Real Time Monitoring of Water Quality for Rural Areas: A Machine Learning and Internet of Things Approach

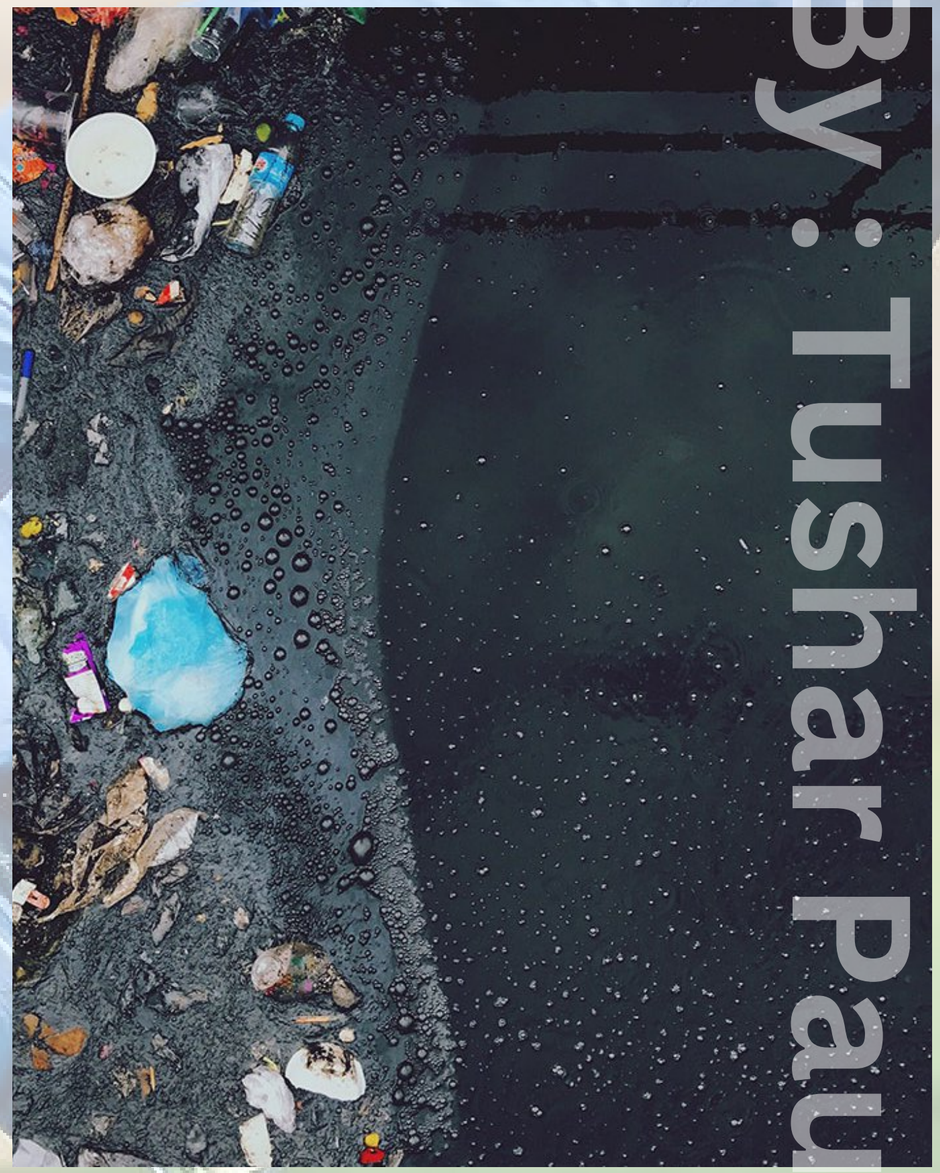
By : Tushar Paul

Email ID : tusharpaul2001@gmail.com

By : Tushar Paul

Introduction

An inventive way to guarantee the public has access to safe drinking water is by utilising IoT and ML in a water quality monitoring system. IoT sensors are used in the system to gather real-time data on water quality and send it to a processing server for analysis. The system can analyse the gathered data to spot potential contaminants, forecast trends in water quality, and even spot abnormalities with the aid of machine learning algorithms. This enables the proactive prevention of water pollution and the maintenance of constant water quality.



By : Tushar Paul

Email ID : tusharpaul2001@gmail.com

Benefits of IoT and ML-based Monitoring System

- **Real-time monitoring-** Continuous data collection and analysis in real-time, provide a timely response to potential water quality hazards.
- **Cost-effective-** Savings on manual labor, laboratory equipment, and transportation costs reduce costs significantly.
- **Increased coverage-** IoT devices and wireless connectivity can extend monitoring capabilities to a wider range of geographical areas, providing a more comprehensive picture of water quality.
- **Early Detection of Water Quality Issues-**ML algorithms can help detect potential water quality concerns such as harmful algal blooms or nutrient pollution before they cause harm.

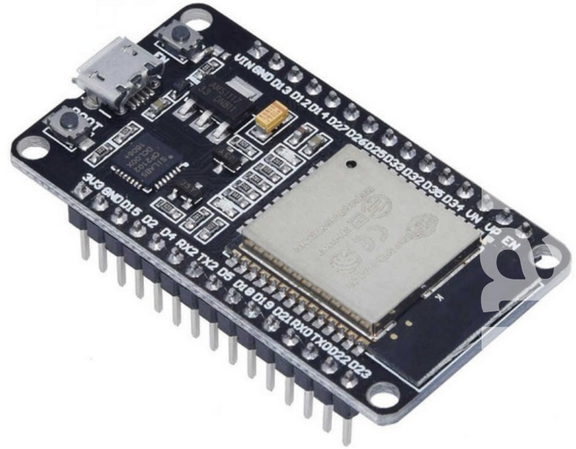
By : TusharPaul

Email ID : tusharpaul2001@gmail.com

IoT COMPONENTS :

- **ESP 32**

  - The ESP32 is a low-cost, low-power system-on-a-chip microcontroller designed by Espressif Systems
  - It is a dual-core 32-bit CPU, Wi-Fi and Bluetooth connectivity, and a rich set of peripherals.
  - The ESP32 features a range of peripherals, including ADCs, DACs, PWM controllers, and communication interfaces , such as UART, SPI, and I2C.



By : Tushar Paul

Email ID : tusharpaul2001@gmail.com

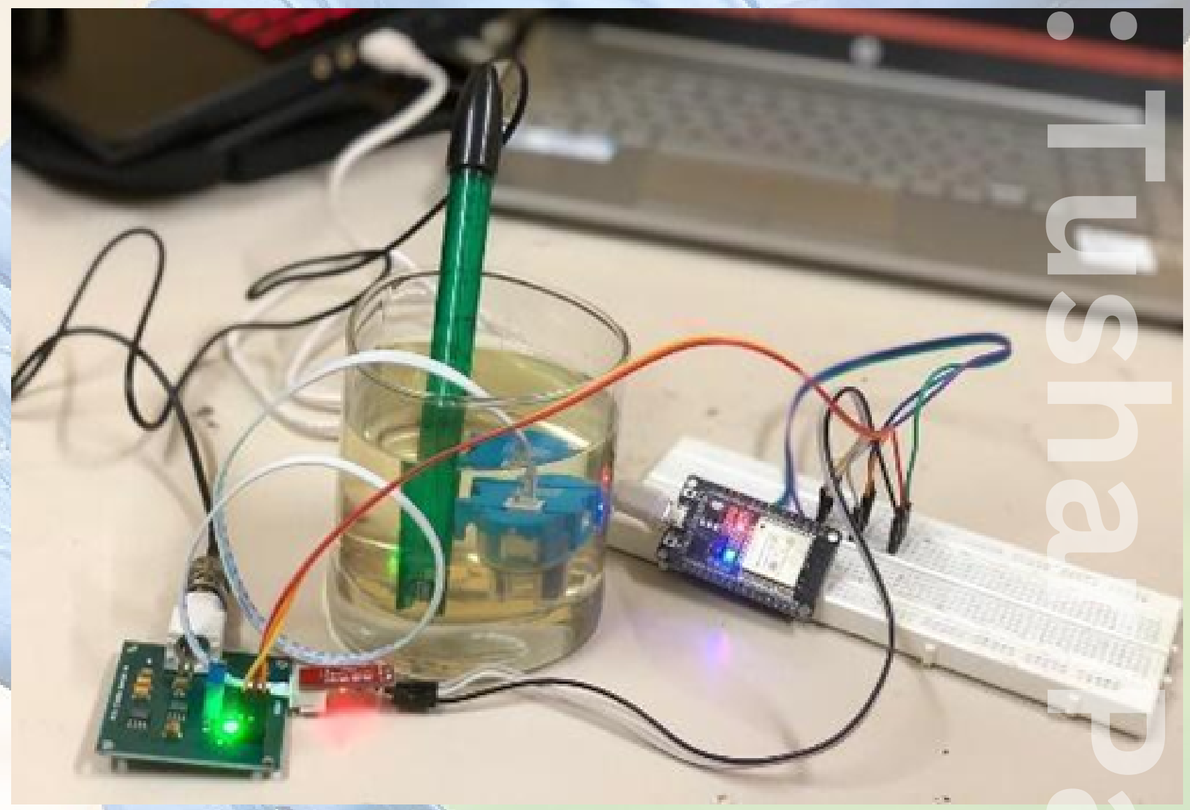
- **pH Sensor**
  - A pH sensor measures the acidity or alkalinity of a solution, commonly referred to as its pH value.
  - A pH sensor is based on the interaction of hydrogen ions with a glass membrane sensing electrode that is selective to hydrogen ions.

- **Turbidity Sensor**
  - A turbidity sensor measures the amount of suspended particles or solids in a liquid or gas.
  - A turbidity sensor measures the attenuation of light as it passes through a sample containing suspended particles or solids.

Email ID : tusharpaul2001@gmail.com

Our Model

- This model is to predict the potability of water using IoT and ML techniques.
- Different sensors and microcontrollers were used to evaluate the pH and turbidity values of water.
- The model was validated using ML techniques.
- The project demonstrates the potential of IoT and ML in predicting water quality and ensuring safe drinking water for communities.



By : Tushar Paul

Email ID : tusharpaul2001@gmail.com

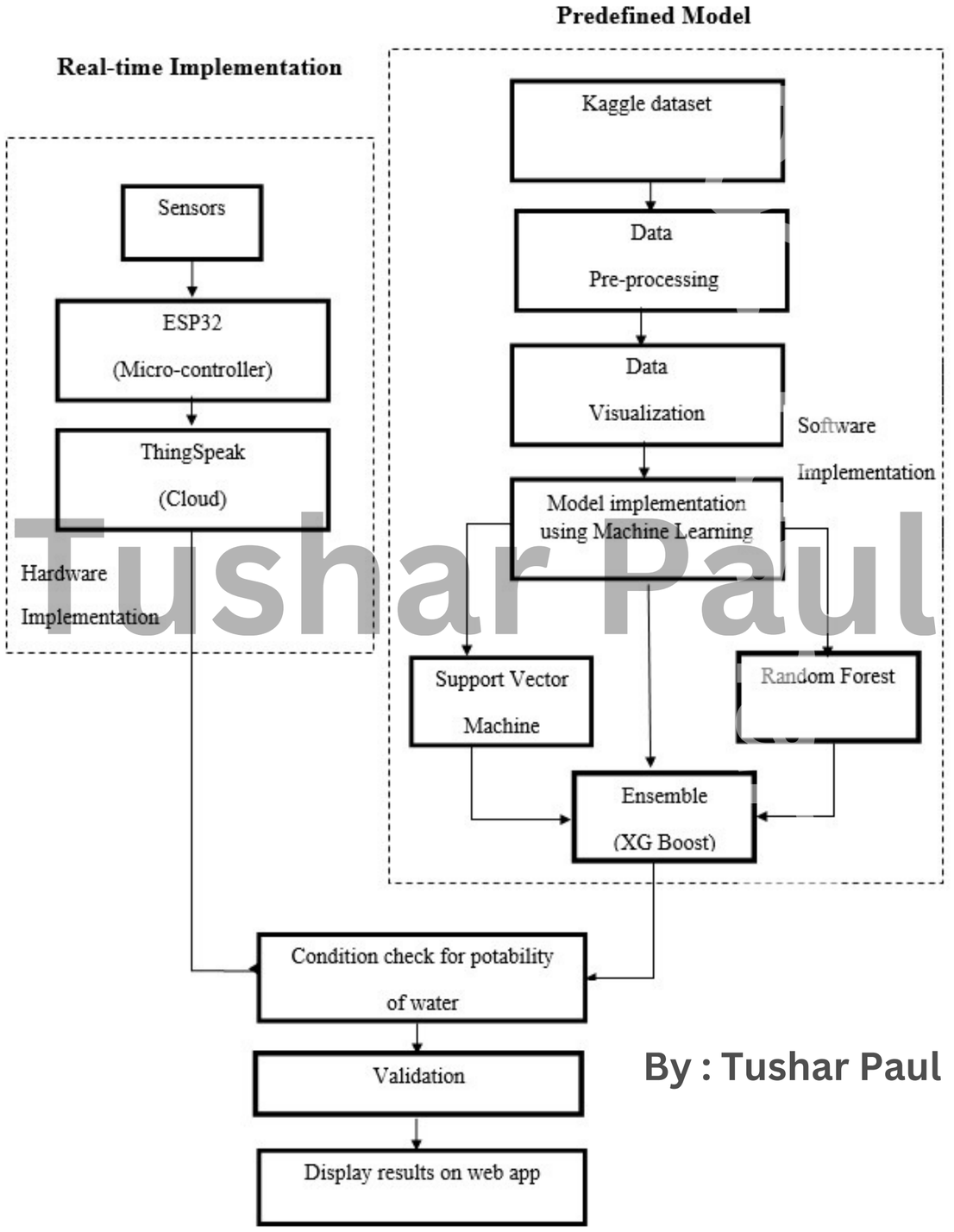
Flowchart

The given figure explains the flow of our work as-

- Mode training and evaluation

.

- Combination of hardware and software components.
- Working of our real-time system.
- Verification using trained model.
- Display on web app.



Tushar Paul

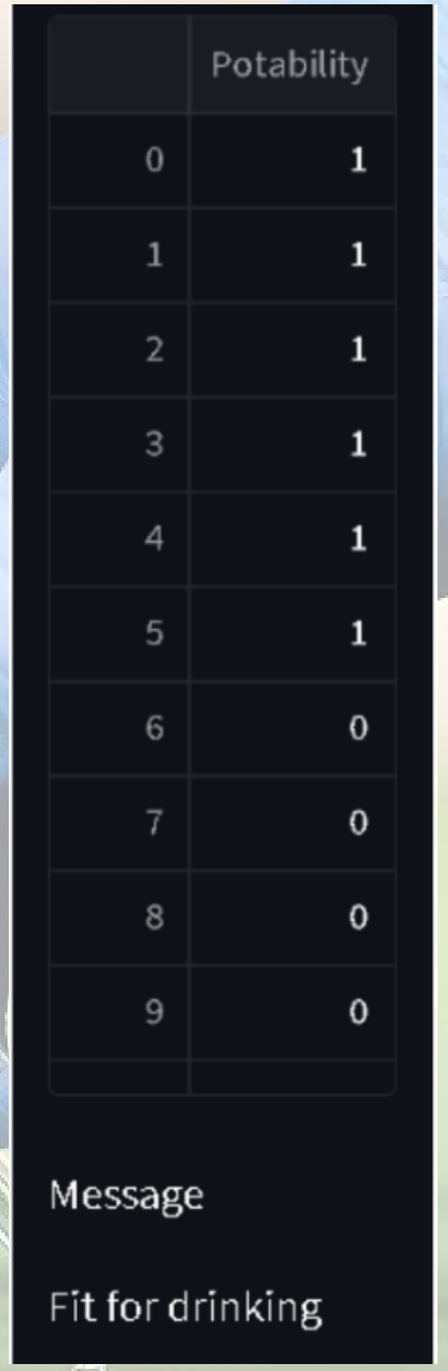
By : Tushar Paul

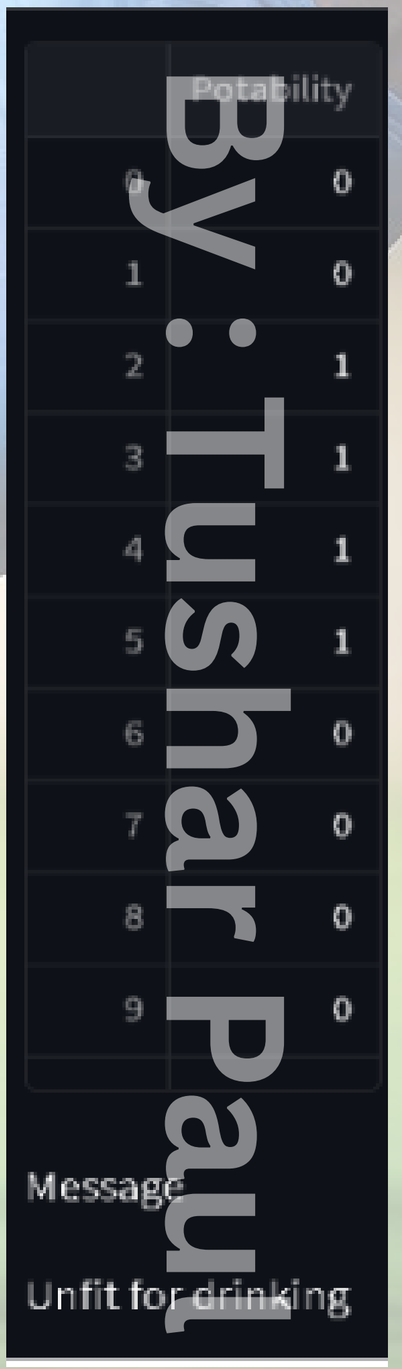
Email ID : tusharpaul2001@gmail.com

Methodology

Our suggested approach's workflow is as follows:

- Model training with SVM, RF, and XGBoost.
- Data captured in real time with the use of sensors.
- Sensor data uploading to ThingsSpeak cloud.
- Information is transferred from ThingsSpeak to our model.
- Preprocess the data that was obtained.
- Data validation.
- Potability levels are predicted using a trained model.
- WebApp displays the predicted values as shown.





By : Tushar Paul

Email ID : tusharpaul2001@gmail.com

Machine Learning Modules

- **sklearn-** Scikit-learn (sklearn) is a popular open-source machine learning library for Python that provides a range of tools for classification, regression, clustering, and dimensionality reduction.
- **Streamlit-** Streamlit is an open-source Python library used for creating web applications for machine learning and data science projects.
- **numpy-** NumPy is a popular open-source Python library for scientific computing that provides support for arrays and matrices, along with a large collection of mathematical functions.
- **pandas-** Pandas is an open-source data analysis and manipulation library for Python that provides fast, flexible, and expressive data structures for working with structured and time-series data.

Email ID : tusharpaul2001@gmail.com

By : Tusharpaul

ML Approach

Our proposed approached has been built using the following models:

- **Support Vector Machine (SVM)** - SVM finds best hyperplane to classify data points which achieved an accuracy of 65.85%.
- **Random Forests (RF)** - Random Forest combines decision trees to improve accuracy and prevent overfitting, which achieved an accuracy of 69.35%
- **XGBoost**- In order to increase the accuracy of our model, we made an ensemble of RF and SVM using the stacking method of XGBoost which gave an accuracy of 95.12%.

After the successful training of our model, we deployed it as an web app using StreamLit library for the real-time monitoring of water quality.

By : Tusharpaul

Email ID : tusharpaul2001@gmail.com

Thank You

By : Tushar Paul

Email ID : tusharpaul2001@gmail.com