iClean: An Intelligent Industrial IoT Framework for Automatic Sustainable Air Quality Monitoring

Lopamudra Samal, Aryan Kumar Samal, Kamalakanta Mahapatra, Senior Member, IEEE, Ayas Kanta Swain, Member, IEEE, and Saraju P. Mohanty, Senior Member, IEEE

***Abstract*—Air pollution monitoring systems are essential for evaluating rural and industrial environmental quality for safeguarding public health. This study presents a comprehensive IoT-based framework that uses low-cost sensors and machine learning algorithms for real-time monitoring of various pollutants, including LPG, methane, CO, alcohol, PM2.5, PM10, temperature, and humidity. The system gathers sensor data from a gateway node, which is then processed using Support Vector Machines (SVM) and Random Forest Regression (RFR) models to predict pollutant concentrations. Our approach features innovative methodologies for data validation, anomaly detection, and predictive modeling, employing Root Mean Squared Error (RMSE) as the performance metric. The model achieved a remarkably low RMSE value of 0.022, significantly improving the accuracy and reliability of air quality assessments. Experimental results highlight the system's capability to capture complex environmental patterns and predict pollutant levels with high precision. This research intends air pollution monitoring from cost-effective Internet of Things (IoT) solutions and machine learning techniques. Additionally, the user interface is designed for mobile applications, offering real-time data access, alerts, and notifications, thereby enabling personalized environmental health management and targeted pollution control strategies in industrial areas.**

***Index Terms*—Air pollution monitoring, IoT, Low-cost sensors, Machine learning, Pollutant prediction, Remote monitoring.**

I. Introduction

TODAY, air pollution is one of the most pressing environmental challenges. It has a significant impact on public health and the ecosystem. The World Health Organization (WHO) estimates that millions of premature deaths each year are linked to exposure to polluted air. This problem becomes worse due to rapid urbanization, industrial activities, and increased vehicle emissions. Information and communications technology (ICT) also contributes to carbon dioxide emissions [1]. These factors highlight the need for effective air quality monitoring and mitigation strategies to reduce health risks and protect the environment. The increasing demand for the Internet of Things (IoT) leads to the growth of connected devices across the world. This growth contributes to air pollution through carbon emissions [2]. In addition, air quality monitoring is essential for greenhouses to promote sustainable agriculture [3]. For these reasons, adopting a sustainable IoT architecture becomes essential [4]. Fig. 1 shows the requirement of an IoT-enabled air pollution monitoring system.

Although traditional air quality monitoring systems are quite effective, they are often expensive and limited in scope. They face problems in providing real-time local pollution data in a variety of environments. Centralized systems generally lack the scalability and responsiveness required to adapt to fluctuating pollution levels in urban and industrial areas. Additionally, they fail to effectively engage users, which limits their ability to raise public awareness and drive action. To address these gaps, this paper presents iClean (intelligent clean-air framework), an industrial IoT system for automatic and sustainable air quality monitoring. The framework integrates low-cost multi-sensors with a Raspberry Pi gateway to acquire PM2.5/PM10, gases (LPG, methane, CO, alcohol), temperature, and humidity, and transmits data to the cloud for storage and visualization. Machine learning models—Support Vector Machines (SVM) and Random Forest Regression (RFR)—are applied to infer pollutant concentrations from historical and real-time streams. A lightweight mobile application provides real-time visualization and alerts, while a soldered ADC/PCB interface and ThingSpeak support ensure reliable operation. Designed with sustainability in mind, iClean uses low-power sensing, fixed-interval telemetry, and durable hardware to minimize energy use and reduce hardware waste, aligning with the principles of the American Council for an Energy-Efficient Economy (ACEEE) [5]. Thus, iClean delivers accurate, real-time monitoring while supporting energy-conscious, long-term deployment in industrial and urban contexts. Beyond real-time air quality monitoring, the framework supports sustainable technology practices through energy-conscious design and long-term operational viability.

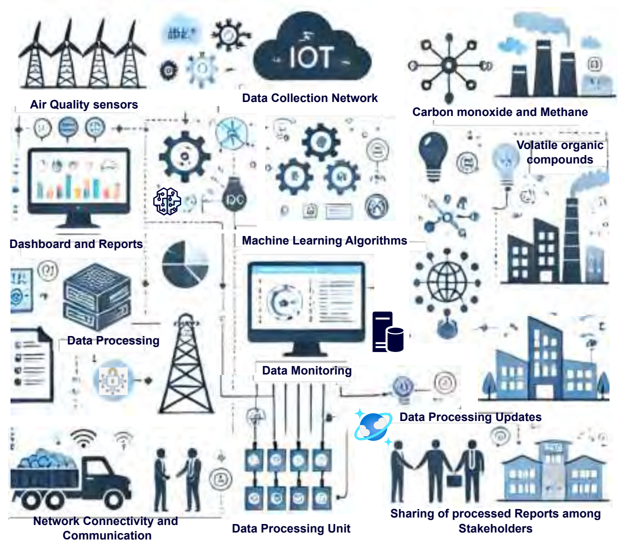

Fig. 1: IoT-enabled air pollution monitoring system

This paper is organized as follows: Section II summarizes prior research and highlights the main contributions of this work, Section III discusses the proposed system architecture, Section IV details prototype development, Section V presents the results and discussion, and Section VI concludes with a summary of the findings.

II. Background and Key Contributions

A. Related Research Work

Several monitoring techniques are recommended and used to measure air quality. These methods vary in complexity, ranging from low-cost sensors to advanced machine learning models. Real-time data collection and spatial distribution estimation are key features in modern systems. Pollution monitoring in smart cities often integrates IoT sensors, AI, and big data analytics to track air quality in real time. Survey studies on IoT-based smart-city systems also outline deployment challenges and opportunities that parallel those found in pollution monitoring applications. The strengths and weaknesses of IoT-enabled waste management models, including their capacity to monitor pollution levels through IoT devices, are discussed in the survey by Anagnostopoulos et al. [20]

1) IoT-Based Air Quality Monitoring Systems: Numerous approaches for monitoring air quality are introduced and implemented, varying from affordable sensor-based systems to sophisticated machine learning models. The 3M'Air project (Mobile Citizen Measurements and Modeling: Air Quality and Urban Heat Islands), as implemented in Fekih et al.'s research [6], involves the use of inexpensive sensors carried by citizens to raise local awareness of temperature and air quality. Similarly, research integrating IoT technologies for real-time monitoring is explored in various studies. For example, a low-cost pollution monitoring tool is featured in the work of Becnel et al. [9], where 50 air quality monitoring nodes are deployed by volunteers within a wireless network, easing maintenance but raising security concerns. Shaban et al. [7] describe a technique for tracking and forecasting urban air pollution using inexpensive, multi-sensor systems that communicate wirelessly, while Al-Ali et al. [8] develop a mobile data acquisition unit and a stationary pollution server for air pollution monitoring, though their system does not include data analysis for prediction or visualization.

2) Low-Cost Sensor Solutions: Research on low-cost sensor nodes gains significant attention. In the research by Ali et al. [10], a unique, low-cost sensor node is presented, featuring PCB-mounted electrochemical sensors capable of short-range, high data rate communication over Wi-Fi and long-range, low-power communication over LoRaWAN IoT networks. A low-cost particulate matter monitoring system, as described by Montrucchio et al. [11], is employed using special-purpose acquisition boards on both stationary and mobile platforms. Although this system is effective on a small scale, maintenance challenges arise for larger implementations. Another study [12] introduces a low-cost, portable air pollution monitoring prototype that features an on-board PM sensor and commercial sensors. A low-cost participatory urban air pollution monitoring system is developed and evaluated in [13], achieving high-resolution pollution surfaces via mobile applications and cloud services, allowing users to measure their monthly pollution exposure.

3) Machine Learning for Pollution Prediction: Machine learning is increasingly integrated into pollution monitoring systems for accurate prediction. Zhang et al. [14] employ both permanent and mobile IoT sensors mounted on vehicles to map air quality fluctuations, with machine learning models being used for accurate measurement and prediction. In the research by Han et al. [15], an improved long short-term memory (LSTM) model is applied for urban air pollution prediction. Nevertheless, high accuracy proves difficult to achieve due to varying topographic and meteorological conditions.

4) Hybrid Approaches for Smart City Pollution Monitoring: This part focuses on research that combines IoT, AI, and big data to monitor pollution in smart cities. In air quality monitoring, a batteryless indoor air quality monitoring device [16] and a solar-powered LoRa-based air quality monitoring device [17] are included. A time-series-based prediction model for Greenhouse Gas (GHG) emissions in smart homes, utilizing LSTM, is designed by Riekstin et al. [18]. The research by Dey et al. [19] introduces a green predictive model (GAP) using a customized big dataset from weather research to forecast winter air quality index (AQI) levels in the Indian subcontinent.

From the survey of related works summarized in Table I, several research gaps are identified. Many existing systems focus on either low-cost sensor deployments or IoT-based frameworks, but not a holistic combination of both. Furthermore, while some research employ machine learning, they are often limited to narrow datasets or specific pollutants, lacking generalization to multiple gases and particulate matter simultaneously. Scalability and integration with mobile applications are also underexplored in prior works, limiting user accessibility and real-time awareness. These gaps motivate the proposed iClean framework, which integrates low-cost sensors, IoT connectivity, and machine learning with a real-time user interface.

B. Issues in the Existing Works and Proposed Solution

Although existing literature report in advance air quality monitoring, many focus on isolated aspects such as low-cost sensing, IoT, or machine learning without offering a complete framework. Challenges like generalization, scalability, and user accessibility to be address comprehensively. To address these gaps, the proposed iClean framework integrates low-cost sensors, IoT connectivity, machine learning, and a real-time mobile interface. The key problems and corresponding solutions are outlined below.

1) Lack of Integrated Frameworks: Many existing systems emphasize either affordable sensing hardware or network-based data transfer, but rarely combine both into a single, unified approach. This separation reduces the overall effectiveness of monitoring solutions. The proposed framework brings together sensing, connectivity, and processing in a holistic manner. This integration ensures both cost-effectiveness and

TABLE I: A brief overview of related works on pollution monitoring

Prior Work	System type	Measured parameter	Computing platform	Methods used	Wireless line	Fixed or mobile	Local Storage
IoT-based air quality monitoring systems							
Fekih et al., 2021 [6]	low-cost environmental monitoring	Temperature, Humidity, NO2, PM1, PM2.5, PM10	Arduino, Raspberry Pi	Four-layer system architecture based on LoRaWAN for IoT applications	LoRaWAN	Mobile	No
Shaban et al., 2016 [7]	Urban air pollution monitoring system with forecasting models	O3,NO2,SO2	A pilot initiative of Qatar Mobility Innovations Center (QMIC)	Multivariate modeling with M5P algorithm for accurate forecasting performances	WiFi	Mobile	Yes
Al-Ali et al., 2010 [8]	Mobile data-acquisition unit	CO,NO2,SO2	16-Bit Single-Chip Microcontroller	Mobile-DAQ fixed internet enabled pollution monitoring server	TCPIP	Mobile	Yes
Becnel et al., 2019 [9]	Distributed low-cost pollution monitoring platform	PM1, PM2.5, PM10, Temperature, Humidity, Light intensity, CO, NO	TI CC3200 WiFi-enabled	Linear calibration, MQTT-based data upload, micro-SD offline data logging,	WiFi (MQTT)	Mobile	Yes
Low-cost sensor solutions							
Ali et al., 2020 [10]	Low-cost sensor node for air pollution monitoring	CO, NO2, PM	MCU	Artificial neural network for sensor calibration with high accuracy	LoRaWAN	Mobile	Yes
Montrucchio et al., 2020 [11]	Low-cost air-quality system	Temperature, Relative humidity, pressur	Raspberry Pi	Calibration, and validation of a low-cost air quality monitoring system	–	Mobile	Yes
Das et al., 2022 [12]	A self-sustainable system with NB-IoT	CO, NO2, SO2, O3, PM, Temperature, Humidity,	ATMEL Studio IDE	On-board PM sensor with power control mechanism with NB-IoT power saving mode for energy efficiency	NB-IoT	Mobile	Yes
Hu et al., 2015 [13]	A metropolitan air pollution sensing system	CO,NO2, O3	HazeWatch node	Calibration with GasAlert Micro 5 for pollutant concentrations	WiFi, 3G	Mobile	Yes
Machine learning for pollution prediction							
Zhang et al., 2020 [14]	Hybrid distributed and fixed IoT sensor system	CO2, PM2.5, PM10.	Raspberry Pi	Linear, GPR Cauchy, GPR Gaussian for air pollution data	WiFi	Mobile	Yes
Han et al., 2019 [15]	Wireless sensor network for environmental monitoring in factories	CO, NO2, SO2, O3, PM, Temperature, Humidity,	ATmegal28 chip	Pollution traceability method and LSTM for pollution period prediction	WiFi	Fixed	Yes
Hybrid approaches for smart city pollution monitoring							
Yue et al., 2017 [16]	IoT-based CO2 gas sensor.	CO2	indoor photovoltaic energy harvesting power module	IPEHPM developed for low-power CO2 IoT node	WiFi	Fixed	–
Jabbar et al., 2022 [17]	LoRaWAN-IoT-AQMS	PM 2.5, CO 2, CO, NO 2, SO 2, temperature, humidity	Arduino uno Processor	Investigated existing systems, reviewed AQMS topics, and selected appropriate materials	LoRaWAN	Mobile	No
Riekstinet al., 2020 [18]	GHG Emissions Prediction	CO2	Computing system	LSTM for prediction	–	Fixed	Yes
Dey et al., 2024 [19]	Green AQI prediction and air pollution epidemiology model	AQI, PM 2.5, NO 2, CO	GPA	LSTM, RNN, CNN, MLP, SVR, KNN, RF, big-dataset augmentation	completely offline	Fixed	–

reliable real-time operation, making the system suitable for deployment across diverse environments.

2) Limited Generalization in Data Analytics: Prior studies applying analytical or predictive models often rely on narrow datasets or address only specific parameters. This restricts their ability to generalize to different environmental conditions or multiple pollutants simultaneously. The framework incorporates advanced analytical methods capable of handling heterogeneous data streams. This enables improved reliability and predictive accuracy across varied pollutants and conditions, addressing the limitations of earlier single-focus approaches.

3) Poor Scalability and User Accessibility: Scalability and accessibility remain underexplored in many available systems. Limited consideration of large-scale deployments and lack of user-friendly interfaces restrict broader adoption and practical utility. The framework is designed to scale from small to large deployments while maintaining efficiency. It also integrates real-time visualization and alert mechanisms through user-friendly platforms, ensuring broader accessibility and timely decision-making.

4) Limited Consideration of Sustainability: Many existing monitoring frameworks overlook sustainability aspects such as energy efficiency, hardware longevity, and overall environmental impact. This results in higher operational costs, frequent hardware replacement, and increased electronic waste, reducing the long-term viability of such systems. The proposed framework emphasizes sustainable design by adopting energy-efficient operation, minimizing unnecessary hardware replacement, and ensuring long-term usability. These choices reduce resource consumption and extend the framework's relevance for both industrial and urban deployments, aligning it with broader sustainability goals.

C. Novel Contribution of Proposed Work

The proposed iClean framework introduces novelty at the architectural, prototyping, and system-integration levels by integrating multiple capabilities into a unified, practical, and sustainable system for air quality monitoring. Specifically, iClean introduces:

- A complete end-to-end architecture and working prototype that integrates a low-cost multi-sensor gateway node, cloud back-end, and mobile application, and validates its performance using real-time deployment data and machine-learning-based prediction of multiple pollutants.
- Combines cost-effectiveness with intelligence by employing affordable, durable sensors together with machine learning models to predict multiple pollutants simultaneously.
- Enhances accessibility through seamless IoT connectivity, cloud integration, and a mobile application that provides real-time visualization, alerts, and user interaction.
- Ensures scalability and adaptability by supporting reliable operation across varied environments, with provisions for local preprocessing and edge-level inference to reduce latency and energy demands.
- Promotes sustainability by emphasizing energy-efficient operation, minimizing hardware waste, and enabling long-term usability for both industrial and urban applications.

Collectively, these contributions establish iClean as a comprehensive and sustainable framework that advances beyond the fragmented approaches of earlier studies.

III. Proposed System Architecture

The proposed system architecture is designed to monitor air quality using a single gateway sensor node. This IoT node is integrated with various sensors, and data from these sensors is sent to the ThingSpeak cloud for real-time monitoring and later downloaded for machine learning-based forecasts. The system is also equipped with a mobile app that allows remote monitoring and alerts users to significant pollution levels. All parts of the system are connected to the internet to transfer and collect data. Fig. 2 shows the block diagram of the architecture of the proposed system with a single gateway.

A. Sensor and Gateway Configuration

The system integrates a variety of sensors into a single IoT-Node, which serves as the gateway for data processing and transmission. The sensor node is equipped with gas sensors for detecting gases like LPG, methane, and alcohol, and environmental sensors for measuring temperature, humidity, and particulate matter. These sensors interface with the IoT-Node via an analog-to-digital converter (ADC), allowing seamless conversion of analog sensor data into digital form for further processing. Each sensor collects specific environmental parameters, and the IoT-Node handles local data processing.

1) Temperature and Humidity Monitoring: The system includes a temperature and humidity sensor, which relies on thermistors and capacitive sensors to measure environmental conditions. The sensor transmits data to the gateway via a single-wire serial interface. The internal analog-to-digital conversion ensures that the data is transmitted accurately with error-checking mechanisms in place. The IoT-Node processes this information for both real-time analysis and long-term storage.

2) Particulate Matter Monitoring: A laser-based scattering technique is used to quantify airborne particulate matter. The sensor interacts with the gateway through a serial interface, and specialized commands trigger data collection. The raw data from the sensor is validated with checksum verification before being processed by the IoT-Node.

3) Gas Detection: The system uses a gas sensor to detect volatile gases such as LPG and methane. Analog data from this sensor is converted into digital signals through an ADC module. The gas concentration is then calculated based on sensor response curves provided in the datasheet, which are pre-calibrated to account for variable gas concentrations. The IoT-Node processes this data for accurate reporting and potential safety.

B. Real-Time Data Transfer to Cloud

The gateway functions as a local hub, collects sensor data, performs basic preprocessing, and temporarily stores values to ensure uninterrupted acquisition during short connectivity losses. Data are transmitted to the ThingSpeak cloud over Wi-Fi at fixed intervals of 30 seconds, which does not hinder real-time monitoring. ThingSpeak employs the MQTT protocol to ensure reliable data transfer and organizes the uploaded information into private or public channels for secure storage. The platform also provides built-in data visualization tools, enabling real-time graphical displays accessible via a web interface or mobile app. Through these visualizations, users monitor air quality levels, identify trends, and configure alerts based on predefined thresholds, thus ensuring continuous monitoring with required action insights. In the current implementation, predictive inference is performed on a client PC; however, the gateway hardware is capable of supporting on-device inference, which further minimizes latency and reduces reliance on continuous cloud connectivity.

C. Data Download and Dataset Preparation

Once the data is stored in the ThingSpeak cloud, it is periodically downloaded to a local device for further processing.

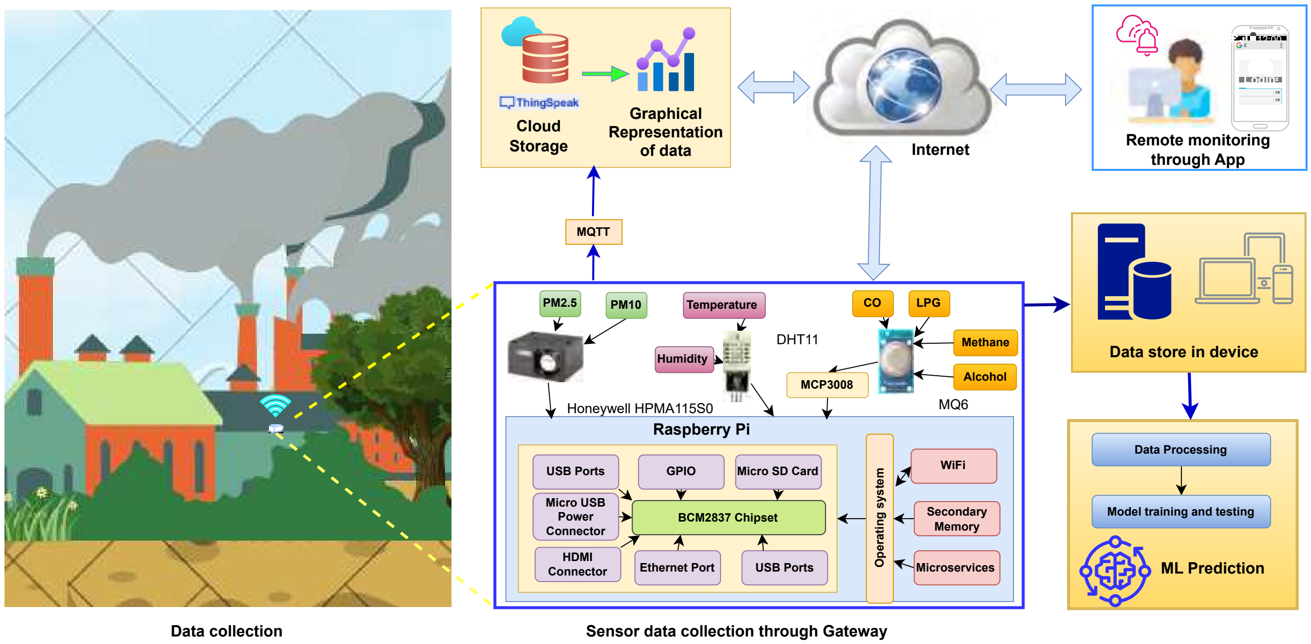

Fig. 2: Block diagram of proposed solution in current work

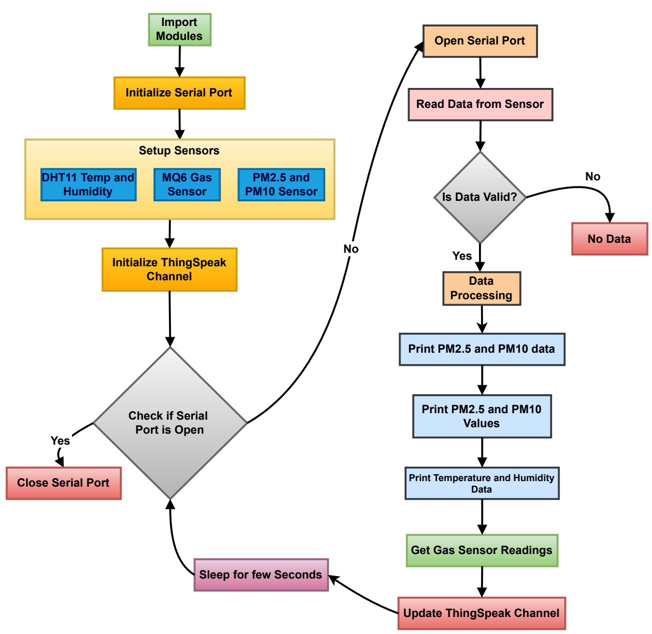

Fig. 3: Flow diagram of sensor data collection through a single gateway.

Fig. 3 shows flow diagram of sensor data collection through a single gateway During our experiments, the collected dataset is generally clean and does not exhibit significant noise or abnormalities. However, in a few instances—particularly when relocating the prototype to different places—some sensors produce abnormally high or low values. These are identified as erroneous and are removed manually. Since such disturbances may occur in real-world deployments, standard preprocessing techniques can also be employed when required. Examples include moving average smoothing to reduce fluctuations, z-score or interquartile range (IQR)-based detection for abnormal values, and normalization to bring heterogeneous sensor data onto a common scale. These measures ensure that the iClean framework remains reliable and adaptable under diverse operating conditions.

D. Machine Learning and Data Analytics

The system leverages advanced machine learning and data analytics techniques to enhance the accuracy and predictive capabilities of the air quality monitoring system. The predictive model is deployed using models such as SVM and RFR to forecast air quality trends by analyzing historical data and real-time inputs, providing valuable insights into future pollution patterns and ensuring that any issues are promptly addressed. Fig. 4 shows the workflow of ML model training, testing, and visualization.

For predictive modeling, we tested prediction algorithms with a 20% test size. The dataset, obtained from ThingSpeak, is used on a client-side PC to employ an ML algorithm for prediction, with RMSE serving as the key performance metric for evaluation. Mathematically, we can denote this as:

$$D = (SensorData_i)_{(i=1)}^{m} \tag{1}$$

Next, the dataset D is divided into training and testing sets using an 80-20 split. Let $Dtrain$ represent the training set and $Dtest$ represent the testing set. Mathematically, we can denote this as:

$$Dtrain = (SensorData_i)_{(i=1)}^{0.8m} \tag{2}$$

$$Dtest = (SensorData_i)_{(i=0.8m+1)}^{m} \tag{3}$$

Algorithm 1 Prediction Model

Step 1: Load Dataset - Import dataset, identify target (Y) and features (X).
Step 2: Split Dataset - Use 80-20 ratio.
function TRAIN_TEST_SPLIT(test_size=0.2, random_state=42)
end function
Step 3: Initialize Models
function RFR_MODEL(n_estimators=100, random_state=42)
end function
function SVR_MODEL(kernel='rbf')
end function
Step 4: Train Models
function FIT(RFR, X_train, Y_train)
end function
function FIT(SVR, X_train, Y_train)
end function
Step 5: Make Predictions
function PREDICT(RFR, X_test)
    Output: predicted_Y_RFR
end function
function PREDICT(SVR, X_test)
    Output: predicted_Y_SVM
end function
Step 6: Evaluate Models
function MEAN_SQUARED_ERROR(Y_test, predicted_Y_RFR)
end function
function MEAN_SQUARED_ERROR(Y_test, predicted_Y_SVM)
end function
Step 7: Visualize Results - Scatter plot for Actual vs. Predicted (RFR, SVM).

In the above expression, m is the total number of data samples collected from the nodes. Once the training and testing sets are defined, an ML model M is developed using the training data $Dtrain$. The model is trained to predict the node number based on the sensor data. Mathematically, this can be represented as:

$$M = Train(Dtrain) \tag{4}$$

After training the model, it is evaluated using the testing data $Dtest$ to assess its performance and accuracy. Mathematically, this can be represented as:

$$Accuracy = Test(M, Dtest) \tag{5}$$

In the above expression, Accuracy represents the accuracy of the ML model in predicting sensor data. RMSE serves as the performance metric in this evaluation. It is frequently employed to gauge the average discrepancy between expected and actual values in a regression problem. The effectiveness of the model and the prediction $\hat{y}_i$ is assessed using the data for the individual sensors, y_i, in accordance with RMSE, which is used as an error criterion and is defined as:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \tag{6}$$

Here, n represents the number of observations. The program flow of the machine learning model is present in Algorithm 1.

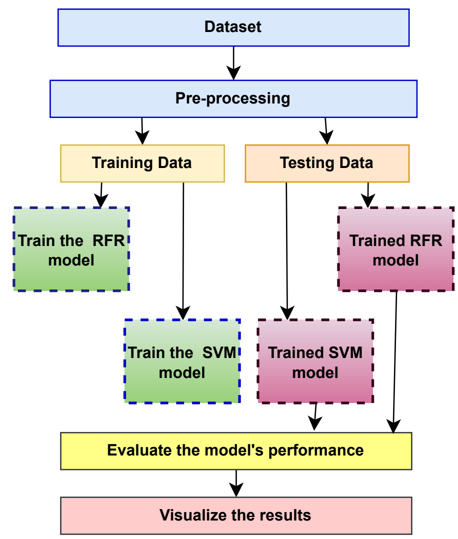

Fig. 4: Workflow of ML model training, testing, and visualization

E. User Interface and Applications

To engage users and provide them with information that induces possible action, the system includes user-friendly interfaces and applications. Mobile applications offer real-time air quality data to reduce pollution exposure. This app is intuitive and accessible, encouraging users to stay informed about their environment. The system also features alerts and notifications, delivering real-time warnings about critical pollution levels to ensure that users are promptly informed of any immediate health risks.

By integrating state-of-the-art sensor technology, IoT connectivity, and machine learning, the proposed system architecture provides a comprehensive and reliable air quality monitoring solution. This system enhances the precision and coverage of air quality data and empowers individuals and communities with actionable insights to improve environmental health.

IV. Prototype Development

The initial phase of the project involves the design and construction of a prototype air quality monitoring system. The system integrates multiple sensors into a central Raspberry Pi gateway to monitor key environmental parameters. The sensor node includes an MQ6 sensor for LPG, methane, CO, and alcohol detection, a DHT22 sensor for temperature and humidity, and a Honeywell HPMA115S0 sensor for PM2.5 and PM10. These sensors are connected to the Raspberry Pi's

GPIO pins, with an MCP3008 ADC chip converting analog sensor data to digital signals for processing.

1) DHT22 Temperature and Humidity Sensor: The DHT22 sensor is powered by the Raspberry Pi's 3.3V and ground, with the data pin connected to a GPIO pin. The Adafruit Python DHT library reads temperature and humidity values. The sensor uses a thermistor for temperature and a capacitive sensor for humidity, with an internal ADC converting measurements into a 40-bit digital signal containing data and a checksum for error checking. This data is sent over a single-wire serial interface, and integrity is ensured through checksum validation.

2) Honeywell HPMA115S0 PM2.5/PM10 Sensor: The Honeywell HPMA115S0 particulate matter sensor is connected via a serial interface. It uses a laser-scattering technique to measure PM2.5 and PM10 concentrations. The sensor processes the scattered light data internally and transmits it to the Raspberry Pi. Commands such as 0x68 0x01 0x01 0x96 initiate readings, with data validated through checksum verification.

3) MQ-6 Gas Sensor: The MQ-6 sensor, which detects gases like LPG and butane, connects to the Raspberry Pi through the MCP3008 ADC for analog-to-digital conversion. The sensor operates on a chemical reaction principle that alters the conductivity of its ceramic element. The output voltage, calibrated using response curves from the datasheet, provides gas concentrations in Parts Per Million (PPM). A voltage divider ensures compatibility between the sensor's output and the Pi's input.

The Raspberry Pi, running its own operating system, with a BCM2837 chipset, secondary memory, and connectivity options (Ethernet, WiFi), serves as a gateway for edge processing and data transmission. Processed data is sent to the ThingSpeak cloud via Wi-Fi and MQTT protocols for real-time visualization and long-term storage. This setup ensures continuous, accurate monitoring of air quality.

Connectivity is established through Wi-Fi, enabling dual-mode communication for efficient data transmission, and ensuring seamless integration into an IoT network for effective device management. Data processing and storage are managed through edge computing for localized data processing at sensor nodes, and ThingSpeak for cloud storage to facilitate long-term analysis. A structured database management system is employed for easy retrieval and analysis of collected data. Fig. 5 shows three main parts of prototype development.

The user interface is designed to be accessible via mobile applications, providing real-time data access, alerts, and notifications. After the prototype is developed, it undergoes extensive field testing to evaluate its performance under various environmental conditions and to ensure its reliability and accuracy. For experimental validation, the prototype is deployed at multiple spots within the Electronics and Communication Engineering Department, NIT Rourkela campus. Sensors are positioned at an approximate height of 1.5 m from the ground to represent human breathing level. Data are collected over several days at regular intervals of 30 s, capturing variations in temperature, humidity, particulate matter (PM2.5 and PM10), and gas concentrations (LPG, methane, CO, alcohol). This mix of indoor laboratory and corridor environments provides diverse environmental conditions for evaluating the robustness and reliability of the proposed system. The performance of the prototype is evaluated on several fronts:

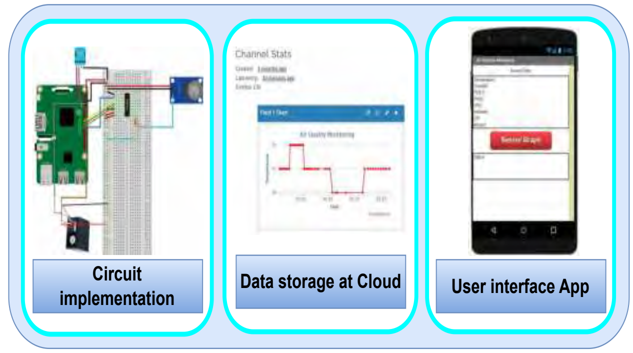

Fig. 5: Three main parts of prototype development.

- Accuracy and Sensitivity: Sensor data is compared against reference-grade equipment to assess the accuracy and sensitivity of the measurements.
- Connectivity: The reliability of data transmission via Wi-Fi is tested to ensure seamless connectivity and data flow.
- Data Processing: The efficiency of edge computing and cloud storage solutions is evaluated to ensure effective data management.
- Machine Learning Models: The calibration and predictive models are tested for accuracy and robustness, using RMSE as the performance metric. RMSE measures the average discrepancy between predicted and actual values, providing a comprehensive assessment of the predictive model's accuracy by calculating the square root of the average squared variations.

Additionally, mobile applications are evaluated for their usability and effectiveness in providing real-time data access, alerts, and personalized health recommendations.

V. Results, Analysis and Evaluation

A custom soldered PCB circuit is designed and developed to support the ADC functionality of the IoT system. The circuit design is translated into a physical soldered PCB layout, which is then fabricated and assembled. After construction, comprehensive testing is conducted to verify the hardware circuit's performance. Each individual component is verified, signal integrity is assessed, and potential issues such as noise interference or voltage fluctuations are checked. Comprehensive coverage and reliable data collection are ensured across various geographical regions.

The components function as expected, and the circuit successfully facilitates the desired data acquisition, processing, and communication between various modules and sensors in the IoT system. The Raspberry Pi is connected to the sensors to measure temperature, humidity, PM2.5, PM10, LPG, methane, alcohol, and CO gas levels. The program is executed in Python, and the Raspberry Pi is accessed via remote desktop, allowing

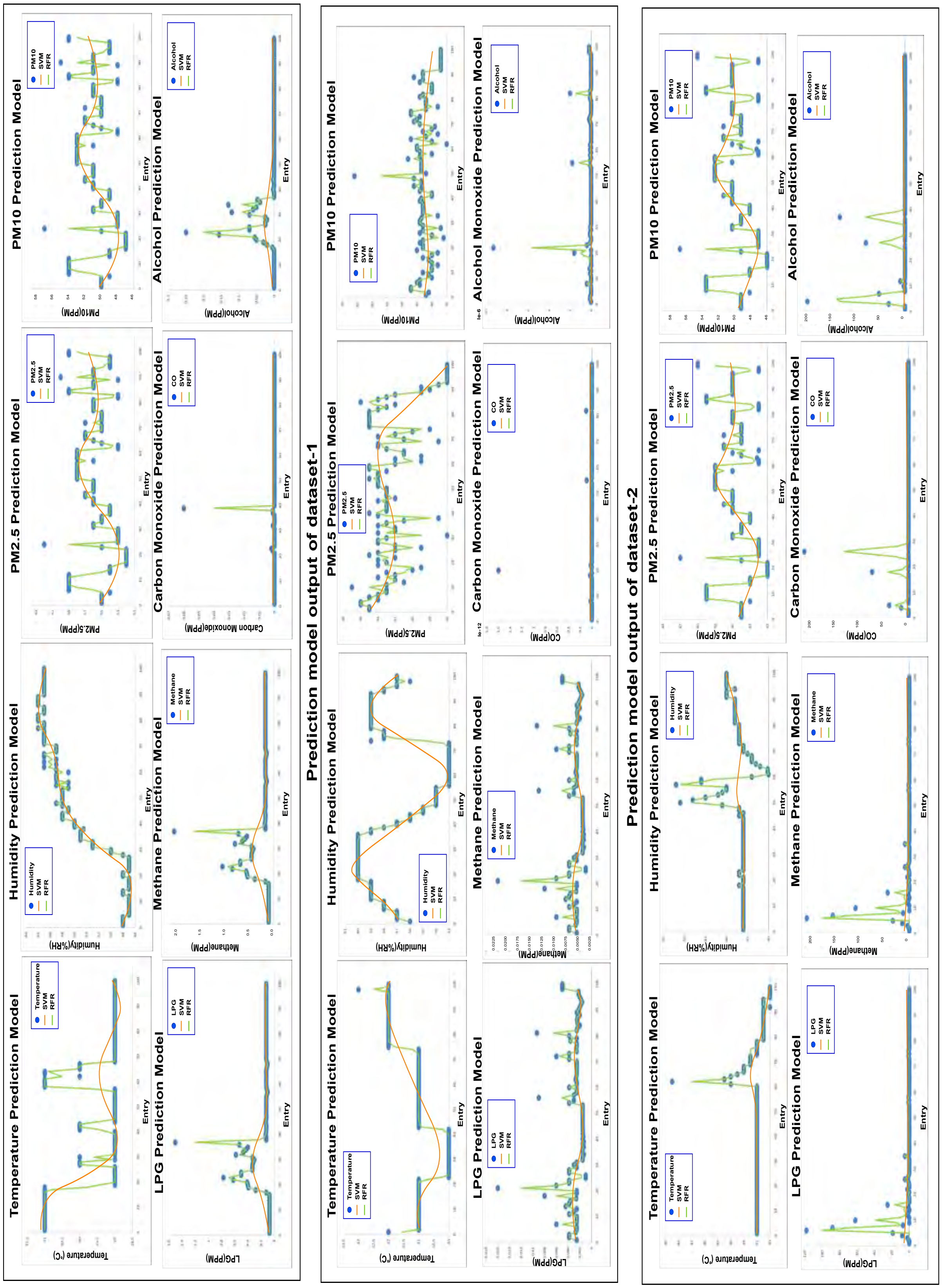

Fig. 6: Prediction model output for all three datasets

TABLE II: Prediction results comparison with different regression algorithms for three datasets

Dataset	ML model	RMSE values							
		Temperature	Humidity	PM2.5	PM10	LPG	Methane	CO	Alcohol
1	SVM	0.62	0.839	2.021	2.671	0.156	0.228	0.006	0.032
	RFR	0.38	0.785	1.826	2.438	0.157	0.232	0.007	0.026
2	SVM	0.32	0.821	2.11	3.527	0.09	0.07	0.022	0.025
	RFR	0.212	0.66	2.058	3.428	0.1	0.08	0.022	0.023
3	SVM	1.559	3.607	1.002	1.453	16.535	29.507	21.36	20.547
	RFR	1.226	2.462	0.625	0.872	18.003	32.215	26.394	27.675

TABLE III: Performance comparison of the proposed system.

References	**Ml model used**	**cloud services utilization**	**User Interface**	**Used dataset generated by system**	**Hardware used**	**Real-Time Monitoring**	**Parameter**	**Least RMSE index**
Zhang et al., 2020 [14]	RFR, SVM, GBR	Yes	yes	Yes	Arduino mega, Raspberry Pi 3B, sensors	Yes	Temperature, Humidity, CO2, PM2.5, PM10	5.7
Yang et al.,2023 [21]	DNN, GBR, RF	No	No	Yes	N/A	No	Temperature, Humidity, PM2.5,	14.39
Naz et al., 2023 [22]	LSTM, GRU	No	No	No	N/A	No	NO2, O3, SO2, PM2.5, PM10	0.59
Shao et al., 2023 [23]	VMD, XGBoost	No	No	No	N/A	No	MP2.5	1.98
Sharma et al., 2020 [24]	CNN LSTM hybrid model	No	No	No	N/A	Yes	Temperature, Humidity, NO2, O3, PM2.5, PM10	7.44
Yelisetti et al., 2023 [25]	RNN, LSTM, GRU, Bi-RNN, Bi-LSTM, Bi-GRU	No	No	Yes	N/A	No	Temperature, CO2	0.102
Lin et al., 2020 [26]	RNN, LSTM, CART, GBM, SVM	No	No	No	N/A	No	PM2.5	3.13
Nguyen et al., 2023 [27]	LSTM-BNN, RNN	No	No	Yes	N/A	No	PM2.5	2.415
Gu et al., 2022 [28]	RBFN, SVR, RF, ELM, DBN, BLS	No	No	Yes	XHAQSN-808 monitor, GPRS, Wi-Fi, Bluetooth, various sensors	No	CO, NO2, SO2, O3, PM2.5, PM10	0.415
Proposed Work (iClean)	SVM, RFR	Yes	Yes	yes	Raspberry Pi 3B, various sensors	Yes	**LPG, methane, CO, alcohol, PM2.5, PM10, temperature, humidity**	0.022

connection to the Pi from anywhere. Fig.7 shows the hardware implementation of the design.

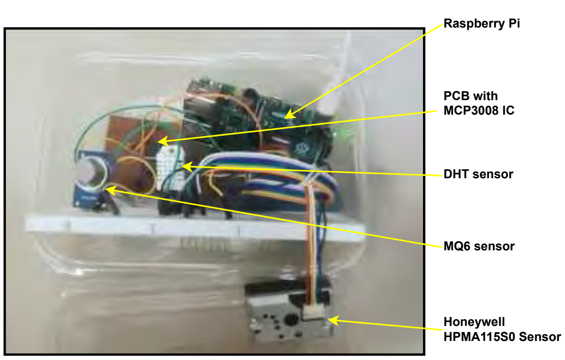

Fig. 7: Hardware implementation of the circuit

In this study, Support Vector Machine (SVM) and Random Forest Regression (RFR) are selected as the primary predictive models since they are widely recognized for handling heterogeneous, nonlinear sensor data. SVM is effective in capturing nonlinear relationships even with limited training data, whereas RFR, as an ensemble method, offers robustness against overfitting and models complex feature interactions. These complementary strengths make them highly suitable for pollutant prediction in the proposed iClean framework. While alternative approaches such as LSTM, CNN, and hybrid CNN+LSTM show high accuracy in pollutant forecasting, they generally require large datasets and substantial computational resources, which limits their applicability in low-power IoT deployments. Table IV presents a comparison of commonly used machine learning models for air quality monitoring. In contrast to deep learning models, SVM and RFR provide a strong balance of accuracy, robustness, and computational efficiency, making them particularly well aligned with the design goals of iClean.

A. Training and testing of Dataset

The dataset obtained from ThingSpeak through a single gateway is used on a client-side PC to predict different models using SVM and RFR. As pollutant monitoring is a regression task, RMSE is selected as the principal evaluation metric.

TABLE IV: Comparison of Machine Learning Models for Air Quality Monitoring and Their Suitability for iClean

References	Algorithm	Strengths	Limitations	Suitability for iClean
Iskandaryan et al., 2020 [29]	SVM	Handles nonlinear data; works with small datasets	Parameter tuning needed	✓ Good for nonlinear sensor data
Gupta et al., 2023 [30]	RFR	Robust to noise; less overfitting	Training cost rises with dataset size	✓ Robust, reliable with noisy IoT data
Wei et al., 2023 [31]	LSTM	Good for time-series data	Needs large datasets; high computation	× Heavy for Raspberry Pi–based setup
Ragab et al., 2020 [32]	CNN	Learns spatial features	Data-hungry; not lightweight	× Not suitable for low-power IoT
Zhang et al., 2022 [33]	Hybrid (CNN+LSTM)	High accuracy	Very resource intensive	× Impractical for edge devices

RMSE penalizes large deviations and directly reflects prediction accuracy in numerical terms, making it more suitable than metrics like accuracy or sensitivity, which are typically applied in classification settings. Other aspects, such as connectivity and data handling, are qualitatively validated during prototype deployment but are not included as quantitative metrics in this study; however, during prototype testing these parameters are observed to perform satisfactorily without noticeable issues. The graphs in Fig. 6 are plotted with the concentration of the sensor data on the y-axis and the entry numbers on the x-axis. All gas sensor and PM sensor data show gas concentrations in PPM, and the models are developed based on this sensor data. Blue dots represent the sensor data collected by the hardware, while the orange line represents the SVM model. Similarly, the RFR model is plotted against the input data as a green line. Three datasets are created and analyzed separately to validate the proposed model.

1) Dataset 1: Data for the first dataset is collected over a few days and is used as a testing dataset, while the historic sensor data serves as the training dataset. Comparing the predicted values from the models with the actual sensor data reveals that both models are effective in capturing general trends and fluctuations. However, upon closer examination, the RFR model demonstrates superior performance based on RMSE. The predicted values from the RFR model closely align with the actual data points, reflecting its ability to capture complex patterns and subtle variations more accurately. Table II shows the RMSE values for all parameters. Although some of these RMSE values are very close to each other, overall, the RMSE values indicate that the prediction models are promising.

2) Dataset 2: Examining the second dataset collected from sensors on different days, similar in size to the previous one. Analyzing the predictions generated by the model, the graphs and RMSE values reveal varying levels of accuracy. The SVM model exhibits a higher RMSE value for temperature, humidity, PM2.5, and PM10, indicating a moderate level of prediction error. Conversely, the random forest model shows a higher RMSE than SVM for LPG, methane, and alcohol. However, for carbon monoxide, both models have the same RMSE error.

3) Dataset 3: Since the areas where the model is tested have stable air quality with minimal sensor fluctuations, an additional dataset is created to simulate potential abnormalities. This third dataset, here data samples collected in very less time interval, that is shorter than the previous ones and is designed to test the model's response to sudden changes in sensor values. Simulating disturbances in a controlled environment is essential for evaluating a machine learning model's performance and robustness. It mimics real-world conditions, assesses the model's response to unexpected events, and highlights areas for improvement, building confidence in its practical reliability.

Disturbances are created using a lighter to release gas near the MQ-6 sensor, increasing gas concentrations without ignition, and by placing hot water near the DHT22 sensor, raising temperature and humidity levels. Upon removing the hot water, humidity levels gradually decrease. Evaluation shows that the RFR model outperforms for temperature, humidity, PM2.5, and PM10, while the SVM model performs better for LPG, methane, carbon monoxide, and alcohol. However, the higher RMSE in Dataset 3 is due to abrupt anomalies, as the models are not optimized for such conditions. This stressed dataset is intended to probe system limits, and anomaly-resilient training or hybrid modeling can potentially improve performance under such disturbances.

Analysis of all three datasets concludes that both SVM and RFR models have strengths and are suitable for different scenarios. The choice of model depends on the dataset's attributes and the prediction task's objectives. The results indicate that RFR generally provides lower RMSE across stable datasets, owing to its ensemble-based robustness. However, under varying conditions (Dataset 2) and synthetically stressed environments (Dataset 3), SVM outperforms RFR for certain gases. This can be attributed to SVM's ability to handle sharp, nonlinear fluctuations, whereas RFR performs better on smoother, more consistent patterns. These findings suggest that model performance depends strongly on dataset characteristics, highlighting the potential benefit of hybrid approaches that combine SVM and RFR for improved robustness.

B. Mobile app as user interface

The Android-based mobile application serves as an intuitive interface for real-time air quality monitoring. It consists of two primary screens: one presents live sensor readings, while the other displays graphical trends of pollutants with axis labels, gridlines, and data markers for clarity. The interface is designed for simplicity and responsiveness, enabling users to easily track changes in air quality. During prototype testing, the app consistently demonstrates reliable performance, handling continuous data streams while maintaining real-time functionality. Its threshold-based alert system promptly notifies users of deviations from normal conditions—for example, a temperature threshold of 40 °C triggers immediate alerts, highlighting its responsiveness. To limit false alarms, the app applies short-term averaging, and further improvements such as anomaly-resilient models and robust statistical filters are possible. Preliminary feedback from 2–3 student users confirms that the app is easy to navigate and responsive. While this small-scale evaluation underscores its usability, a larger user study is planned to comprehensively assess user experience and effectiveness. Overall, the app provides actionable insights and empowers users to respond promptly to environmental changes. Fig. 8 illustrates the mobile app interface.

Fig. 8: Screen shorts of mobile App interface.

C. Discussion

All collected sensor data are validated before being used for training or prediction. For particulate matter sensors, checksum verification ensures the integrity of transmitted values, while gas and temperature readings are cross-checked against calibration curves. During experimentation, the dataset does not show significant noise or outliers, so no additional filtering is applied. In larger deployments, however, preprocessing can be extended with techniques such as moving average smoothing, normalization, or statistical methods like z-score and IQR analysis to detect and remove anomalies. These measures further enhance data reliability for machine learning applications.

Evaluation results show that the proposed approach achieves an RMSE of 0.022, reflecting close agreement between predicted and actual pollutant levels. A comparative overview with existing studies is provided in Table III. Reported values for other works are taken directly from their publications and are based on different datasets and pollutants; therefore, the comparison serves as a contextual reference rather than a strict benchmark. In this study, RMSE is used as the primary quantitative metric, while aspects such as connectivity and mobile app responsiveness are qualitatively verified during prototype testing.

When interpreting these findings, it is important to recognize that dataset characteristics, pollutant diversity, model architectures, and preprocessing strategies can all affect RMSE values. A fairer comparison would require evaluation on standardized datasets under similar conditions.

It is important to acknowledge that the current evaluation is based on a single prototype node deployed over a limited duration. While this setup is adequate to demonstrate the feasibility and baseline performance of the proposed framework, long-term stability aspects—such as sensor drift, environmental exposure, and coordinated multi-node behavior—is not discussed in the present paper as it remain beyond the scope of this prototype-level study. These aspects would be explored in future work through extended deployments and larger-scale implementations. We note that the architectural design already incorporates provisions for multi-node scalability through mesh networking, hierarchical aggregation, and interoperability with existing systems.

Beyond prediction accuracy, the iClean framework advances sustainability by employing low-power sensors, fixed-interval data transmission, and durable hardware, reducing both energy consumption and hardware waste. These design choices align with energy-efficiency principles and extend the framework's suitability for long-term use in both industrial and urban settings. Overall, iClean demonstrates competitive predictive performance and practical advantages—such as multi-gas and particulate monitoring, cost-effective hardware, machine learning–driven prediction, and a real-time mobile application for accessibility—while also promoting energy-conscious, sustainable monitoring. Limitations remain, including reduced performance under stressed sensor conditions (as observed in Dataset 3) and potential latency from cloud dependence. In the context of the present prototype-level study, sustainability is addressed qualitatively through the design choices implemented in the system. Features such as lightweight communication, and the use of modular and easily replaceable components naturally support reduced energy use and longer device lifespan. While a detailed quantitative evaluation of power consumption and maintenance cycles would require extended multi-month field deployment, the current investigation reflects how the proposed framework incorporates sustainability considerations at the design and operational levels.Taken together, these results highlight that iClean delivers a cost-effective, intelligent, sustainable, and user-centric solution for air quality monitoring, with scope for further improvements in robustness and scalability.

VI. Conclusion

This study addresses the critical challenge of air quality monitoring by developing iClean, an integrated IoT-based framework that combines advanced sensors, reliable connectivity, and machine learning–driven data analysis. The system effectively monitors multiple environmental parameters—including particulate matter, gases, temperature, and humidity—while SVM and RFR enable accurate pollutant concentration prediction, achieving an RMSE as low as 0.022. These results demonstrate improved accuracy, comprehensiveness, and user accessibility compared with existing solutions, highlighting the potential of iClean for deployment in both urban and rural contexts.

Beyond predictive performance, the framework emphasizes sustainability through the use of low-power sensors, fixed-interval data transmission, and durable hardware. These choices reduce energy demand, minimize hardware waste, and align with established sustainability principles [5]. The architecture is also designed for scalability, supporting larger deployments through mesh networking, hierarchical aggregation, and interoperability with existing monitoring systems via standard APIs (MQTT/HTTP). While challenges such as latency, synchronization, and energy consumption arise in large-scale deployments, they are mitigated through edge-level preprocessing and adaptive transmission strategies, ensuring efficient and sustainable operation. Overall, iClean demonstrates that a low-cost, accurate, scalable, and sustainable framework significantly enhances real-time air quality monitoring and contributes to long-term environmental management.

Future research enhances the scalability and interoperability of iClean by integrating emerging sensor technologies and standardized communication protocols. Broader machine learning approaches, including deep learning and hybrid models, are investigated to improve robustness and predictive accuracy. The mobile application is refined with an improved interface and public health alert system to increase participation and accessibility. For industrial deployments, security is strengthened through device authentication [34] and encrypted communication [35], while the integration of energy-harvesting modules further reduces the environmental footprint. Together, these developments advance iClean toward a more secure, accurate, and sustainable solution for real-time air quality management, contributing to healthier communities.

REFERENCES

[1] A. Noureddine, M. D. Lodeiro, N. Bru, and R. Chbeir, "The Impact of Green Feedback on Users' Software Usage," *IEEE Transactions on Sustainable Computing*, vol. 8, no. 2, pp. 280–292, 2023, doi: 10.1109/TSUSC.2022.3222631.

[2] P. Mohanty, U. C. Pati, K. Mahapatra, and S. P. Mohanty, "bslight: Battery-Less Energy Autonomous Street Light Management System for Smart City," *IEEE Transactions on Sustainable Computing*, vol. 9, no. 1, pp. 100–114, 2024, doi: 10.1109/TSUSC.2023.3310884.

[3] P. K. Tripathy, A. K. Tripathy, A. Agarwal, and S. P. Mohanty, "My-green: An IoT-Enabled Smart Greenhouse for Sustainable Agriculture," *IEEE Consumer Electronics Magazine*, vol. 10, no. 4, pp. 57–62, 2021, doi: 10.1109/MCE.2021.3055930.

[4] P. Mohanty, U. C. Pati, K. Mahapatra, and S. P. Mohanty, "bslight 2.0: Battery-free Sustainable Smart Street Light Management System," *IEEE Transactions on Sustainable Computing*, pp. 1–15, 2024, doi: 10.1109/TSUSC.2024.3408630.

[5] American Council for an Energy-Efficient Economy (ACEEE), "Energy efficiency in corporate sustainability: Linking sustainability to specific actions," ACEEE, Washington, DC, USA, Tech. Rep., [Online; accessed 25-Aug-2025]. [Online]. Available: https://www.aceee.org/sites/default/files/pdfs/eecs-smmcg_0.pdf

[6] M. A. Fekih, W. Bechkit, H. Rivano, M. Dahan, F. Renard, L. Alonso, and F. Pineau, "Participatory Air Quality and Urban Heat Islands Monitoring System," *IEEE Transactions on Instrumentation and Measurement*, vol. 70, pp. 1–14, 2021, doi: 10.1109/TIM.2020.3034987.

[7] K. Bashir Shaban, A. Kadri, and E. Rezk, "Urban Air Pollution Monitoring System With Forecasting Models," *IEEE Sensors Journal*, vol. 16, no. 8, pp. 2598–2606, 2016, doi: 10.1109/JSEN.2016.2514378.

[8] A. R. Al-Ali, I. Zualkernan, and F. Aloul, "A Mobile GPRS-Sensors Array for Air Pollution Monitoring," *IEEE Sensors Journal*, vol. 10, no. 10, pp. 1666–1671, 2010, doi: 10.1109/JSEN.2010.2045890.

[9] T. Becnel, K. Tingey, J. Whitaker, T. Sayahi, K. Lê, P. Goffin, A. Butterfield, K. Kelly, and P.-E. Gaillardon, "A Distributed Low-Cost Pollution Monitoring Platform," *IEEE Internet of Things Journal*, vol. 6, no. 6, pp. 10 738–10 748, 2019, doi: 10.1109/JIOT.2019.2941374.

[10] S. Ali, T. Glass, B. Parr, J. Potgieter, and F. Alam, "Low Cost Sensor With IoT LoRaWAN Connectivity and Machine Learning-Based Calibration for Air Pollution Monitoring," *IEEE Transactions on Instrumentation and Measurement*, vol. 70, pp. 1–11, 2021, doi: 10.1109/TIM.2020.3034109.

[11] B. Montrucchio, E. Giusto, M. G. Vakili, S. Quer, R. Ferrero, and C. Fornaro, "A Densely-Deployed, High Sampling Rate, Open-Source Air Pollution Monitoring WSN," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 12, pp. 15 786–15 799, 2020, doi: 10.1109/TVT.2020.3035554.

[12] P. Das, S. Ghosh, S. Chatterjee, and S. De, "A Low Cost Outdoor Air Pollution Monitoring Device With Power Controlled Built-In PM Sensor," *IEEE Sensors Journal*, vol. 22, no. 13, pp. 13 682–13 695, 2022, doi:10.1109/JSEN.2022.3175821.

[13] K. Hu, V. Sivaraman, B. G. Luxan, and A. Rahman, "Design and Evaluation of a Metropolitan Air Pollution Sensing System," *IEEE Sensors Journal*, vol. 16, no. 5, pp. 1448–1459, 2016, doi: 10.1109/JSEN.2015.2499308.

[14] D. Zhang and S. S. Woo, "Real Time Localized Air Quality Monitoring and Prediction Through Mobile and Fixed IoT Sensing Network," *IEEE Access*, vol. 8, pp. 89 584–89 594, 2020, doi: 10.1109/ACCESS.2020.2993547.

[15] Q. Han, P. Liu, H. Zhang, and Z. Cai, "A Wireless Sensor Network for Monitoring Environmental Quality in the Manufacturing Industry," *IEEE Access*, vol. 7, pp. 78 108–78 119, 2019, doi: 10.1109/ACCESS.2019.2920838.

[16] X. Yue, M. Kauer, M. Bellanger, O. Beard, M. Brownlow, D. Gibson, C. Clark, C. MacGregor, and S. Song, "Development of an Indoor Photovoltaic Energy Harvesting Module for Autonomous Sensors in Building Air Quality Applications," *IEEE Internet of Things Journal*, vol. 4, no. 6, pp. 2092–2103, 2017, doi: 10.1109/JIOT.2017.2754981.

[17] W. A. Jabbar, T. Subramaniam, A. E. Ong, M. I. Shu'Ib, W. Wu, and M. A. de Oliveira, "LoRaWAN-Based IoT System Implementation for Long-Range Outdoor Air Quality Monitoring," *Internet of Things*, vol. 19, p. 100540, 2022. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2542660522000427

[18] A. C. Riekstin, A. Langevin, T. Dandres, G. Gagnon, and M. Cheriet, "Time Series-Based GHG Emissions Prediction for Smart Homes," *IEEE Transactions on Sustainable Computing*, vol. 5, no. 1, pp. 134–146, 2020, doi: 10.1109/TSUSC.2018.2886164.

[19] S. Dey, K. Chatterjee, R. P. Kumar, A. Bandyopadhyay, S. Swain, and N. Kumar, "Apict: Air Pollution Epidemiology Using Green AQI Prediction During Winter Seasons in India," *IEEE Transactions on Sustainable Computing*, vol. 9, no. 3, pp. 559–570, 2024, doi: 10.1109/TSUSC.2023.3343922.

[20] T. Anagnostopoulos, A. Zaslavsky, K. Kolomvatsos, A. Medvedev, P. Amirian, J. Morley, and S. Hadjieftymiades, "Challenges and Opportunities of Waste Management in IoT-Enabled Smart Cities: A Survey," *IEEE Transactions on Sustainable Computing*, vol. 2, no. 3, pp. 275–289, 2017.

[21] Q. Yang, X. Zhang, Q. Luan, S. Sun, Y. Zhao, C. Fang, X. Mi, and M. Li, "Optimization of Hourly PM2.5 Inversion Model Integrating Upper-Air Meteorological Elements," *IEEE Access*, vol. 11, pp. 35 421–35 428, 2023, doi: 10.1109/ACCESS.2023.3263790.

[22] F. Naz, C. Mccann, M. Fahim, T.-V. Cao, R. Hunter, N. T. Viet, L. D. Nguyen, and T. Q. Duong, "Comparative Analysis of Deep Learning and Statistical Models for Air Pollutants Prediction in Urban Areas," *IEEE Access*, vol. 11, pp. 64 016–64 025, 2023, doi: 10.1109/ACCESS.2023.3289153.

[23] Q. Shao, J. Chen, and T. Jiang, "A Novel Coupled Optimization Prediction Model for Air Quality," *IEEE Access*, vol. 11, pp. 69 667–69 685, 2023, doi: 10.1109/ACCESS.2023.3293249.

[24] E. Sharma, R. C. Deo, R. Prasad, A. V. Parisi, and N. Raj, "Deep Air Quality Forecasts: Suspended Particulate Matter Modeling With Convolutional Neural and Long Short-Term Memory Networks," *IEEE Access*, vol. 8, pp. 209 503–209 516, 2020, doi: 10.1109/ACCESS.2020.3039002.

[25] S. Yelisetti, V. K. Saini, R. Kumar, R. Lamba, and A. Saxena, "Uncertainty Aware Learning Model for Thermal Comfort in Smart Residential Buildings," *IEEE Transactions on Industry Applications*, vol. 60, no. 1, pp. 1909–1918, 2024, doi: 10.1109/TIA.2023.3328574.

[26] L. Lin, C.-Y. Chen, H.-Y. Yang, Z. Xu, and S.-H. Fang, "Dynamic System Approach for Improved PM2.5 Prediction in Taiwan," *IEEE Access*, vol. 8, pp. 210 910–210 921, 2020, doi: 10.1109/ACCESS.2020.3038853.

[27] H. A. D. Nguyen, Q. P. Ha, H. Duc, M. Azzi, N. Jiang, X. Barthelemy, and M. Riley, "Long Short-Term Memory Bayesian Neural Network for Air Pollution Forecast," *IEEE Access*, vol. 11, pp. 35 710–35 725, 2023, doi: 10.1109/ACCESS.2023.3265725.

[28] K. Gu, H. Liu, J. Liu, X. Yu, T. Shi, and J. Qiao, "Air Pollution Prediction in Mass Rallies With a New Temporally-Weighted Sample-Based Multitask Learner," *IEEE Transactions on Instrumentation and Measurement*, vol. 71, pp. 1–15, 2022, doi: 10.1109/TIM.2022.3157365.

[29] D. Iskandaryan, F. Ramos, and S. Trilles, "Air Quality Prediction in Smart Cities Using Machine Learning Technologies Based on Sensor Data: A Review," *Applied Sciences*, vol. 10, no. 7, 2020. [Online]. Available: https://www.mdpi.com/2076-3417/10/7/2401

[30] N. S. Gupta, Y. Mohta, K. Heda, R. Armaan, B. Valarmathi, and G. Arulkumaran, "Prediction of Air Quality Index Using Machine Learning Techniques: a Comparative Analysis," *Journal of Environmental and Public Health*, vol. 2023, no. 1, p. 4916267, 2023.

[31] Y. Wei, J. Jang-Jaccard, W. Xu, F. Sabrina, S. Camtepe, and M. Boulic, "LSTM-Autoencoder-Based Anomaly Detection for Indoor Air Quality Time-Series Data," *IEEE Sensors Journal*, vol. 23, no. 4, pp. 3787–3800, 2023.

[32] M. G. Ragab, S. J. Abdulkadir, N. Aziz, Q. Al-Tashi, Y. Alyousifi, H. Alhussian, and A. Alqushaibi, "A Novel One-Dimensional CNN with Exponential Adaptive Gradients for Air Pollution Index Prediction," *Sustainability*, vol. 12, no. 23, 2020. [Online]. Available: https://www.mdpi.com/2071-1050/12/23/10090

[33] Q. Zhang, Y. Han, V. O. K. Li, and J. C. K. Lam, "Deep-AIR: A Hybrid CNN-LSTM Framework for Fine-Grained Air Pollution Estimation and Forecast in Metropolitan Cities," *IEEE Access*, vol. 10, pp. 55 818–55 841, 2022.

[34] G. Sharma, A. M. Joshi, and S. P. Mohanty, "strade 2.0: Efficient Mutual Authentication Scheme for Energy Trading in V2G Using Physically Unclonable Function," *IEEE Journal of Emerging and Selected Topics in Industrial Electronics*, pp. 1–11, 2024, doi: 10.1109/JESTIE.2024.3384921.

[35] D. Puthal, S. P. Mohanty, P. Nanda, and U. Choppali, "Building Security Perimeters to Protect Network Systems Against Cyber Threats [Future Directions]," *IEEE Consumer Electronics Magazine*, vol. 6, no. 4, pp. 24–27, 2017, doi: 10.1109/MCE.2017.2714744.

Lopamudra Samal Received B.Tech degree in Electronics and Telecommunication Engineering from BPUT, Rourkela, and M.Tech degree in VLSI and Embedded Systems Design, from National Institute of Technology Rourkela, in 2012. She is currently pursuing Ph. D. in Electronics and Communication Engineering from NIT Rourkela. Her research interests include Cyber-Physical System (CPS) Security, IIoT Security, Edge Computing Security, Attack Detection and Mitigation and Machine Learning.

Aryan Kumar Samal He received the bachelor's degree in Electrical and Electronics Engineering from the International Institute of Information Technology, Bhubaneswar, in 2017, and the master's degree in Electronics and Instrumentation Engineering from National Institute of Technology Rourkela, in 2023. He is currently with Tata Consultancy Services, Kolkata. His research interests include Internet of things and ML.

Kamalakanta Mahapatra (Senior Member, IEEE) Dr. Kamalakanta Mahapatra is a Professor (HAG) in Electronics and Communication Engineering Department of National institute of Technology, Rourkela. He assumed professor position since February 2004. He obtained his B. Tech degree with Honours from Regional Engineering College, Calicut in 1985, Master's from Regional Engineering College, Rourkela in 1989 and Ph. D. from IIT Kanpur in 2000. He is a senior member of the IEEE and a fellow of the institution of Engineers (India) in ECE Division. Presently , he has served as the Chairman of IEEE Rourkela Sub-section and mentor of IEEE CTSoC chapter of Kolkata section. He has published several research papers in National and International Journals. He received coveted J. C. Bose award for best engineering oriented research in the year 2014. His research interests include Embedded Computing Systems, VLSI Design, Hardware Security and Industrial/Consumer/Power Electronics. He has supervised 24 PhD dissertations and 92 Master's thesis.

Ayas Kanta Swain (Member, IEEE) received the bachelor's degree (Hons.) in Electrical Engineering from the Indira Gandhi Institute of Technology, Utkal University, Bhubaneswar, in 2001, the master's degree (research) in Electronics and Communication Engineering from National Institute of Technology Rourkela, in 2009, and the Ph.D. degree in Electronics and Communication Engineering from the National Institute of Technology Rourkela. He is an Assistant Professor with NIT Rourkela. His research area includes VLSI design, secured network-on-chip design, and embedded systems.

Saraju P. Mohanty (Senior Member, IEEE) received the bachelor's degree (Honors) in electrical engineering from the Orissa University of Agriculture and Technology, Bhubaneswar, in 1995, the master's degree in Systems Science and Automation from the Indian Institute of Science, Bengaluru, in 1999, and the Ph.D. degree in Computer Science and Engineering from the University of South Florida, Tampa, in 2003. He is a Professor with the University of North Texas. His research is in "Smart Electronic Systems" which has been funded by National Science Foundations (NSF), Semiconductor Research Corporation (SRC), U.S. Air Force, IUSSTF, and Mission Innovation. He has authored 550 research articles, 5 books, and invented 10 granted/pending patents. His Google Scholar h-index is 58 and i10-index is 269 with 15,000 citations. He is a recipient of 19 best paper awards, Fulbright Specialist Award in 2021, IEEE Consumer Electronics Society Outstanding Service Award in 2020, the IEEE-CS-TCVLSI Distinguished Leadership Award in 2018, and the PROSE Award for Best Textbook in Physical Sciences and Mathematics category in 2016. He has delivered 30 keynotes and served on 15 panels at various International Conferences. He has been the Editor-in-Chief of the IEEE Consumer Electronics Magazine during 2016-2021 and currently serves on the editorial board of 8 journals/transactions. He has supervised 3 post-doctoral researchers, 17 Ph.D. dissertations, 28 M.S. theses, and 28 undergraduate projects.