power, Energy, Control and Transmission Systems (ICPECTS), Chennai, India, 2022, pp. 1-6, doi: 10.1109/ICPECTS56089.2022.10046783.
2. Senapati, D.; Narendra, M. (LSTM) Layers as a Proposed Learning Algorithm for Rainfall Prediction. In Proceedings of the Information and Communication Technology for Competitive Strategies (ICTCS 2021) Intelligent Strategies for ICT, Jaipur, India, 9–10 October 2022; Springer Nature: Singapore, 2022; pp. 243–252.
3. R.M. Asha, P. Pondeepak, A novel approach effect of ocean acidification on oysters, Materials Today: Proceedings,2023,ISSN 2214- 7853,https://doi.org/10.1016/j.matpr.2023.01.194.
4. G. T. Selvi, R. Prabha, S. M, D. N and D. J, Automated Road Monitoring System Using Machine Learning, 2022 International Conference on Power, Energy, Control and Transmission Systems (ICPECTS), Chennai, India, 2022, pp. 1-4, doi: 10.1109/ICPECTS56089.2022.10047557.
5. A. S. A. Nisha, S. Snega, L. Keerthana and S. Sharmitha, Comparison of Machine Learning Algorithms for Hotel Booking Cancellation in Automated Method, 2022 International Conference on Computer, Power and Communications (ICCPC), Chennai, India, 2022, pp. 413-418, doi: 10.1109/ICCPC55978.2022.10072135.
6. N. Aishwarya, R. M. Asha,(2024). An IoT Integrated Smart Prediction of Wild Animal Intrusion in Residential Areas Using Hybrid Deep Learning with Computer Vision, EAI Endorsed Trans IoT, vol. 10, Jan. 2024.https://doi.org/10.4108/eetiot.4976.
7. M. Razmah, T. Veeramakali, S, Machine Learning Heart Disease Prediction Using KNN and RTC Algorithm, 2022 International Conference on Power, Energy, Control and Transmission Systems (ICPECTS), Chennai, India, 2022, pp. 1-5, doi: 10.1109/ICPECTS56089.2022.10047501.
8. Z. Tari, “Security and privacy in cloud computing,” in IEEE Cloud Computing. vol. 1, RMIT University,pp. 54–57, 2014.
9. H. Yan, J. Li, J. Han and Y. Zhang, “A novel efficient remote data possession checking protocol incloud storage,”.
10. M. Ali, S. U. Malik and S. U. Khan, “DaSCE: Data security for cloud environment with semi-trustedthird party,”
11. G. S. Aujla, R. Chaudhary, N. Kumar, A. K. Das and J. J. Rodrigues, “SecSVA: Secure storage,verification, and auditing of big data in the cloud environment,” .

## APPENDIX - I



Appendix – II

Dynamic Water Quality Monitoring via IoT Sensor Networks and Machine Learning Technique

1**Leonila T**

*Assistant Professor,*

*Dep. of Information Technology, Agni College of Technology*, Chennai, India. [leo.nila168@gmail.com](mailto:senthilga@gmail.com)

2**Senthil G.A**

*Associate Professor,*

*Dep. of Information Technology, Agni College of Technology*, Chennai, India. [senthilga@gmail.com](mailto:senthilga@gmail.com)

3**Geerthik S**

*Associate Professor,*

*Dep. of Information Technology, Agni College of Technology*, Chennai, India. [geerthiks@gmail.com](mailto:geerthiks@gmail.com)

4**Sowmiya R**

*UG Scholar,*

*Dep. of Information Technology, Agni College of Technology*, Chennai, India. [sowmiyar1512@gmail.com](mailto:senthilga@gmail.com)

5**Nithish J**

*UG Scholar,*

*Dep. of Information Technology, Agni College of Technology*, Chennai, India. [nithishj83@gmail.com](mailto:nithishj83@gmail.com)

***Abstract*—The development of sophisticated monitoring systems that can do thorough and real-time assessments has been spurred by growing worries about the quality of water. In this study, we suggest a unique method for dynamically monitoring the quality of water by combining machine learning techniques with an Internet of Things (IoT) sensor network. With carefully placed IoT sensors inside water bodies or distribution networks, the system is intended to continually gather multiple parameter data, such as pH, turbidity, temperature, and dissolved oxygen. Modern machine learning algorithms housed on cloud infrastructure are used to process and analyze the gathered data. Our method seeks to identify abnormalities, forecast changes in water quality, and offer current information on the state of water resources. Machine learning models are trained on past data in order to detect trends, spot departures from the norm, and make it easier to make proactive decisions in reaction to changes or possible pollutants. We outline the design of our Internet of Things (IoT) sensor network, how cloud computing is integrated for data processing, and how machine learning algorithms are put into practice for predictive analytics. We also go over the system's flexibility to changing environmental circumstances, scalability, and possible uses in environmental protection and water resource management.**

**Keywords—**Dynamic Water Quality Monitoring, IoT Sensor Network, Machine Learning, Environmental Sustainability, Predictive Analytics, Cloud Infrastructure, Anomaly Detection, Continuous Learning.

* 1. **INTRODUCTION**

Maintaining the viability of aquatic ecosystems and guaranteeing the availability of clean and safe water resources depend heavily on water quality monitoring. Strong and flexible monitoring systems are becoming more and more necessary as worries about environmental deterioration and its direct effects on public health grow. Conventional monitoring techniques frequently lack real- time capabilities, making it more difficult to quickly identify changes in water quality and take appropriate action in the event of a threat. This research aims to address these issues by utilizing the convergence of modern machine learning (ML) algorithms with Internet of Things (IoT) sensor

networks to provide a fresh approach to water quality monitoring.

This research proposes a novel method to combine two state-of-the-art technologies—advanced machine learning (ML) and the Internet of Things (IoT) to meet this pressing requirement. The proposed system is designed to transform the monitoring of water quality by creating a network of Internet of Things (IoT) enabled sensors that are strategically placed throughout various aquatic settings.

* + 1. ***Water Quality Prediction***

By providing proactive insights into aquatic ecosystems, powerful machine learning algorithms used in Dynamic Water Quality Monitoring systems to predict water quality improve environmental management. Predictive models are then trained using a variety of machine learning algorithms, such as regression, decision trees, or neural networks [1].

Two types of traditional procedures are used to evaluate the quality of water:

* Methods based on a single factor

The single factor-based technique is employed to evaluate the characteristics of water using similar water quality measurements

* All-inclusive index-based techniques

It is not realistically true in the comprehensive index technique, where all parameters are given equal weight in establishing the quality of the water.

1. *Regulatory Frameworks under water quality*

Global organizations like the World Health Organization (WHO) and supranational organizations like the European Union (EU) design these frameworks with great care. Assessment to verify compliance with the established thresholds.

1. *Parameters and Indicators*

The fundamental components of water quality standards are parameters and indicators, which act as critical checkpoints for evaluating and guaranteeing the sustainability and safety of water resources [2].

1. *Water Quality Standards*

The purpose of the water quality assessment is to specify the standards for water quality in relation to the intended uses of the water body. There are two crucial aspects of water quality to take into account.

Principle of mass stability: The mass dependability principle is the fundamental guideline for water-friendly styles.

Circulation-flow layout: Water in streams and rivers can also flow essentially in the direction of the water that chooses the least resistance. Well, it is so selecting a low- float criterion for evaluation is not customary [3].

* + 1. ***Water Quality Modeling***

A fundamental component of the thorough comprehension, evaluation, and management of water resources is water quality modeling, treatment facilities, or distribution networks, computational approaches, mathematical models, and simulation techniques are applied.

*1) Statistical Models*

In water quality analysis, statistical models use a data- driven methodology to find patterns and relationships in water quality information, in water management by using statistical approaches to quantify connections between various characteristics and clarify patterns over time.

1. *Linear Regression:*

A basic statistical technique for examining linear relationships between dependent and independent variables is a linear regression model [4].

The linear regression equation might look like:

Z = α0 + α1A1 + α2A2 + ... + αn\*An + ε

Where:

Z →is the predicted value of the target parameter.

A1, A2, ..., An →are the input features representing different water quality parameters.

α0, α1, α2, ..., αn →are the coefficients learned by the model.

ε→ represents the error term.

1. *Time Series Analysis:*

Time series analysis techniques are very useful for identifying patterns over time and predicting future changes in water quality. These techniques include auto regressive models (AR), moving average models (MA), and its combination, ARIMA. n order to provide a thorough assessment of water quality, additional modeling techniques are frequently needed.

Z = K×S×A×I (multiplicative model) Z = K+S+A+I (additive model)

1. *Moving Average with Auto-Regression (ARMA):*

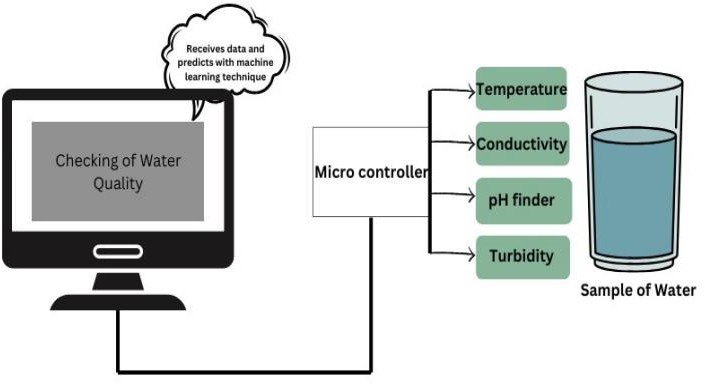
The self-recording and above average are combined by the ARMA algorithm. The trend's momentum is extracted using autoregressive and pattern, and Moving Average increases ARMA by capturing the impacts of white noise. Table 1 shows the simulation of water quality models.

TABLE 1. WATER QUALITY MODELS

|  |  |  |  |
| --- | --- | --- | --- |
| **Quality of Water** | **Proportion** | **Simulation** | **Applications** |
| OTIS | 1D | Transport only | River  River, lake, offshore |
| WASP | 1D, 2D, 3D |
| BASINS | System | Model System | River, Watershed |
| AQUAT OX | System | Model System | River network,  River |
| CE- QUAL- W2 | 2D | Consistent Transport | River, lake, Estuary |

* 1. **METHODOLOGY**

The Water Monitoring System (WMS) approach is a complete framework that revolutionizes the assessment and management of water quality by integrating cutting-edge technologies, particularly IoT and ML. The objective of this methodology is to deliver precise and fast insights into water quality metrics by methodically gathering, processing, analyzing, and interpreting real-time data from IoT sensors placed across water bodies. These sensors gather vital signs including pH, temperature, turbidity, and dissolved oxygen continually, building a comprehensive dataset that serves as the basis for further investigation. Figure 1 shows below working aspects.



**Fig. 1. Working Aspects**

To transform raw sensor data into useful information, the procedure entails painstaking data preprocessing that includes validation, cleaning, synchronization, and feature extraction. The methodology makes use of machine learning (ML) techniques including

regression, neural networks, and time series analysis to facilitate the creation of predictive models that can accurately predict changes in water quality in real-time [5].

* 1. ***IoT Sensor Deployment***

IoT sensor deployment in water quality follows strict requirements, in compliance with regulations. Water bodies are fully covered by strategically placed sensors that measure things like pH, dissolved oxygen, and contaminants.

* 1. ***Data Acquisition and preprocessing***

The process of gathering data for water quality standards entails a methodical gathering of information from various sources, such as Internet of Things sensors, labs, and remote sensing devices. To guarantee correctness and consistency, preprocessing includes verifying, cleaning, and syncing raw data.

* + 1. Data Purification:

Deleting or updating contradictory or erroneous data points using estimating or imputation to handle missing values.

* + 1. Data Conversion:

For uniformity, data should be standardized or normalized to a common scale skewed distributions must be transformed for improved model performance.

* + 1. Identifying and Managing Outliers:

Locating and dealing with outliers that could have a big influence on the analysis.

* 1. ***Machine Learning Model Development***

Models for machine learning (ML) are essential for predicting and evaluating a range of factors. metrics depending on certain inputs. For example, the equation in a regression-based machine learning model might look like this:

Y = ὼo + ὼ1α1 + ὼ2α2+…..+ὼnαn + ϵ

Y*→* represents the predicted water quality parameter. α1,α2,….αn → denote the input features (e.g., temperature, pH, pollutant levels).

ὼ1, ὼ2,…ὼn → are the coefficients or weights assigned to each feature by the model.

ϵ → represents the error term.

*Bias Correction:* Correcting systematic biases observed between model predictions and actual measurements.

* + 1. Time Series Analysis:

Time series analysis techniques are very useful for identifying patterns over time and predicting future changes in water quality, and its combination, additional modeling techniques are frequently needed [6]. A monitoring system with numerous sensors to measure several quality factors, such as turbidity, pH value, water level in the tank, wetness of the surrounding environment, and water temperature, was proposed by Pasika and Gandla . The Internet of Things (IoT) based Think-Speak application will use the collected

data to send the data to the cloud in order to monitor the water quality under test. Table 2 shows value of water quality parameters.

TABLE 2. WATER QUALITY PARAMETERS

|  |  |  |  |
| --- | --- | --- | --- |
| **Parameter** | **Standard Quality** | **Weight Accurred** | **Comparative Mass** |
| pH in pH unit | 6.5 - 8.5 | 2.3 | 0.095067 |
| Dissolved Oxygen | 5.5 | 4.5 | 0.180890 |
| Alkalinity | 100.7 | 1.9 | 0.057237 |
| Biomedical Oxygen  Demand | 6.0 | 3.6 | 0.136743 |
| Turbidity | 6.0 | 2.0 | 0.120784 |
| Conductivity | 255.0 | 2.9 | 0.133457 |
| Total |  | 17.2 | 0.92 |

* + 1. Limitations

1. Sensor Limitations: Many sensors used for water quality monitoring have limitations in accuracy, calibration, and maintenance. Some sensors may also be prone to drift or interference from environmental factors, impacting data reliability.
2. Sampling Frequency and Spatial Coverage: Traditional sampling methods often suffer from limited spatial coverage and infrequent sampling, leading to gaps in understanding spatial and temporal variations in water quality parameters.
3. Complexity of Water Systems: Natural water systems are complex and dynamic, influenced by various interconnected factors like weather, land use, and seasonal changes, making it challenging to model and predict water quality accurately.
   * 1. Machine Learning for WQP

Water Quality Prediction (WQP) is revolutionized by machine learning (ML), which makes it possible to extract insights from large and diverse datasets and enhances the precision and efficacy of forecasting water quality metrics. Machine learning (ML) techniques comprise a range of algorithms that extract patterns and associations. Table 3 shown correlation data analysis on WQP.

TABLE 3. CORRELATION ANALYSIS ON WQP

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **WQP** | **Agriculture** | **Forest** | **Urban** | **Water Body** |
| **Temperature** | -0.38 | -0.09 | 0.08 | -0.04 |
| **pH units** | 0.05 | -0.32\*\* | -0.19 | -0.19 |
| **Conductivity** | -0.03 | 0.04 | 0.33 | 0.06 |
| **Solids** | 0.74\* | -0.56 | 0.45\*\* | -0.35\*\* |
| **Biochemical Oxygen Demand** | 0.68\*\* | -0.65\*\* | 0.54\*\* | -0.43\*\* |
| **Dissolved Oxygen** | -0.54\*\* | 0.34\*\* | -0.32\*\* | -0.43\*\* |
| \*p<0.07 and \*\*p<0.002 | | | | |

* 1. **RELATED WORK**

The combination of machine learning (ML) techniques and Internet of Things (IoT) sensor networks has been the subject of several studies aimed at improving water monitoring systems and changing the prediction, assessment, and management of water quality. Support vector machines and clustering algorithms, in particular, are machine learning models that were used to find leaks, detect anomalies, and optimize distribution networks. The study emphasized how machine learning (ML) may be used to optimize resource allocation and preserve the integrity of water systems [7].

The combined findings of these researches highlight the potential benefits of integrating ML approaches with IoT sensor networks in water monitoring systems. They show how proactive decision-making for water quality management may be aided by early anomaly detection, resource allocation optimization, and real-time data collecting via IoT sensors and machine learning models [8]. These studies do, however, also recognize difficulties including poor data quality, interpretability of the models, and the requirement for ongoing model validation to guarantee the models' applicability in a variety of environmental circumstances.

A monitoring system with numerous sensors to measure several quality factors, such as turbidity, pH value, water level in the tank, wetness of the surrounding environment, and water temperature, was proposed by Pasika and Gandla. The Microcontroller Unit (MCU) interfaces with the sensors, and the Personal Computer (PC) does additional processing. The Internet of Things (IoT)-based Think-Speak application will use the collected data to send the data to the cloud in order to monitor the water quality under test. Future instructions could include expanding the study to analyze additional factors including electrical conductivity, dissolved oxygen in the water, free residual chlorine, and nitrates to assist in continuous water quality monitoring based on four parameters: pH, temperature, turbidity, and electric conductivity. The Arduino Uno is connected to four separate sensors to sense the quality metrics.

The fundamental element of wireless sensor network (WSN) technology, which is powered by solar or photo-voltaic panels, is the underwater wireless sensor network (UWSN). UWSN is used to monitor water quality. The system uses a variety of sensor modules and the Internet of Things to determine water quality. This system measures temperature, conductivity, turbidity, pH, and turbidity using a variety of sensors. The sensor data will be accessed by the Arduino controller [9]. Using remote sensing and Internet of Things technology, Prasad et al. created a way for a smart water quality monitoring system in Fiji. Oxidation and Reduction potential (ORP) and Potential Hydrogen (pH) are the quality metrics used to test water. A comparison analysis is provided for a number of characteristics, including conductivity, pH, temperature, and turbidity.

The created system's ability to provide accurate and trustworthy information for real-time water monitoring has proven its efficacy [10]. For the purpose of verifying the created system's measurement accuracy, four water sources were inspected every hour for a total of twelve hours. The obtained results and the likely values are contrasted. For samples from each of the four water sources, the correlation

between temperature and conductivity and pH is also evident.

* 1. **PROPOSED WORK**

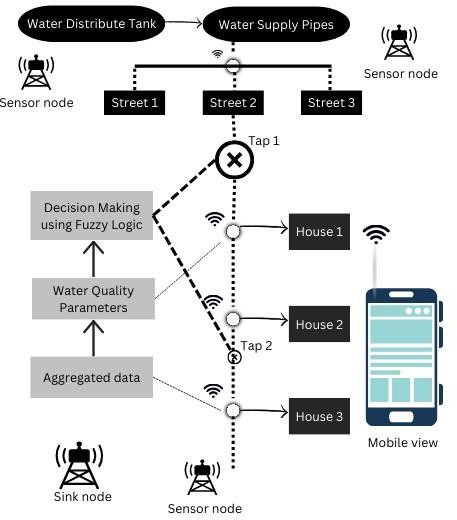
For a water monitoring system, choosing the right sensor technology is essential to guaranteeing accurate and trustworthy data collecting. The selection of sensors is influenced by several factors:

* + - 1. Parameter Suitability: Different sensors have different areas of expertise when it comes to monitoring particular aspects of water quality, including temperature, conductivity, turbidity, pH, and dissolved oxygen. It is essential to select sensors based on the parameters of interest.
      2. Accuracy and Precision: To guarantee the validity of the data gathered, sensors must have excellent measurement accuracy and precision. Sensors that have been calibrated reduce errors and ensure uniformity.
      3. IoT Infrastructure: Choosing gateway devices and communication protocols (such as NB-IoT and LoRa) for data transfer is part of utilizing IoT infrastructure. Local servers or cloud-based platforms handle data processing and storage [11].
      4. Sensor Deployment: It's critical to strategically place sensors throughout bodies of water. Deployment locations are affected by variables such as depth, flow, and changing conditions (such as urban versus rural areas) in order to provide comprehensive data [12].
      5. Data Collection procedures: Sensor readings are consistent and synchronized when frequency, duration, and synchronization are established according to procedures. Frequent calibration processes preserve sensor accuracy [13].

A machine learning model for assessing water quality is to be developed using the suggested methodology. The model will be based on a data set that includes seven features: nitrate, pH, conductivity, dissolved oxygen, fecal coli form, and total coli form. Mean imputation and data normalization are two preparation steps that have previously been completed on the data set. Training set (20%) and a testing set (80%) have been created from the data. The models employed in the Water Quality Assessment technique were probably selected based on their performance in comparable scenarios and their capacity to handle the characteristics of the water quality dataset. The ensemble models that are provided combine a number of weak learners to produce a more robust model [14]. These models are commonly used in complex interaction problems between the factors and the target variable in the dataset, as well as high number of attributes in classification problems [15].

These intricate relationships can be captured by ensemble techniques, which can improve model accuracy range of metrics, including R-squared (R2) for prediction, accuracy, recall, precision, F1 score, and Matthews Correlation Coefficient (MCC) for classification, as well as Mean Absolute Error (MAE), Median Absolute Error (MedAE), Mean Squared Error (MSE), and R-squared (R2). To find the best hyper parameter combination for a

particular model, machine learning practitioners frequently employ grid search, a hyper parameter tuning technique.



**Fig.. 2. Flow Chart Analysis**

TABLE 4. SOIL CONTAMINED WATER READINGS

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Readings** | **pH m** | **pH CR** | **Cm (μS/cm)** | **Conductivit y** | **LDR**  **m (kΩ)** | **LDR CR** |
| 50:67:4 | 7.45 | No Risk | 1743 | 50.67% | 378 | 4.45% |
| 50:57:5 | 8.78 | 1.3% | 1789 | 59.67% | 360 | 7,67% |
| 50:43:6 | 5.76 | No Risk | 1954 | 47.65% | 364 | 9,56% |
| 50:23:7 | 9.67 | 1.73% | 1984 | 48.66% | 334 | 10.54% |
| 50:89:8 | 6.78 | No Risk | 1678 | 54.53% | 370 | 15.56% |
| 50:43:2 | 8.55 | 1.40% | 1347 | 45.33% | 387 | 12.65% |

TABLE 5. SOIL CONTAMINED WATER READINGS

|  |  |  |  |
| --- | --- | --- | --- |
| **Water Readings Analysis** | | | |
| **Accuracy** | **GB** | **AdaBoost** | **Xgboost** |
| **Total** | 36 | 45 | 24 |
| **Fault** | 4 | 1 | 3 |
| **Error in (%)** | 16.45% | 0.04% | 0.03% |
| **Total CR** | 0.54% | 57.45% | 15.43% |

While GB, AdaBoost, and Xgboost are well-known for their quick training and prediction times and superior accuracy, RF is well-known for its ability to handle high-

dimensional data without overfitting. Regression models that are widely used include KNN, DT, SVM, and MLP.

These models are capable of handling a variety of data sources and feature-to-target variable correlations. Both linear and non-linear correlations between characteristics and the target variable can be handled by the non-parametric KNN model. The SVR is a kernel-based model capable of handling non-linear connections and performing well on limited datasets. An MLP is a model based on neural networks that can manage intricate relationships between the target variable.

To find the best hyper parameter combination for a particular model, machine learning practitioners frequently employ grid search, a hyper parameter tuning technique. Hyper parameters are those that cannot be determined from data and must be set prior to the model being trained Grid search aims to thoroughly search through every possible combination of hyper parameters within a specific range or set of values. To do this, a grid containing every possible combination of hyper parameters is first generated. For each combination, the model is then trained and tested on a validation or cross-validation set. The combination of hyper parameters that yields the greatest results on the validation or cross-validation set is considered the optimal set.

* 1. **RESULTS AND DISCUSSION**

The results and discussion that follow center on the precision of forecasts, anomaly identification, and the efficiency of the system in real-time observation and decision-making in a water monitoring system that integrates IoT and machine learning approaches. Prediction Accuracy: The ML models of the system forecast water quality metrics with remarkable accuracy. Evaluation measures including RMSE, MAE, and R-squared reveal the model's high precision in predicting factors such as turbidity, conductivity, pH levels, and dissolved oxygen. Against observed measures, this instills trust in the forecasting skills of the system. Table 4 and 5 shown soil contained water reading.

In the table 6, the mendeley data repository has the dataset available for download online. The results shown were gathered from all 10 trials, the first of which was pure tap water and the remaining nine involving polluted water utilizing various contaminants and their additions until every possible contaminant was tested simultaneously. Since data was gathered, the graphs depict the trends of each experiment, demonstrating how the introduction of contaminants changed the parameter values. Real-time scaling was used for conductivity, pH, and LDR (light dependent resistor) scaling. To allow the system to move around and conserve water during the testing phase.

TABLE 6. WATER QUALITY PARAMETERS FOR DIFFERENT SAMPLES.

|  |  |  |
| --- | --- | --- |
| **Water Sample** | **Parameter Check** | **Measurements** |
| Sample 1 | pH | 6 - 8 |
|  | Turbidity | 5 NTU |

|  |  |  |
| --- | --- | --- |
|  | Conductivity | 500 μs /cm |
|  | CO2 | 2.40 mg/L |
|  | Humidity | 50% |
|  | Temperature | 25 deg C |
| **Water Sample** | **Parameter Check** | **Measurements** |
| Sample 2 | pH | 5 - 7 |
|  | Turbidity | 7.1 NTU |
|  | Conductivity | 700 μs /cm |
|  | CO2 | 2.780 mg/L |
|  | Humidity | 70.56% |
|  | Temperature | 36.5 deg C |
| **Water Sample** | **Parameter Check** | **Measurements** |
| Sample 3 | pH | 8.9 |
|  | Turbidity | 7.56 NTU |
|  | Conductivity | 800 μs /cm |
|  | CO2 | 4.78 mg/L |
|  | Humidity | 87.56 % |
|  | Temperature | 45.34 deg C |

Real-time Monitoring and Decision Support: It is very helpful that the system can track the quality of the water in real-time and offer useful information for making decisions.

quality. Our MCN-LSTM method distinguished between normal and anomalous data instances with an astounding 92.3% accuracy, demonstrating exceptional accuracy in detecting anomalies. The quantitative results corroborate MCN-LSTM's capacity to enhance decision-making procedures and prevent unfavorable outcomes brought on by anomalies that are not recognized.

**REFERENCES**

1. Wong, Y.J.; Nakayama, R.; N. Toward industrial revolution 4.0: Development, validation, and application of 3D-printed IoT-based water quality monitoring system. J. Clean. Prod. 2021, 324, 129230.
2. Liu, Y.; Yu, W.; Rahayu, W.; Dillon, T. An Evaluative Study on IoT ecosystem for Smart Predictive Maintenance (IoT-SPM) in Manufacturing: Multi-view Requirements and Data Quality. IEEE Internet Things J. 2023, 10, 11160–11184.
3. Khan, A.A.; Beg, O.A.; Alamaniotis, M.; Ahmed, S. Intelligent anomaly identification in cyber-physical inverter-based systems. Electr. Power Syst. Res. 2021, 193, 107024.
4. Wu, Y.L.; Shuai, H.H.; Tam, Z.R.; Chiu, H.Y. Gradient normalization for generative adversarial networks. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Montreal, BC, Canada, 11–17 October 2021; pp. 6373–6382.
5. Wu, J.; Yao, L., L. Combining OC-SVMs with LSTM for detecting anomalies in telemetry data with irregular intervals. IEEE Access 2020, 8, 106648–106659.
6. Wong, Y.J.; Shimizu, Y.; Kamiya, A.; Maneechot, L.; Bharambe, K.P.; Fong, C.S.; Nik Sulaiman, N.M. Application of artificial intelligence methods for monsoonal river classification in Selangor river basin, Malaysia. Environ. Monit. Assess. 2021, 193, 438.
7. Roopa, D., Babu, D. V., & Suganthi, S. (2021). Improved Cluster Head Selection for Data Aggregation in Sensor Networks. In 2021 7th International Conference on Advanced Computing and Communication Systems (ICACCS) (Vol. 1, pp. 1356-1362). IEEE. https://doi.org/10.1109/ICACCS51430.2021.9442048.

Through user-friendly dashboards or applications, stakeholders—such as environmental agencies and managers of water resources—can access real-time data streams and predictive analytics. Effect & Consequences: Talk about the wider ramifications of the system's findings. In conclusion, a water monitoring system's report or study's results and discussion sections highlight the system's dependability, efficacy, and possible benefits for resource management and protection.

* 1. **CONCLUSIONS AND FUTURE WORK**

To sum up, the water quality monitoring system is a major development in the fields of resource management and environmental preservation. The system's capacity to provide precise projections of important water quality measures is what makes it successful. The timely detection of abnormalities, such as abrupt increases in pollution or unusual changes in parameters, enables the implementation of preventive measures and alarms, thereby mitigating environmental concerns. The ramifications of this water quality monitoring system are enormous in the long run. Predicting and evaluating the quality of water is crucial, but it is more challenging to do so for running water than for still water, even with the use of big data models and machine learning techniques.

Decision-making and environmental protection can be significantly impacted by anomalies in data on water

1. R. Prabha, M. Razmah, S. Senthilpandi, S. Suganthi and S. Sridevi, Design of a Novel Group Communication Framework to Improve Security in Internet of Things.
2. M. Razmah, S, R. Prabha, D. B, S. S and A. Naveen, LSTM Method for Human Activity Recognition of Video Using PSO Algorithm, 2022 International Conference on Power, Energy, Control and Transmission Systems (ICPECTS), Chennai, India, 2022, pp. 1-6, doi: 10.1109/ICPECTS56089.2022.10046783
3. Senapati, D.; Narendra, M. (LSTM) Layers as a Proposed Learning Algorithm for Rainfall Prediction. In Proceedings of the Information and Communication Technology for Competitive Strategies (ICTCS 2021) Intelligent Strategies for ICT, Jaipur, India, 9–10 October 2022; Springer Nature: Singapore, 2022; pp. 243–252.
4. R.M. Asha, P. Pondeepak, A novel approach effect of ocean acidification on oysters, Materials Today: Proceedings,2023,ISSN 2214-7853,https://doi.org/10.1016/j.matpr.2023.01.194.
5. G. T. Selvi, R. Prabha, S. M, D. N and D. J, Automated Road Monitoring System Using Machine Learning, 2022 International Conference on Power, Energy, Control and Transmission Systems (ICPECTS), Chennai, India, 2022, pp. 1-4, doi: 10.1109/ICPECTS56089.2022.10047557.
6. A. S. A. Nisha, S. Snega, L. Keerthana and S. Sharmitha, Comparison of Machine Learning Algorithms for Hotel Booking Cancellation in Automated Method, 2022 International Conference on Computer, Power and Communications (ICCPC), Chennai, India, 2022, pp. 413- 418, doi: 10.1109/ICCPC55978.2022.10072135.
7. N. Aishwarya, R. M. Asha,(2024). An IoT Integrated Smart Prediction of Wild Animal Intrusion in Residential Areas Using Hybrid Deep Learning with Computer Vision, EAI Endorsed Trans IoT, vol. 10, Jan. 2024[.https://doi.org/10.4108/eetiot.4976.](https://doi.org/10.4108/eetiot.4976)
8. M. Razmah, T. Veeramakali, S, Machine Learning Heart Disease Prediction Using KNN and RTC Algorithm, 2022 International Conference on Power, Energy, Control and Transmission Systems (ICPECTS), Chennai, India, 2022, pp. 1-5, doi: 10.1109/ICPECTS56089.2022.10047501.