IJNRD.ORG ISSN : 2456-4184
INTERNATIONAL JOURNAL OF NOVEL RESEARCH AND DEVELOPMENT (IJNRD) | IJNRD.ORG
An International Open Access, Peer-reviewed, Refereed Journal

MONITORING OF AIR AND WATER QUALITY USING IOT AND MACHINE LEARNING

Mr. S. Bosu Babu [1], K. Ajay Kumar [2], P. Hema [3], N. Likhitha [4], B. Harsha Vardhan[5], T. Santhi Salomy [6]

[1] Assistant Professor, Dept of Computer Science and Engineering, Anil Neerukonda Institute of Technology and Sciences, Visakhapatnam, India

[2,3,4,5,6] Students, Dept of Computer Science and Engineering, Anil Neerukonda Institute of Technology and Sciences, Visakhapatnam, India

ABSTRACT

Ensuring the quality of air and water is crucial to maintain a healthy environment and to prevent the spread of various diseases. With the increasing concerns about environmental pollution, it is necessary to continuously monitor the parameters of air and water quality to avoid any adverse effects on human health and the ecosystem. It provides accurate and timely information on the quality of air and water in a particular location, enabling users to make informed decisions regarding their health and well-being. The system employs multiple sensors and devices to collect environmental data, which is then stored in an Excel sheet by connecting devices through a data streamer. It aims to provide precise and real-time information on parameters such as harmful gasses, particulate matter, and total dissolved salt levels. This uses the Random Forest classifier as a model to predict air and water quality. The proposed IoT-based system provides a cost-effective and scalable solution for monitoring air and water quality in real time, allowing for timely intervention to mitigate pollution and enhance overall public health.

Keywords: Internet of Things, Sensors, Microcontroller

1. INTRODUCTION

The quality of air and water is of utmost importance to our environment and has a significant impact on human health and the ecosystem. Due to the increasing levels of pollution in both air and water, there is a growing need for continuous monitoring to ensure that the environment remains healthy and safe. The development of new sensor capabilities in wireless communication and sensor networks has allowed the emerging technology of the Internet of Things (IoT) to show great potential in addressing this challenge. Industrial areas, in particular, are more susceptible to air pollutants, and the drinking water in such areas is often contaminated and not suitable for consumption. Common air pollutants such as CO, CO_2, NH_3, and CH_4 pose serious health risks by causing diseases like asthma and bronchitis. Similarly, water pollution results from the release of industrial waste, faecal bacteria, and radioactive substances into rivers, lakes, and oceans, causing diseases like cholera, typhoid, and polio.

Compared to traditional methods of monitoring air and water quality, this IoT-based system offers numerous benefits. It provides real-time and accurate information on various parameters, including harmful gas ppm levels, particulate matter in the air, and total dissolved solids levels in the water. The system is scalable and cost-effective, making it suitable for deployment in various locations, ranging from homes to large industrial sites. It offers a scalable and cost-effective solution for collecting and analyzing data in real time, enabling prompt actions to protect the environment and human health.

2. LITERATURE SURVEY

The Internet of Things (IoT) is an unexpectedly developing subject with capacity programs in numerous domains, which include environmental monitoring. Air and water quality monitoring are critical for environmental protection and public health. In recent years, IoT-based systems for monitoring air and water quality have become increasingly popular due to their affordability, ability to provide real-time monitoring, and accuracy.

Pattar Sunil Mahesh et. Al (2018) developed an IoT-based air pollution monitoring system that can display air quality information on a web server using the internet and trigger an alarm when air quality drops below a certain threshold due to the presence of harmful gases such as CO2, smoke, alcohol, benzene, and NH3. The system also displays air quality data in parts per million (PPM) on an LCD screen and website, enabling easy monitoring of air components.

Meghana P. Gowda et.ak (2021) proposed an IoT-based system for monitoring real-time air pollution levels of various pollutants, which enables people to identify which components of the air are being polluted. Our system utilizes a module node MCU esp8266, equipped with a Wi-Fi module that allows remote monitoring of air pollutants and provides continuous updates on air quality.

Santosh and Shankar[4] A Smart Water Quality Monitoring (SWQM) system has been developed using an IoT environment, utilizing a Field Programmable Gate Array (FPGA), sensors, and a Zigbee wireless communication module. The proposed model is designed to measure six different water quality parameters in real-time, including turbidity, pH, humidity, water level, water temperature, and carbon dioxide (CO2). This method has the potential to decrease the time required for determining water quality in water resources.

Omar Faruq[5] developed a water quality monitoring system to cater to the needs of people residing in the outskirts of Bangladesh where safe drinking water is not readily accessible. This device has been engineered with a high degree of precision and sensitivity to detect multiple water parameters, including temperature, turbidity, and hydrogen potential. Additionally, the device is equipped with an LED screen that displays the pH content

3. DATASET USED

The dataset obtained is by connecting the sensors through the microcontroller Arduino Uno with ESP8266 Wifi module. The sensors MQ4 (methane gas sensor), MQ7 (carbon monoxide sensor), MQ135 sensor (detects acetone, carbon dioxide and ammonia), PMS5003 PM2.5 (particulate matter sensor), and Total dissolved salts (water quality sensor) connecting the sensors with the Jumper wires. The data is collected by exploring through different locations and various gas levels are obtained. On the whole, there are 3500 records with 8 attributes including the location.

A1 | Location

	A	B	C	D	E	F	G	H
1	Location	CO	CH4	NH3	acetone	C02	pm2.5	tds
2	Tagarapuv	0.92	49	597	540.95	591	73.24	104
3	Tagarapuv	0.97	49	596	540.95	590	83.13	104
4	Tagarapuv	17.4	49	596	541.86	590	79.84	104
5	Tagarapuv	1.7	49	596	540.95	590	94.52	105
6	Tagarapuv	22.1	49	596	540.95	590	135.99	105
7	Tagarapuv	45.41	49	596	540.95	590	178.33	105
8	Tagarapuv	112.16	49	596	541.86	590	139.7	105
9	Tagarapuv	80.87	49	597	540.95	590	80.65	105
10	Tagarapuv	29.16	48	596	541.86	589	58.36	105
11	Tagarapuvalasa		48	597	541.86	589	79.29	105
12	Tagarapuv	132.07	46	596	540.95	589	88.7	105
13	Tagarapuv	52.04	46	597	540.95	589	74.28	105
14	Tagarapuv	48.82	46	597	541.86	589	113.93	104
15	Tagarapuv	19.2	46	596	540.95	589	105.39	104
16	Tagarapuv	0.6	46	596	541.86	589	66.52	105
17	Tagarapuv	1.63	45	596	541.86	588	65.04	105
18	Tagarapuv	11.44	46	596	541.86	588	103.36	105
19	Tagarapuv	6.1	45	596	541.86	588	177.33	105
20	Tagarapuv	2.51	45	596	541.86	588	113.25	105
21	Tagarapuv	7.92	44	595	541.86	588	99.7	105
22	Tagarapuv	9.52	45	595	541.86	588	80.61	105
23	Tagarapuv	9.05	46	597	541.86	588	100.79	105
24	Tagarapuv	22.53	45	8	541.86	588	107.99	105
25	Tagarapuv	2.03	45	638	541.86	588	140.29	105
26	Tagarapuv	1.42	45	638	541.86	586	239.96	104
27	Tagarapuv	2.27	46	637	540.95	586	142.96	105

3.1 SYSTEM ARCHITECTURE:

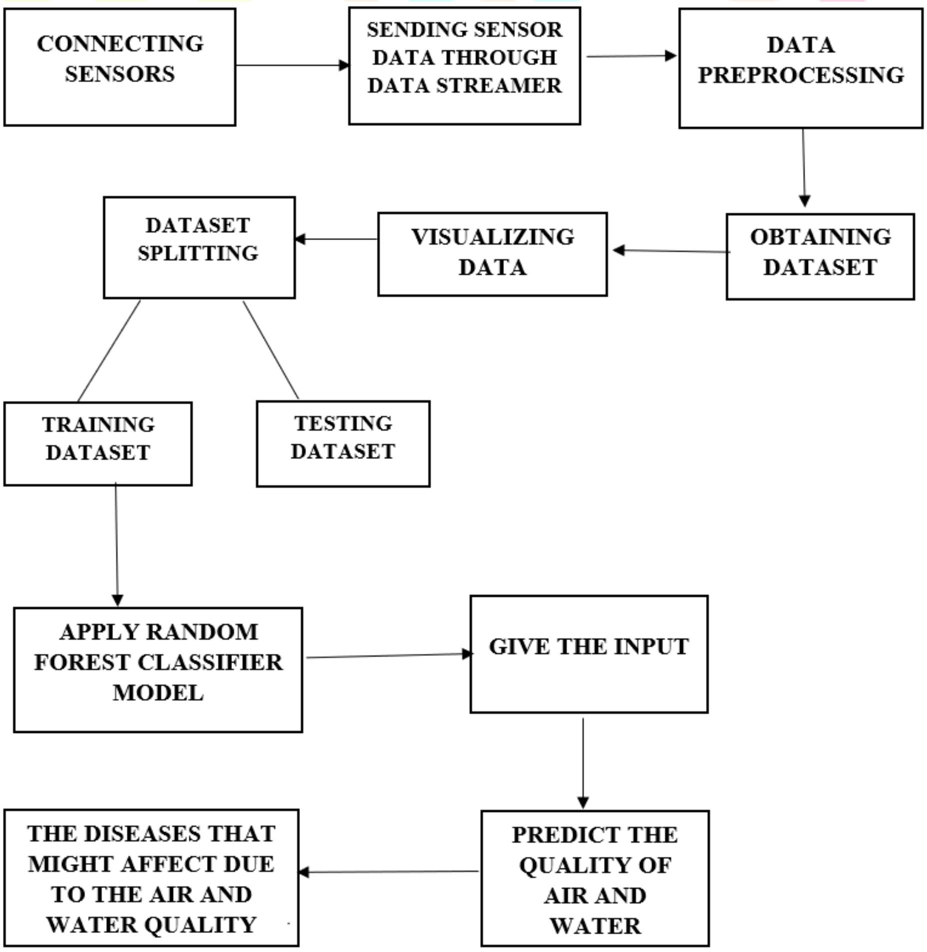

Fig. System Architecture

3.2 MODULES DIVISION:

1. **Collecting data:** Connecting sensors to the IoT network can be done using wired or wireless connections, depending on the type of sensor and the network infrastructure. Our project's air and water sensors include MQ135, MQ4, MQ7, PM2.5, and total

dissolved salts sensors. Before connecting the sensors to a microcontroller the sensors are calibrated to reduce the error while generating the output. Now the sensors are connected through the Arduino Uno microcontroller with ESP8266 wifi-module. The data recorded by the sensors is transferred to Microsoft Excel through the Data Streamer in excel.

2. **Data Preprocessing:** In machine learning, data preprocessing plays a crucial role in converting raw data into a more effective and efficient format that can be readily comprehended and analyzed by machine learning algorithms. This process encompasses various techniques that aid in enhancing, cleaning, and transforming the data. Common techniques employed in data preprocessing include Data Cleaning, Data Transformation, Feature Selection, Data Integration, and Data Reduction.
3. **Data Visualization:** Data visualization plays a crucial role in machine learning since it enables us to extract valuable insights from data and uncover patterns that might not be apparent in its raw form. Essentially, data visualization refers to the technique of presenting information in graphical or visual formats, making it easily understandable and facilitating the extraction of meaningful insights.
4. **Training data through random forest classifier:** The random forest algorithm is a versatile machine learning technique that can handle both regression and classification tasks. As a type of ensemble learning, it merges several models to enhance the overall model's accuracy and robustness. Specifically, the random forest classifier creates numerous decision trees and aggregates their outputs to classify data. By constructing each decision tree with a random subset of the features and data samples, the algorithm prevents overfitting and improves tree diversity. Ultimately, the final prediction is determined by majority voting across all trees.
5. **Predict the air and water quality index:** Based on the concentration of various pollutants the Random Forest Classifier will predict the air and water quality index. With the obtained index value we can get to know the extent upto which the pollutants can affect the health.

3.3 IMPLEMENTATION AND ALGORITHM:

To implement the code for Arduino microcontroller boards, we typically use the open-source Arduino IDE software, which is an integrated development environment that facilitates programming and software development. The software boasts a user-friendly interface, allowing users to write, compile, and upload code to their Arduino boards. Compatible with various operating systems, such as Windows, macOS, and Linux, the Arduino IDE is popularly used by professionals, students, and hobbyists for developing and prototyping a range of projects.

Once the sensors are connected through the Arduino microcontroller, the written code must be verified and uploaded to the board. To view the data and check the results, one can select the serial monitor button.

Upon receiving the results, the data is transferred to an Excel file via the data streamer in Excel. The data is subsequently analyzed and presented visually. A Random Forest Classifier model is then constructed to train the data and predict air and water quality, which can be used to gauge the health conditions of people living in a specific area.

The webpage has been designed to provide continuous monitoring of pollutant concentrations. Real-time data recorded by sensors can be viewed on the webpage, allowing for up-to-date information on pollutant levels.

Fig. Webpage for pollutants concentration

4. DESIGN:

DATA FLOW DIAGRAM

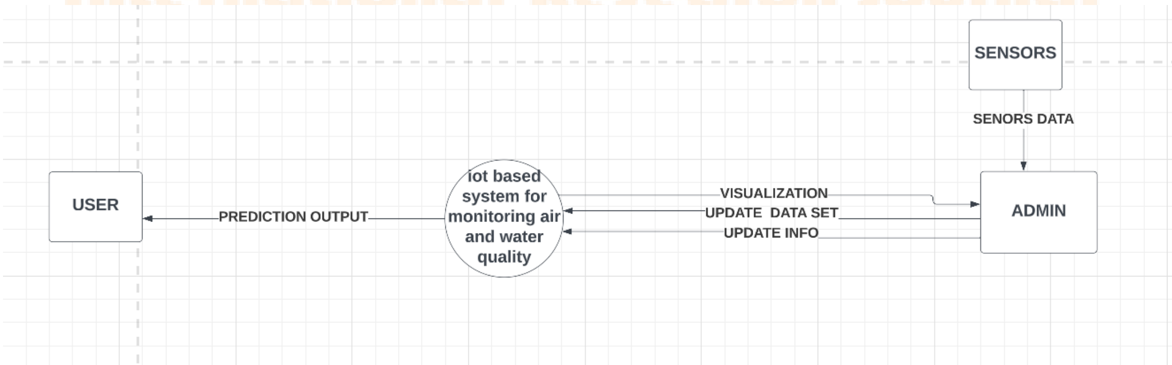

Fig. DFD Level 0

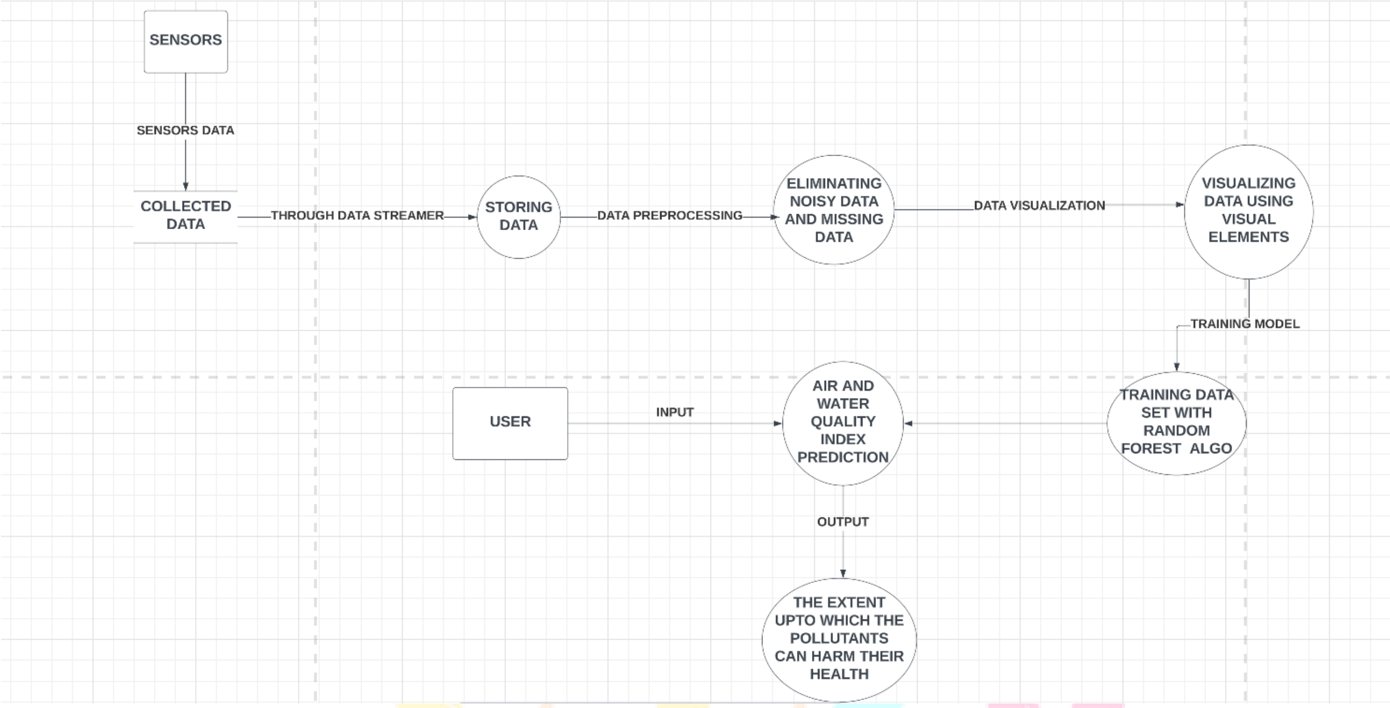

Fig. DFD Level 1

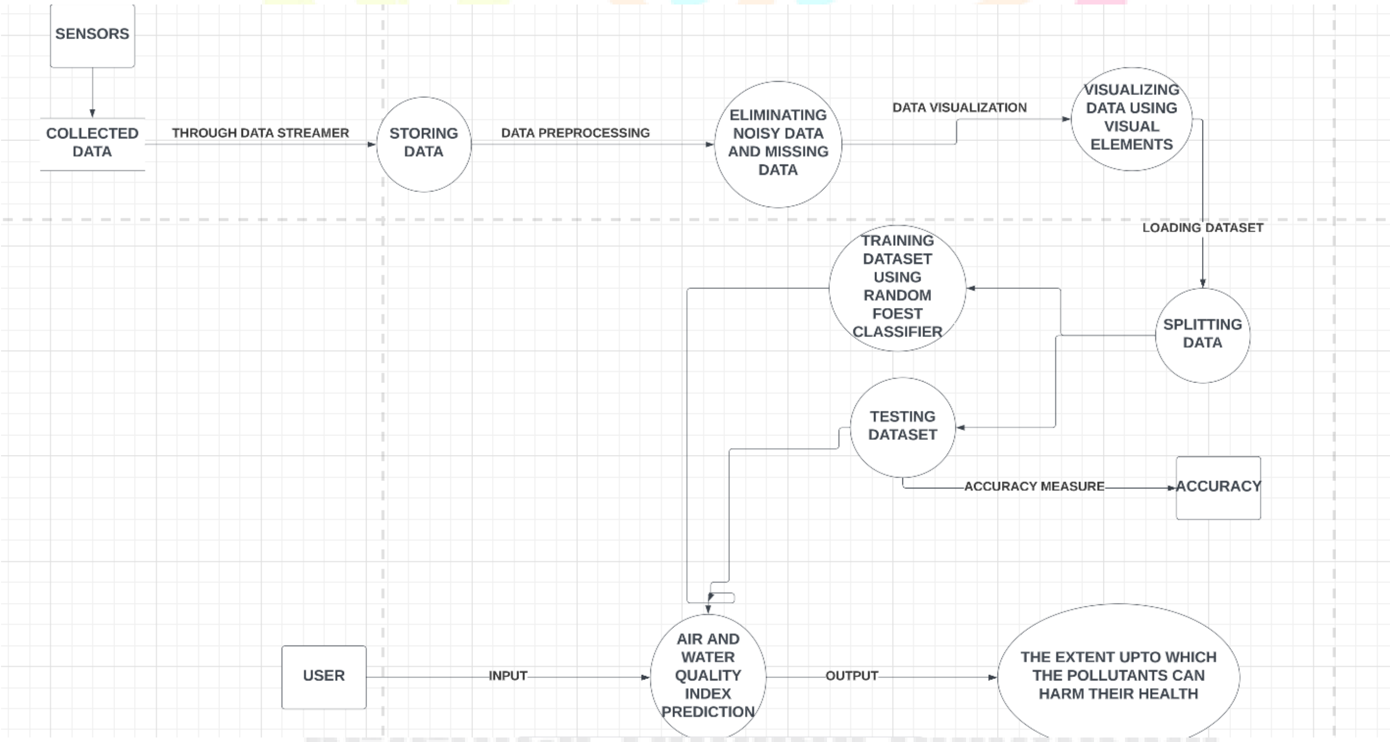

Fig. DFD Level 2

4.1 FUNCTIONAL REQUIREMENTS:

1. **Sensor Integration**: The system should be able to integrate multiple sensors for monitoring different parameters like temperature, humidity, air quality, water quality, and pollutants.
2. **Real-time data acquisition**: The system should collect and transmit data in real-time. This is important to identify anomalies quickly and take action before any damage occurs.
3. **Data storage**: The system should store the collected data in an excel file, which can be accessed later for analysis.

4. **Data analytics**: The system should have the capability to analyze the data collected and generate meaningful insights. The data analytics can include trend analysis, anomaly detection, and correlation analysis.
5. **Webpage** : The system should provide an intuitive webpage for the users to view the real-time data
6. **Data visualization**: The system should provide an easy-to-understand visual representation of the data collected. This can include graphs, charts, and maps.

5. INPUT AND EXPECTED OUTPUT:

INPUT:

The input for the trained model in this system will consist of the air and water quality indices. Specifically, the AQI will be derived from the concentration of pollutants, while the WQI will be determined based on the total dissolved salts (TDS) concentration in the water.

EXPECTED OUTPUT:

Once the model is given the input of the air quality index (AQI) and total dissolved salts (TDS) level of water, it generates predictions for the AQI bucket and the water quality index (WQI). These predictions are indicative of the condition of the air and water, and can provide information on whether they are safe for human health. Additionally, data visualization can be used to determine the degree to which these pollutants may pose a threat to human health.

6. CONCLUSION

An IoT-based system for monitoring air and water quality can be a valuable tool for ensuring human health and safety. By collecting and analyzing data on pollutant concentrations and TDS levels, the system can generate accurate predictions of the AQI and WQI, indicating the overall condition of the air and water. With the help of data visualization, it can also provide insight into the extent of the threat posed by these pollutants. Overall, such a system has the potential to contribute significantly to the promotion of environmental sustainability and the protection of public health.

7. FUTURE SCOPE

As the aforementioned system has the capability to only display real-time data via the webpage, it is also feasible to send notifications to the residents of that locality, enabling them to stay aware of their health status.

8. REFERENCES

1. T. S. Kitchlian authored a research paper on the employment of IoT and Machine Learning techniques for the monitoring and prediction of air quality, as stated in the publication available at https://www.ijsrp.org/research-paper-0322.php?rp=P12312043.
2. The utilization of IoT and Machine Learning for air quality monitoring and disease prediction was investigated by Darshini Rajasekhar and Aravind Sekhar, as mentioned in their paper, accessible at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3777352
3. N. Prasad's work on a Smart Water Quality Monitoring System is discussed in the publication found on ResearchGate, which can be accessed at https://www.researchgate.net/publication/287647835_Smart_Water_Quality_Monitoring_System, wherein he details the system's functionalities.