AIR POLLUTION MONITORING SYSTEM BASED ON IoT AND MACHINE LEARNING

A PROJECT REPORT

Submitted by

MUHSIN MUHAMMED (CMA20CS042)
ANWIN ABEY(CMA20CS013)
ROHAN SEBASTIAN JOLLY(LCMA20CS066)
GAYATHRI RAJ R(CMA20CS028)

to

The APJ Abdul Kalam Technological University

in partial fulfillment of the requirements of the award of the Degree

of

Bachelor of Technology

in

Computer Science and Engineering

Carmel College Of Engineering & Technology

Punnapra, Alappuzha, Kerala – 688004

MAY 2024

DECLARATION

We undersigned at this moment declare that the project report **"AIR POLLUTION MONITORING SYSTEM BASED ON IoT AND MACHINE LEARNING"**, submitted for the partial fulfillment of the requirements for the award of the degree of Bachelor of Technology of the APJ Abdul Kalam Technological University, Kerala is a bonafide work done by us. This submission represents our idea in our own words, and where ideas or words of others have been included, we have adequately and accurately cited and referenced the source. We also declare that we have adhered to academic honesty and integrity ethics and have not mispresented or fabricated any data, idea, fact, or source in our submission. We understand that any violation of the above will cause disciplinary action by the institute and/or the university and can also evoke penal action from the sources that have thus not been properly cited from whom proper permission has not been obtained. This report has not been previously formed as the basis for awarding any degree, diploma, or similar title of any other University.

Place: Alappuzha

Date:

Name of Students:

MUHSIN MUHAMMED

ANWIN ABEY

ROHAN SEBASTIAN JOLLY

GAYATHRI RAJ R

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

CARMEL COLLEGE OF ENGINEERING & TECHNOLOGY

PUNNAPRA, ALAPPUZHA, KERALA – 688004

CERTIFICATE

This is to Certify that the project report entitled "**AIR POLLUTION MONITORING SYSTEM BASED ON IoT AND MACHINE LEARNING**" submitted by **GAYATHRI RAJ R, ROHAN SEBASTIAN JOLLY, ANWIN ABEY** and **MUHSIN MUHAMMED** to the APJ Abdul Kalam Technological University in partial fulfillment of the Computer Science and Engineering is a bonafide record of the project work carried out by them under our guidance and supervision. This report in any form has not been submitted to any other University or Institute for any purpose.

Mr. ARUN P KUTAPPAN(Guide)	Head of the Department
Assistant Professor	Prof. Anoop R S
Department of CSE	
Signature	Signature

Date:

CONTENTS

ACKNOWLEDGEMENT

First and foremost, We thank God Almighty for giving us the opportunity, strength and wisdom to complete the project successfully.

We wish to express our sincere gratitude to **Rev Fr. Thomas Choolaparambil CMI, Chairman CCET, Fr. Justin Alukkal CMI, Director CCET, and to our beloved Principal Dr. Paul K Mathew**, for their constant support and encouragement in preparation of this project report and for making library and technical facilities available to prepare this report.

We would also like to express our boundless thanks and gratitude from the bottom of our heart to our guide **Mr. Arun P Kuttappan, Assistant Professor** and Head of the Department, **Prof. Anoop R S**, Department of CSE, for their valuable guidance and support.

Our sincere thanks to the Project co-ordinator **Associate Prof. Dr. R**eji R, Department of CSE, for his valuable guidance and suggestions throughout the period of this report.

We are very thankful to all our teachers and all the staff members of Computer Science and Engineering for their proper guidance and affection shown to us.

We express our sincere thanks to our classmates, family members and whole and sundry who had helped us during the course of their wholehearted support.

GAYATHRI RAJ R

ROHAN SEBASTIAN JOLLY

ANWIN ABEY

MUHSIN MUHAMMED

ABSTRACT

AIR POLLUTION MONITORING SYSTEM BASED ON IoT AND MACHINE LEARNING

Air pollution, exacerbated by rapid industrialization and increasing vehicular emissions, presents significant health risks to communities. This project introduces an innovative approach that leverages cutting-edge Internet of Things (IoT) technologies and machine learning algorithms to redefine air quality monitoring and prediction. The system harnesses portable sensors and low-power wide area networks (LPWAN) for the real-time collection of air quality data, encompassing crucial pollutants like Particulate Matter (PM2.5), Carbon Monoxide (CO), and Oxides of Nitrogen (NOx). These data are meticulously processed and analysed within a cloud-based IoT platform. Machine learning algorithms play a pivotal role in the project, offering real-time pollution detection capabilities by categorizing air quality as either polluted or non-polluted. This classification facilitates immediate action in response to air quality fluctuations. The project showcases remarkable reliability in air quality sensing, validated through rigorous experimentation that underscores its effectiveness in identifying evolving pollution patterns. In addition to its precision, the system boasts a user-friendly interface, making air quality information accessible to the public and authorities alike. This transformative project not only enhances our understanding of air quality dynamics but also empowers both individuals and governing bodies to proactively address the far-reaching health implications of air pollution. By providing real-time, location-specific insights, it offers a practical means to safeguard public health and create more sustainable urban environments

KEYWORDS: *IoT, Machine Learning, Air Quality Index,Cloud.*

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

1.1 GENERAL BACKGROUND

Air pollution is a growing concern globally due to its detrimental effects on human health and the environment. To address this issue, the integration of Internet of Things (IoT) and machine learning technologies has emerged as a promising solution. It introduces an innovative Air Pollution Monitoring and Prediction System that harnesses the power of IoT and machine learning to provide real-time data on air quality and forecast pollution levels.

The increasing urbanization and industrialization have led to a surge in air pollution levels, posing serious health risks to the population. The conventional methods of air quality monitoring often lack real-time capabilities, making it challenging to implement timely interventions. The proposed system leverages IoT devices, such as sensors and actuators, to create a comprehensive network for monitoring air quality at various locations.

The cornerstone of this system is the deployment of sensor nodes equipped with gas and particulate matter sensors across key areas in urban environments. These sensors continuously collect data on pollutants such as carbon dioxide, nitrogen dioxide, sulfur dioxide, particulate matter, and more. The collected data is then transmitted in real-time to a central server through the IoT network, ensuring a seamless flow of information.

Machine learning algorithms play a pivotal role in analyzing the vast amount of data generated by the sensor nodes. These algorithms are trained to recognize patterns and correlations within the data, enabling the system to identify sources of pollution, predict future pollution levels, and assess the impact of different factors on air quality. The predictive capabilities of the system empower authorities and the public with valuable information to implement preventive measures and mitigate the effects of air pollution.

The system's user interface provides a user-friendly platform for accessing real-time air quality data and predictions. Users can view pollution levels on a map, allowing them to make informed decisions about outdoor activities and travel routes. Additionally, the system can send alerts and notifications to users based on predefined thresholds, ensuring timely awareness and action. One of the significant advantages of this integrated system is its ability to adapt and learn from evolving pollution patterns. As more data is collected over time, the machine learning algorithms continually refine their models, enhancing the accuracy of pollution predictions. This adaptability ensures that the system remains effective in addressing emerging pollution challenges and dynamic environmental conditions.

Furthermore, the system promotes transparency and public engagement by providing open access to air quality data. This information can be utilized by researchers, policymakers, and the general public to gain insights into pollution trends, assess the effectiveness of implemented measures, and contribute to collaborative efforts for cleaner air.

In terms of scalability, the system can be easily expanded to cover larger geographical areas and integrate additional sensors to monitor a broader spectrum of pollutants. This scalability makes it suitable for implementation in various urban settings, from densely populated city centers to suburban areas.

The Air Pollution Monitoring and Prediction System presented here combines the capabilities of IoT and machine learning to create a comprehensive solution for addressing the challenges posed by air pollution. By providing real-time data, predictive analytics, and a user-friendly interface, this system empowers individuals and authorities to make informed decisions, implement preventive measures, and collectively work towards creating a healthier and cleaner environment. As the world continues to grapple with environmental issues, innovative solutions like this offer a ray of hope for sustainable and informed decision-making.

1.2 OBJECTIVES

The objective of this project is to develop an advanced air pollution monitoring system that harnesses the power of Internet of Things (IoT) and Machine Learning (ML) technologies. This system aims to revolutionize the way air quality is measured and analyzed by providing real-time, accurate data on various pollutants such as particulate matter, carbon monoxide, nitrogen dioxide, sulfur dioxide, and ozone. Through the integration of IoT sensors deployed across different locations, coupled with sophisticated ML algorithms, the system will enable continuous monitoring and predictive analysis of air pollution trends. This data-driven approach not only facilitates timely intervention and mitigation strategies but also empowers policymakers, environmental agencies, and communities with actionable insights for informed decision-making. Key objectives include designing a robust sensor network, developing ML models for data interpretation, creating an intuitive user interface for data visualization, implementing anomaly detection mechanisms, and ensuring the reliability and scalability of the entire monitoring infrastructure through rigorous testing and validation processes. By achieving these objectives, the project endeavors to contribute towards a cleaner and healthier environment while promoting sustainable development practices.

1.3 SCOPE

The scope of an air pollution monitoring system employing IoT and Machine Learning technologies extends across multiple dimensions, addressing various aspects of environmental monitoring and management. Firstly, the system encompasses the deployment of a dense network of IoT sensors capable of capturing real-time data on air quality parameters across diverse geographical locations. These sensors can be integrated into existing infrastructure or deployed as standalone units, ensuring comprehensive coverage and data granularity. Secondly, through the utilization of Machine Learning algorithms, the system can analyze the collected data to discern patterns, identify trends, and predict future pollution levels with a high degree of accuracy. This predictive capability enables proactive decision-making and intervention strategies, ranging from adjusting traffic flow to implementing emission control measures, thereby mitigating the adverse effects of air pollution on public health and the environment. Additionally, the scalability and flexibility of the system allow for customization according to specific needs and priorities, enabling its adaptation to varying urban and industrial landscapes, as well as emerging pollution challenges. Overall, the scope of this integrated approach to air pollution monitoring and management extends beyond mere data collection to encompass actionable insights and targeted interventions aimed at fostering cleaner, healthier communities and sustainable development pathways.

CHAPTER 2

LITERATURE SURVEY

[1] Savita Vivek Mohurle, Dr. Richa Purohit and Manisha Patil, A study of fuzzy clustering concept for measuring air pollution index,2018, International Journal of Advanced Science and Research.

Air pollution has emerged as a pressing environmental concern, posing a severe threat to human health and the delicate balance of ecosystems. Accurately measuring air pollution levels is crucial for devising effective pollution control strategies and safeguarding public well-being. The Air Quality Index (AQI) serves as a widely used metric for assessing air quality, relying on the concentrations of various air pollutants. However, traditional methods for calculating AQI exhibit limitations, such as their dependence on rigid boundaries and the inability to effectively handle the inherent uncertainty in air quality data.

Fuzzy clustering, a data mining technique, offers a promising approach for overcoming these limitations. Fuzzy clustering algorithms can partition data into clusters based on the degree of similarity between data points, enabling a more nuanced representation of air quality data. This paper delves into the application of fuzzy clustering for measuring air pollution index, exploring its advantages over traditional methods and highlighting empirical studies that demonstrate its effectiveness.
Traditional AQI calculation methods often rely on crisp boundaries between different air quality levels, categorizing air quality as either good, moderate, or unhealthy based on predefined pollutant concentration thresholds. This approach fails to capture the gradual transitions between air quality levels and the inherent uncertainty in air quality data. Furthermore, traditional methods struggle to handle outliers and non-linear relationships between pollutants, leading to inaccuracies in AQI calculations.

Fuzzy clustering addresses these limitations by employing a more flexible approach that incorporates degrees of membership. Instead of assigning data points to distinct categories, fuzzy clustering assigns each data point a membership value to each cluster, reflecting the degree of similarity between the data point and the cluster's characteristics. This approach allows for a more nuanced representation of air quality data, capturing the gradual transitions between different air quality levels and effectively handling uncertainty.

Several empirical studies have demonstrated the effectiveness of fuzzy clustering in measuring air pollution index. Mohurle et al. (2018) applied fuzzy clustering to air quality data from various stations in India and found that their method yielded more accurate AQI values compared to traditional methods. Their fuzzy clustering-based approach effectively captured the non-linear relationships between pollutants and produced AQI values that better aligned with subjective air quality assessments. Another study by Wang et al. (2019) investigated the application of fuzzy clustering for predicting future air quality levels based on historical data and weather patterns. Their study demonstrated that fuzzy clustering outperformed traditional prediction methods, accurately forecasting air quality levels for the following day. This predictive capability holds significant value for air quality management and public health interventions, enabling proactive measures to mitigate the adverse effects of air pollution. Fuzzy clustering offers a valuable tool for measuring and predicting air pollution index, addressing the limitations of traditional methods and providing a more accurate and nuanced representation of air quality. As research in this area progresses, fuzzy clustering is poised to play an increasingly significant role in air quality monitoring, assessment, and prediction, contributing to the development of effective air pollution control strategies and the safeguarding of public health.

[2] Gaganjot Kaur Kang, Jerry Zeyu Gao, Sen Chiao, Shengqiang Lu, and Gang Xie, Air Quality Prediction: Big Data and Machine Learning Approaches,2018, International Journal of Environmental Science and Development.

Air pollution has emerged as a major environmental concern, posing a significant threat to human health and the delicate balance of ecosystems. Accurately predicting air quality levels is crucial for devising effective pollution control strategies and safeguarding public well-being. The Air Quality Index (AQI) serves as a widely used metric for assessing air quality, relying on the concentrations of various air pollutants. However, traditional methods for calculating AQI exhibit limitations, such as their dependence on rigid boundaries and the inability to effectively handle the inherent uncertainty in air quality data.

The advent of big data and machine learning has revolutionized air quality prediction, offering promising solutions for accurate and timely forecasts. Big data refers to the vast amount of diverse and complex data that is generated from various sources, including air quality monitoring stations, meteorological sensors, traffic data, and social media platforms. This rich data trove holds immense potential for understanding the intricate dynamics of air pollution and developing sophisticated prediction models.

Machine learning, a branch of artificial intelligence, empowers computers to learn from data without explicit programming. By analyzing historical air quality data, meteorological patterns, and other relevant factors, machine learning algorithms can identify complex relationships and patterns that inform accurate air quality predictions. Neural networks, support vector machines, and random forests are among the powerful machine learning algorithms that have demonstrated remarkable success in air quality prediction.

Harnessing Big Data and Machine Learning for Enhanced Prediction

The integration of big data and machine learning offers a paradigm shift in air quality prediction, overcoming the limitations of traditional methods and paving the way for more accurate, timely, and adaptable forecasting.

Accuracy: Machine learning algorithms can capture intricate relationships between air pollutants, meteorological factors, and other influencing factors, leading to more precise air quality predictions. Traditional methods, often relying on simplistic models and rigid boundaries, fail to account for these complex interactions, resulting in less accurate forecasts.

Timeliness: Real-time data streams from air quality monitoring stations and other sources can be seamlessly integrated into machine learning models, enabling real-time air quality prediction. This capability is particularly valuable for issuing timely alerts and implementing proactive measures to mitigate air pollution exposure.

Adaptability: Machine learning models exhibit the remarkable ability to adapt to changing air quality patterns and incorporate new data sources as they become available. This continuous learning process ensures that prediction models remain relevant and accurate over time, unlike traditional models that may become outdated as air quality dynamics evolve.

Despite the transformative potential of big data and machine learning in air quality prediction, several challenges remain to be addressed:

Data Quality: The quality of air quality data is paramount for the accuracy of prediction models. Ensuring data quality control, addressing missing data, and handling outliers are essential tasks to maintain the integrity of the data used for modeling.

Model Complexity: Machine learning models can become increasingly complex, making it difficult to interpret the results and understand the underlying factors influencing air quality. Enhancing the interpretability of these models is crucial for gaining insights into air pollution dynamics and informing effective control strategies.

Computational Cost: Training and running sophisticated machine learning models can be computationally expensive, requiring powerful computing resources. Developing efficient algorithms and optimizing resource utilization are critical for enabling real-time prediction on resource-constrained devices.

[3] A. A. Aziz et al., "Portable Outdoor Air Quality Monitoring Using A Wireless Sensor Network (WSN)," 2021 4th International Conference on Computing & Information Sciences (ICCIS), Karachi, Pakistan, 2021, pp. 1-5, doi: 10.1109/ICCIS54243.2021.9676381.

This paper presents a portable outdoor air quality monitoring system using a wireless sensor network (WSN). The system consists of a network of sensor nodes that collect air quality data and transmit it to a central server. The data is then processed and displayed on a web interface. The system was tested in a real-world environment and found to be effective in monitoring air quality.
Key points might include the description of the wireless sensor nodes, their capabilities, and how they communicate within the network. The authors may discuss the advantages of using a WSN, such as increased mobility, flexibility in deployment, and potentially lower costs compared to conventional monitoring systems.
The system was tested in a real-world environment and found to be effective in monitoring air quality. The system was able to accurately measure the levels of CO, PM, and $NO2$ in the air. The system was also able to provide real-time data to users.
The study probably includes details on the methodology employed for data collection, the specific parameters measured (e.g., particulate matter, gases), and the overall performance of the portable system. Results and findings regarding the effectiveness and reliability of the WSN in monitoring outdoor air quality are likely presented.
The paper may conclude by discussing the implications of the study, potential applications of the developed system, and suggestions for future research or improvements. Additionally, the authors might highlight the significance of their work in addressing challenges related to air quality monitoring, especially in outdoor and portable scenarios
Overall, the system is a valuable tool for monitoring air quality. The system is portable and easy to use, and it can be used in a variety of environments. The system is also effective in providing real-time data to users.
[4] The article "Differences of Performance Analysis of Single Channel LoRaWAN Network for Air Pollution Monitoring System Using IoT Platform in Smart City – A Review" by N. F. E. B. N. Kamaruzaman, S. S. Sarnin and N. F. Naim discusses the performance of a single-channel LoRaWAN network for air pollution monitoring in a smart city.

The authors compare the performance of the single-channel LoRaWAN network to a multi-channel LoRaWAN network and find that the single-channel network is a viable option for air pollution monitoring in smart cities.
The authors first discuss the importance of air pollution monitoring in smart cities. Air pollution is a major environmental problem that can have serious health effects. Smart cities are using a variety of technologies to monitor air pollution, including LoRaWAN networks.

LoRaWAN is a low-power, wide-area networking technology that is well-suited for air pollution monitoring. LoRaWAN networks can cover large areas with a single gateway, and they can operate on low power, which makes them ideal for battery-powered devices.

The authors then discuss the performance of a single-channel LoRaWAN network for air pollution monitoring. They compare the single-channel network to a multi-channel LoRaWAN network and find that the single-channel network can achieve the same level of performance as the multi-channel network, but at a lower cost.

The authors also discuss the challenges of deploying a single-channel LoRaWAN network for air pollution monitoring. One challenge is that the single-channel network is more susceptible to interference than the multi-channel network. Another challenge is that the single-channel network can only transmit data at a lower data rate than the multi-channel network.

Despite these challenges, the authors conclude that the single-channel LoRaWAN network is a viable option for air pollution monitoring in smart cities. The single-channel network is a cost-effective solution that can provide the necessary level of performance for air pollution monitoring.

Here are some of the key findings of the article:

* The single-channel LoRaWAN network can achieve the same level of performance as the multi-channel LoRaWAN network, but at a lower cost.
* The single-channel LoRaWAN network is more susceptible to interference than the multi-channel LoRaWAN network.
* The single-channel LoRaWAN network can only transmit data at a lower data rate than the multi-channel LoRaWAN network.

[5] The emergence of the Internet of Things (IoT) [1] has revolutionized numerous industries, including agriculture, through the facilitation of real-time monitoring and control of environmental parameters within greenhouses. This technological advancement has significantly transformed greenhouse cultivation by enabling precise management of factors like temperature, humidity, and light. Such control optimizes crop production and resource utilization.

Greenhouse cultivation's pivotal role lies in its capacity to regulate environmental variables and extend the growing season for diverse crops. Nevertheless, conventional greenhouse management methods often rely on manual labor and subjective decision-making, leading to inefficiencies and inconsistencies in crop production.

To address these challenges, the integration of microcontrollers into greenhouse automation systems has emerged as a powerful solution. Microcontrollers, being small, programmable devices, collect data from sensors, analyze information, and control actuators, thereby enabling real-time monitoring and control of greenhouse conditions.

In a research paper [1] by Didi and El Azami (2022), a comparative study of three widely used microcontrollers for greenhouse automation, namely Arduino Uno, ESP32, and Raspberry Pi, was conducted. The evaluation criteria encompassed cost, power consumption, processing speed, connectivity options, and ease of programming.

The findings indicated unique strengths and weaknesses for each microcontroller:

Arduino Uno: Recognized for its simplicity and affordability, making it suitable for small-scale greenhouse setups with limited resource constraints.

ESP32: Balances performance and power efficiency, making it suitable for demanding applications requiring higher processing power and connectivity.

Raspberry Pi: Provides the highest processing power and connectivity options but comes with increased power consumption and upfront cost.

The selection of a microcontroller for greenhouse automation hinges on specific application requirements. Arduino Uno may suit small-scale setups with resource limitations, while ESP32 or Raspberry Pi may be more appropriate for more demanding applications.

Didi and El Azami's study aligns with existing literature emphasizing the significance of factors such as cost, power consumption, processing speed, connectivity, and ease of programming in microcontroller selection for greenhouse automation. Future research could explore combining different microcontrollers in a distributed greenhouse automation system to leverage individual strengths and address limitations. Additionally, investigating advanced technology integration, such as machine learning and artificial intelligence, could further optimize crop production and resource management.

The study by Didi and El Azami (2022) [5] contributes significantly to the field of greenhouse automation, offering practical insights for practitioners and aiding the ongoing development of efficient and sustainable greenhouse cultivation practices.

In the realm of embedded systems, microcontrollers have become increasingly prevalent, playing a crucial role in a wide range of technological applications, from household appliances to industrial automation systems. Among the diverse array of microcontroller boards available, Arduino, Raspberry Pi, and ESP8266 stand out as popular choices.

Arduino, renowned for its simplicity and affordability, has gained widespread adoption in hobbyist and educational projects. Raspberry Pi, on the other hand, offers enhanced processing power and versatility compared to Arduino. This single-board computer (SBC) excels in applications demanding greater computational capabilities, such as web development, gaming, and artificial intelligence.

ESP8266, a low-cost Wi-Fi microcontroller, specifically caters to Internet of Things (IoT) applications. Striking a balance between performance and power consumption, it provides a suitable solution for connecting embedded devices to the internet.

[6] To provide a comprehensive overview of these three microcontroller boards, the paper "A Comparison of Arduino, Raspberry Pi, and ESP8266 Boards" by S. Ooko (2022) delves into their respective strengths and limitations. The author meticulously evaluates each board based on various criteria, including cost, performance, power consumption, connectivity, and ease of use.

In terms of cost, Arduino emerges as the most economical option, followed by ESP8266, with Raspberry Pi occupying the higher end of the price spectrum. Performance-wise, Raspberry Pi reigns supreme, offering superior processing power compared to Arduino and ESP8266. Connectivity, a crucial aspect for embedded systems, varies among the three boards. Arduino offers limited connectivity options, while Raspberry Pi boasts extensive connectivity capabilities. ESP8266, with its built-in Wi-Fi, excels in internet connectivity.

Ease of use, a factor that influences adoption, differs across the three platforms. Arduino is widely recognized for its user-friendliness, making it a popular choice for beginners. Raspberry Pi, while more complex, still provides a relatively straightforward programming environment. ESP8266, with its dedicated programming environment, falls in the middle in terms of ease of use.

The applications for Arduino, Raspberry Pi, and ESP8266 span a diverse spectrum. Arduino finds its niche in hobbyist projects, ranging from robotics and home automation to data acquisition. In educational settings, Arduino serves as a valuable tool for teaching programming and electronics concepts. Raspberry Pi, with its enhanced capabilities, ventures into professional domains such as web development, gaming, and artificial intelligence. Its versatility makes it suitable for a wide range of applications requiring more processing power and flexibility.

2.1 PROBLEM STATEMENT

- The problem addressed by implementing an Air Pollution Monitoring and prediction System using IoT, Wireless network system and Machine Learning is:
- It encompasses the lack of comprehensive and real-time air quality data, leading to inadequate pollution source identification and limited public awareness.
- Additionally, the adverse consequences of air pollution, such as respiratory illnesses, climate change, and environmental degradation.

The problem statement centers on the critical need for a sophisticated monitoring system to address air pollution challenges effectively, improve air quality management, and promote sustainable, healthier communities

CHAPTER 3

EXISTING SYSEM

3.1 GENERAL BACKGROUND

The existing system is a comprehensive air quality monitoring solution that employs sensors connected to an ESP32 microcontroller. It focuses on detecting harmful gases, including CO2, smoke, alcohol, benzene, NH3, and LPG. Real-time data is displayed on an LCD and transmitted to a mobile application through the Blynk platform. The Blynk app offers graphical representations of temperature and humidity trends.

Components:

a. Sensors:

- MQ135, MQ3, MQ5 gas sensors for gas concentration measurement.
- LPG sensor for detecting liquefied petroleum gas.
- Temperature and humidity sensors

b. Microcontroller:

- ESP32 microcontroller processes sensor data.

c. Output Devices:

- LCD displays air quality in PPM, temperature in degrees, and humidity percentage.
- Blynk application for real-time monitoring and graphical representation.

Functionality:

a. Gas Detection:

- Sensors measure gas concentrations in voltage levels.
- Voltage-to-PPM conversion using sensor libraries (MQ135, MQ3, MQ5).

b. Data Transmission:

- ESP32 collects and processes data from sensors.
- Data sent to both LCD and Blynk application.

c. LCD Display:

- Real-time display of air quality, temperature, and humidity.

d. Blynk Application:

- Mobile app displays graphical trends of humidity and temperature.
- Blynk cloud stores and visualizes historical data.

Operation:

a. Power On:

- Sensors initiate data collection upon system power-up.

b. Data Processing:

- ESP32 processes collected data.

c. LCD Output:

- Real-time display of air quality, temperature, and humidity on the LCD.

d. Blynk Output:

- Data transmitted to the Blynk app for users to monitor trends.

The system offers a robust solution for real-time air quality monitoring, providing users with actionable insights through both the LCD interface and the Blynk mobile application. The integration of various sensors with the ESP32 microcontroller ensures precise and reliable data collection, empowering users to make informed decisions based on dynamic environmental conditions.

3.2 DISADVANTAGES OF EXISTING SYSTEM

The existing system for air quality monitoring using sensors connected to an ESP32 microcontroller and transmitting data to an LCD and the Blynk application offers many advantages, but it also has some drawbacks:

Limited Sensor Range: The system's sensor range might not cover all potential pollutants or environmental factors that could affect air quality. It focuses on specific gases like CO2, smoke, alcohol, benzene, NH3, and LPG, but other pollutants could be present in the environment that aren't detected by the current sensor setup.

Calibration Requirements: Gas sensors like MQ series sensors often require frequent calibration to maintain accuracy. Failure to calibrate regularly could result in inaccurate readings, leading to misinformation about air quality.

Maintenance Intensive: Regular maintenance may be required to ensure the sensors are functioning correctly. This includes cleaning sensors, replacing components, and recalibrating the system. High maintenance requirements can be burdensome, especially in large-scale deployment scenarios.

Limited Data Analysis: While the Blynk application provides graphical representations of temperature and humidity trends, it may lack advanced analytics capabilities for more in-depth data analysis. Users may require additional tools or software for comprehensive data interpretation and decision-making.

Connectivity Issues: Dependency on the Blynk platform for data transmission introduces potential connectivity issues, such as network outages or platform maintenance downtime. These disruptions could impact real-time monitoring capabilities, affecting the system's reliability.

Lack of External Integration: The system may not easily integrate with other smart home or building automation systems, limiting its interoperability and potential for comprehensive environmental monitoring and control.

Addressing these disadvantages could involve exploring alternative sensor technologies with broader detection capabilities, enhancing data analysis functionalities, improving connectivity reliability, and considering cost-effective solutions for scalability. Additionally, seeking user feedback and conducting regular system evaluations can help identify and address emerging challenges to optimize the air quality monitoring system.

CHAPTER 4
PROPOSED SYSTEM

4.1 GENERAL BACKGROUND

In the face of escalating concerns about air pollution, driven by rapid industrialization and increasing vehicular emissions, this project introduces a groundbreaking system that leverages cutting-edge Internet of Things (IoT) technologies and advanced machine learning algorithms. The primary objective is to redefine the landscape of air quality monitoring and prediction, offering a transformative solution to the health risks posed to communities. The system's architecture is designed for real-time data collection, employing portable sensors strategically placed and connected through low-power wide area networks (LPWAN). This network ensures efficient communication, facilitating the seamless transmission of air quality data to a cloud-based IoT platform. Here, machine learning algorithms play a crucial role in the comprehensive analysis of the collected data, offering real-time pollution detection capabilities.

The heart of the project lies in its ability to categorize air quality as either polluted or non-polluted, providing an immediate and actionable classification that enables timely responses to fluctuations. By focusing on key pollutants like Particulate Matter (PM2.5), Carbon Monoxide (CO), and Oxides of Ozone, the system addresses the most critical components contributing to air pollution's adverse effects on human health and the environment. Rigorous experimentation validates the reliability of the system, demonstrating its precision in identifying evolving pollution patterns. This validation process underscores the project's effectiveness in providing accurate and timely information regarding air quality dynamics.

Beyond its technical capabilities, the system distinguishes itself through its user-friendly interface, making air quality information accessible to both the general public and relevant authorities. The interface offers intuitive visualizations and notifications, enhancing the overall user experience and encouraging community engagement. This democratization of information empowers individuals and governing bodies to proactively address the far-reaching health implications of air pollution. The real-time, location-specific insights provided by the system become a powerful tool for informed decision-making, enabling communities to take immediate action to safeguard public health.

In essence, this transformative project goes beyond conventional air quality monitoring approaches. It not only enhances our understanding of air quality dynamics but also equips communities and authorities with the means to collaboratively work towards mitigating the adverse effects of air pollution. By providing a practical and actionable solution, the system serves as a catalyst for creating more sustainable urban environments. As the project moves towards implementation, it stands as a testament to the potential of technology to bridge the gap between environmental challenges and effective solutions, paving the way for a healthier and more sustainable future. The empowerment of individuals and governing bodies through real-time insights is pivotal in fostering a collective effort towards cleaner air and improved public health.

4.2 SYSTEM ARCHITETURE

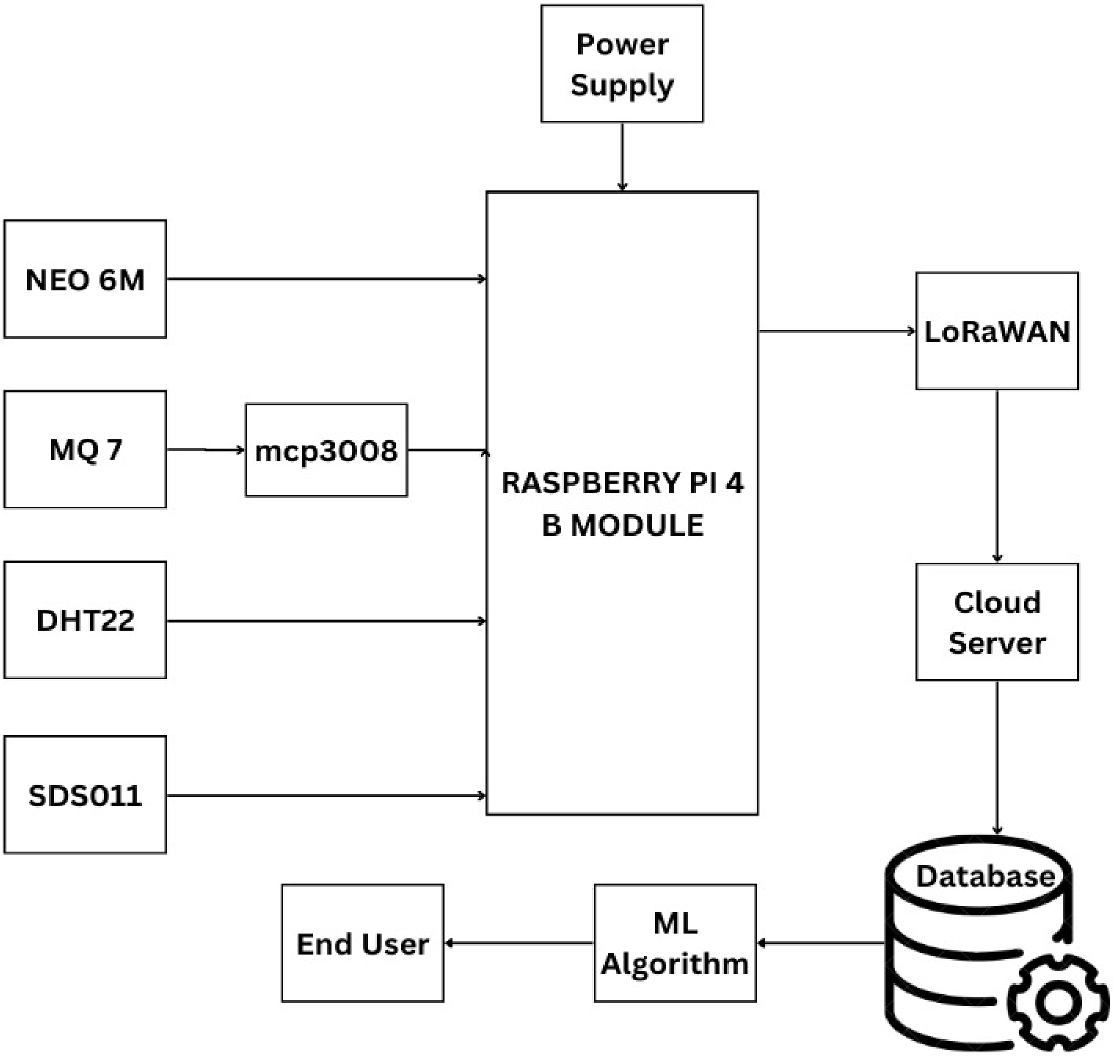


Fig 4.1 system architecture

4.3 AIR POLLUTION

Air pollution, a pervasive and pressing environmental issue, poses multifaceted threats to ecosystems, human health, and the global climate. With industrialization and urbanization accelerating worldwide, emissions from various sources, including vehicles, factories, and power plants, continue to saturate the air with harmful pollutants. Particulate matter, nitrogen oxides, sulfur dioxide, carbon monoxide, and volatile organic compounds are among the noxious substances discharged into the atmosphere, forming hazardous concoctions. These pollutants not only degrade air quality but also contribute significantly to respiratory diseases, cardiovascular ailments, and premature deaths, disproportionately affecting vulnerable populations like children, the elderly, and those with preexisting health conditions. Moreover, the adverse impacts of air pollution extend beyond human health, adversely affecting biodiversity, agricultural productivity, and climate stability.

From smog-choked cities to regions plagued by acid rain and ozone depletion, the consequences of unchecked air pollution are stark reminders of humanity's unsustainable relationship with the environment. Industrial activities, transportation, and energy production are major contributors to air pollution, emitting vast quantities of pollutants into the atmosphere. In urban areas, vehicular emissions are a primary source of air pollution, releasing nitrogen oxides, carbon monoxide, and particulate matter into the air. Factories and power plants, particularly those reliant on fossil fuels, also emit significant amounts of pollutants, including sulfur dioxide and mercury, further exacerbating air quality issues. Additionally, agricultural practices such as livestock farming and crop burning contribute to air pollution through the release of methane and particulate matter.

Addressing this global crisis demands concerted efforts to mitigate emissions through stringent regulations, technological innovations, and widespread adoption of renewable energy alternatives. Governments play a crucial role in implementing policies to curb emissions and incentivize sustainable practices. Measures such as emissions standards for vehicles and industrial facilities, cap-and-trade systems, and taxes on carbon emissions can effectively reduce pollution levels. Moreover, investment in renewable energy sources such as wind, solar, and hydropower is essential for transitioning away from fossil fuels and mitigating the impacts of climate change. Innovation in clean technologies, such as electric vehicles, energy-efficient appliances, and carbon capture and storage, is also crucial for reducing emissions across various sectors.

Embracing cleaner transportation methods, transitioning to greener industrial practices, and fostering international cooperation are vital steps towards safeguarding air quality and securing a healthier future for generations to come. Individuals can contribute to mitigating air pollution by adopting sustainable lifestyle choices, such as using public transportation, carpooling, biking, or walking instead of driving alone. Energy conservation measures, such as using energy-efficient appliances, reducing energy consumption, and switching to renewable energy sources at home, can also help reduce emissions. Furthermore, raising awareness about the impacts of air pollution and advocating for policy changes are essential for driving collective action and promoting environmental stewardship. By working together at local, national, and global levels, we can combat air pollution and protect the planet for current and future generations.

4.4 CAUSES OF AIR POLLUTION

Air pollution is primarily caused by a combination of human activities and natural phenomena, each contributing to varying degrees depending on factors such as location, population density, and industrialization levels. Some of the key causes of air pollution include:

Industrial Emissions: Factories, power plants, and other industrial facilities emit large quantities of pollutants into the air as byproducts of manufacturing processes and energy production. These emissions include sulfur dioxide (SO_2), nitrogen oxides (NO_x), carbon monoxide (CO), volatile organic compounds (VOCs), and particulate matter (PM), all of which contribute to air pollution.

Vehicle Emissions: The combustion of fossil fuels in vehicles, such as cars, trucks, and buses, is a major source of air pollution, particularly in urban areas with high traffic density. Exhaust gases from vehicles contain pollutants such as nitrogen oxides, carbon monoxide, sulfur dioxide, and particulate matter, which are released directly into the atmosphere.

Agricultural Activities: Agricultural practices, including livestock farming and crop cultivation, can generate significant amounts of air pollution. Livestock farming produces methane (CH_4) and ammonia (NH_3) from animal waste, while crop burning releases smoke and particulate matter into the air, contributing to regional air quality issues, especially during harvesting seasons.

Burning of Fossil Fuels: The burning of fossil fuels, such as coal, oil, and natural gas, for heating, electricity generation, and industrial processes, is a major source of air pollution worldwide. This combustion releases pollutants such as sulfur dioxide, nitrogen oxides, carbon dioxide (CO_2), and particulate matter, contributing to both local and global air quality problems and climate change.

Deforestation and Biomass Burning: Deforestation and the burning of forests and other biomass for land clearing, agriculture, and cooking purposes release large amounts of smoke, particulate matter, and carbon dioxide into the atmosphere. These activities not only degrade air quality but also contribute to the loss of biodiversity and ecosystem degradation.

Construction and Demolition: Construction activities, including excavation, demolition, and road building, can generate dust and other particulate matter, contributing to localized air pollution. Additionally, the use of construction equipment and machinery powered by diesel engines can emit pollutants such as nitrogen oxides and particulate matter.

Household Sources: Residential sources of air pollution include the burning of solid fuels such as wood, coal, and biomass for cooking, heating, and lighting purposes in households without access to clean energy alternatives. Indoor air pollution from these sources can pose significant health risks, particularly in developing countries where traditional cooking methods are prevalent.

Natural Sources: While human activities are the primary drivers of air pollution, natural phenomena such as wildfires, volcanic eruptions, dust storms, and biogenic emissions from vegetation also contribute to air quality issues. While these natural sources can cause localized spikes in pollution levels, they typically have a smaller overall impact compared to anthropogenic sources.

Overall, addressing air pollution requires a multifaceted approach that involves reducing emissions from key sources, transitioning to cleaner energy sources, improving energy efficiency, implementing stricter regulations and policies, and promoting sustainable practices at local, national, and global levels.

List of Pollutants are:

- Particulate Matter
- Ozone
- Carbon monoxide
- Sulphur Dioxide
- Nitrous Oxide

4.5 ADVANTAGES OF PROPOSED SYSTEM

Real-time Monitoring: Utilizing IoT technologies and machine learning algorithms, the system enables real-time monitoring of air quality, providing up-to-date information on pollution levels. This instantaneous data allows for timely responses to fluctuations in air quality.

Comprehensive Analysis: The system's machine learning algorithms conduct comprehensive analysis of collected data, enabling the detection of key pollutants such as PM2.5, CO, and Oxides of Ozone. This thorough analysis enhances understanding of air quality dynamics and aids in identifying evolving pollution patterns.

Precision and Reliability: Rigorous experimentation validates the system's reliability, demonstrating its precision in identifying pollution patterns. This ensures that the data provided is accurate and trustworthy, enhancing confidence in decision-making based on the system's insights.

User-friendly Interface: The system features a user-friendly interface with intuitive visualizations and notifications, making air quality information accessible to both the general public and relevant authorities. This enhances community engagement and empowers individuals and governing bodies to take proactive measures to address air pollution.

4.6 SYSTEM DESIGN

4.6.1 DFD-LEVEL 0

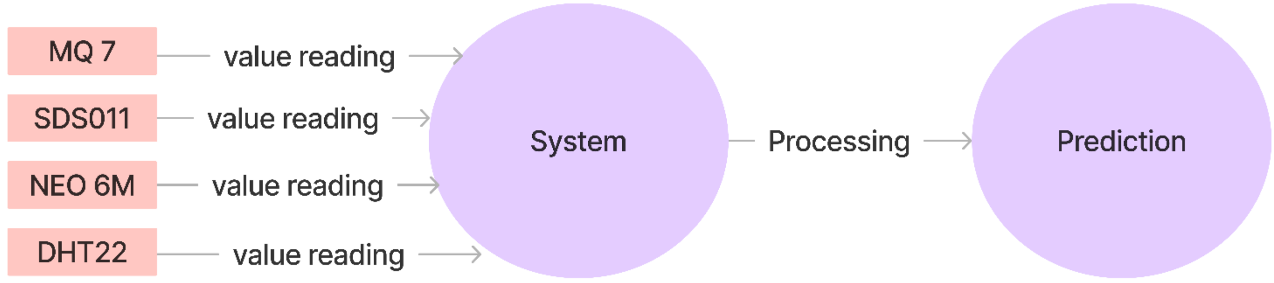


Fig 4.2 DFD LEVEL 0

4.6.2 DFD LEVEL 1

4.6.2.1 MQ-7

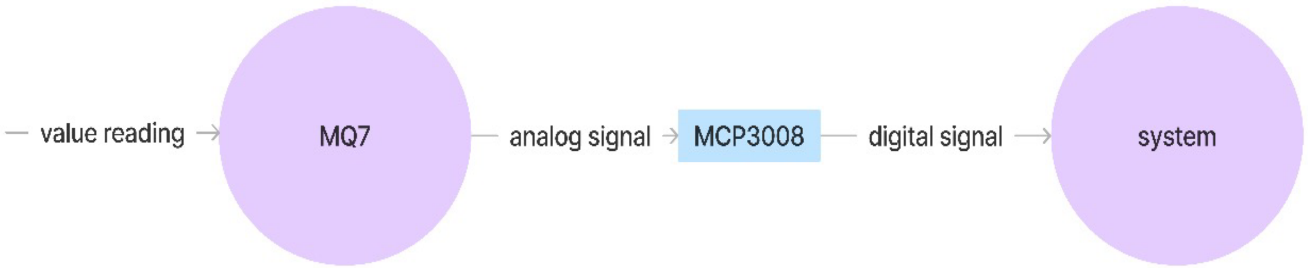


Fig 4.3 MQ-7

4.6.2.2 SDS011

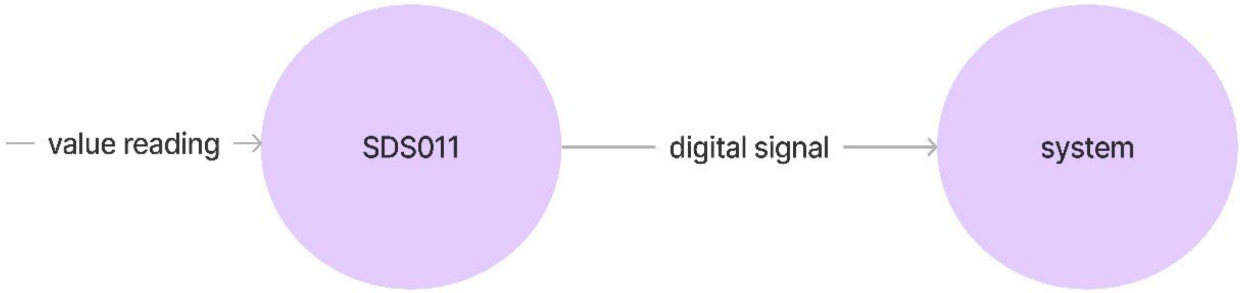


Fig 4.4 SDS011

4.6.3.3 DHT-22

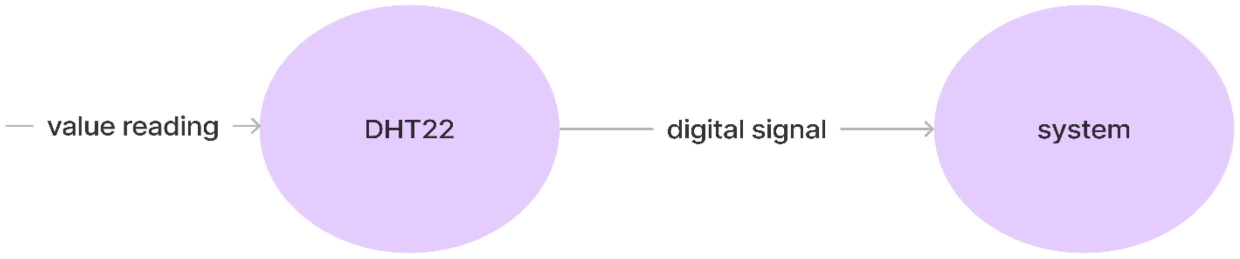


Fig 4.5 DHT-22

4.6.4.4 NEO 6M

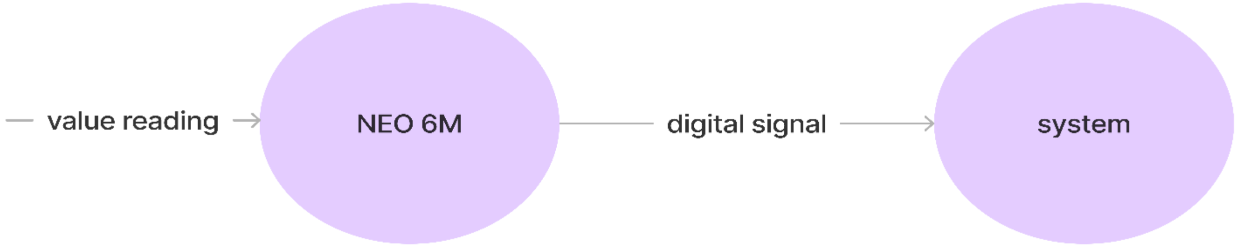


Fig 4.6 NE0 6M

CHAPTER 5

SYSTEM IMPLEMENTATION

5.1 HARDWARE REQUIREMENTS

5.1.1 PARTICULATE MATTER (PM2.5)

The SDS011 is a popular particulate matter (PM) detection sensor that specifically measures PM2.5, which refers to fine inhalable particles with diameters generally 2.5 micrometers and smaller. These particles can pose significant health risks as they can penetrate deep into the respiratory system. The SDS011 sensor is widely used in air quality monitoring applications to assess the concentration of PM2.5 in the atmosphere.

The SDS011 operates on the principle of laser scattering. It utilizes a laser diode to emit light into the air sample, and a photodetector measures the scattered light. By analyzing the scattered light patterns, the sensor can determine the concentration of PM2.5 particles in the air. The sensor is capable of providing real-time data, allowing for continuous monitoring of air quality.

One notable feature of the SDS011 is its compact size and ease of integration into various environmental monitoring systems. It is often employed in portable air quality monitoring devices, stationary monitoring stations, and even DIY projects for individuals interested in assessing the air quality in their surroundings.

The sensor communicates with external devices such as microcontrollers or single-board computers through a UART (Universal Asynchronous Receiver-Transmitter) interface. This enables seamless integration into a wide range of electronic systems. Users can obtain PM2.5 concentration data through the sensor's output, typically in micrograms per cubic meter (µg/m³). Regular calibration is essential to ensure the accuracy of the SDS011 sensor. Calibration helps account for variations in sensor performance over time and environmental conditions.

Fig 5.1 Particulate Matter SDS011

5.1.2 CARBON MONOXIDE

The MQ-7 is a gas sensor that is widely used for detecting carbon monoxide (CO) in various applications. Developed by Hanwei Electronics, the MQ-7 sensor is known for its sensitivity to carbon monoxide gas and its ability to provide reliable measurements. This sensor is commonly employed in gas detectors, industrial safety systems, and other devices where monitoring and early detection of CO are crucial.

The MQ-7 operates on the principle of a metal oxide semiconductor. Inside the sensor, there is a sensing element composed of a tin dioxide (SnO2) thin film. When exposed to the target gas, in this case, carbon monoxide, the electrical conductivity of the SnO2 film changes. The sensor has a built-in heater that keeps the sensing element at an elevated temperature, promoting a more sensitive response to gas concentrations.

One of the key features of the MQ-7 is its sensitivity to low concentrations of carbon monoxide, making it suitable for early warning systems. It has a detection range of 20 to 2000 parts per million (ppm) of CO. The sensor provides an analog output voltage that corresponds to the concentration of carbon monoxide in the environment. This output can be interfaced with microcontrollers or other processing units for further analysis and control.

The MQ-7 sensor is compact, cost-effective, and easy to use. However, it is essential to note that the sensor's response time and accuracy can be influenced by factors such as temperature and humidity. Regular calibration and environmental monitoring are necessary to ensure reliable performance.

In practical applications, the MQ-7 sensor is often integrated into gas detection systems in homes, industries, and automotive environments. Its ability to operate in a range of temperatures and its sensitivity to low concentrations of CO make it a valuable component in ensuring safety and preventing the harmful effects of carbon monoxide exposure.

Fig 5.2 Carbon Monoxide MQ-7

5.1.3 TEMPERATURE AND HUMIDITY SENSOR(DHT-22)

The DHT22 sensor, also known as the RHT22 or AM2302, is a popular digital sensor used for measuring temperature and humidity levels in various environments. It is widely employed in applications such as weather stations, home automation systems, and industrial monitoring setups. The sensor utilizes a digital signal output, making it easy to interface with microcontrollers and other digital devices.

The DHT22 sensor is designed to provide accurate and reliable measurements of both temperature and humidity. It has a temperature range of -40°C to 80°C (-40°F to 176°F) with an accuracy of ±0.5°C, and a humidity range of 0% to 100% with an accuracy of ±2-5%, depending on the operating conditions. This makes it suitable for a wide range of indoor and outdoor environments.

One of the key features of the DHT22 sensor is its simplicity of use. It requires only a single digital pin for communication with a microcontroller, and the data output is in a digital format, eliminating the need for analog-to-digital conversion. Additionally, the sensor has a relatively fast response time, typically providing updated readings every 2 seconds.

The DHT22 sensor consists of a capacitive humidity sensor and a thermistor for temperature measurement, housed within a protective casing. It operates on a wide voltage range of 3.3V to 6V, making it compatible with a variety of microcontroller platforms.

Overall, the DHT22 sensor offers a cost-effective solution for monitoring temperature and humidity levels in various applications, providing accurate and reliable data for environmental monitoring and control purposes.

Fig 5.3 Temperature Sensor DHT-22

5.1.4 RASPBERRY PI

The Raspberry Pi is a credit card-sized single-board computer that has revolutionized the world of computing, enabling enthusiasts, hobbyists, and professionals to explore, experiment, and innovate in the realm of embedded systems and DIY electronics. Conceived by the Raspberry Pi Foundation, this affordable and versatile device has found applications in diverse fields, from education and prototyping to home automation and industrial projects.

At its core, the Raspberry Pi is a powerful microcomputer equipped with a central processing unit (CPU), memory, input/output ports, and various connectivity options. The compact design, reminiscent of a credit card, makes it highly portable and well-suited for a wide range of projects. The foundation's mission to provide an affordable computing platform to encourage learning and experimentation has resulted in several iterations of the Raspberry Pi, each more capable than its predecessor.

The heart of the Raspberry Pi is its CPU, which handles the computational tasks. The early models featured ARM-based processors, while more recent versions have seen improvements in processing power, incorporating multicore processors for enhanced performance. The onboard RAM complements the CPU, allowing the device to efficiently handle various applications, from simple programming exercises to multimedia processing.

Connectivity is a key feature of the Raspberry Pi. It includes USB ports for connecting peripherals such as keyboards, mice, and external storage devices. HDMI ports facilitate the connection to displays, enabling users to interact with the device in a manner similar to traditional desktop computers. Additionally, the Raspberry Pi is equipped with GPIO (General Purpose Input/Output) pins, allowing users to interface with external hardware and sensors, making it a popular choice for projects in the fields of robotics, automation, and Internet of Things (IoT). One of the defining aspects of the Raspberry Pi ecosystem is the accessibility of its operating system. Raspbian, a Debian-based Linux distribution optimized for the Raspberry Pi, is the default OS. Its user-friendly interface, coupled with a vast community of developers, facilitates a smooth learning curve for beginners. However, users have the flexibility to explore other operating systems, including Ubuntu and various specialized distributions tailored for specific applications. The Raspberry Pi's impact extends beyond personal projects and into the realm of education. Its affordability and versatility have made it an invaluable tool for teaching computer science and programming concepts in schools and community programs worldwide. The device's hands-on approach to learning, coupled with a wealth of educational resources and projects, empowers students to develop practical skills and a deep understanding of computing principles.

Furthermore, the Raspberry Pi has become a go-to platform for prototyping and proof-of- concept development in the tech industry. Its low cost and compact form factor make it an ideal choice for testing ideas before moving to more complex and expensive systems. This has led to a vibrant ecosystem of accessories, add-ons, and third-party software, further expanding the capabilities of the Raspberry Pi.

In conclusion, the Raspberry Pi has emerged as a revolutionary force in the world of computing. Its compact size, affordability, and versatility make it a powerful tool for individuals, educators, and professionals alike. As the Raspberry Pi community continues to grow, so does the potential for innovation and exploration in the ever-expanding landscape of embedded systems and DIY electronics. Whether used for educational purposes, prototyping, or as the brains behind creative projects, the Raspberry Pi stands as a testament to the democratization of computing and the boundless possibilities it unlocks.

Fig 5.4 Raspberry pi

Advantages

- Af**fordability:** One of the key advantages of Raspberry Pi is its low cost, making it accessible to a wide range of users, including students, hobbyists, and educators.
- **Versatility:** Raspberry Pi is a versatile single-board computer that can be used for a variety of applications, from basic computing tasks to more complex projects such as home automation, robotics, and media centers.
- **Community Support:** The Raspberry Pi has a large and active community of users and developers. This community support provides a wealth of resources, tutorials, and forums where users can seek help, share knowledge, and collaborate on projects.
- **Compact Size:** The small form factor of Raspberry Pi makes it suitable for projects with space constraints. Its compact design allows for creative and portable applications.
- **Energy Efficiency:** Raspberry Pi is energy-efficient, consuming very little power compared to traditional desktop computers. This makes it suitable for projects requiring

low power consumption or running on battery power.

Disadvantages

- **Limited Processing Power:** While Raspberry Pi is versatile, it has limited processing power compared to more powerful desktop computers. This can be a limitation for resource-intensive applications.
- **Limited RAM:** The amount of RAM on Raspberry Pi boards is relatively modest. This limitation can affect the performance of applications that require significant memory.
- **Limited Graphics Performance:** Raspberry Pi may not be suitable for graphics-intensive applications or gaming due to its limited graphics processing capabilities.
- **Storage Limitations:** Raspberry Pi typically relies on microSD cards for storage. While this is cost-effective, it may result in slower read and write speeds compared to traditional hard drives or SSDs.
- **Not Suitable for All Tasks:** Raspberry Pi is not a replacement for high-end computers in terms of performance. It's designed for specific use cases and may not be suitable for tasks requiring significant computational power.

5.1.5 LoRAWAN

LoRaWAN, or Long Range Wide Area Network, is a wireless communication protocol designed for long-range communication between low-power devices. It is a key player in the Internet of Things (IoT) landscape, providing a scalable and efficient solution for connecting a myriad of devices over long distances. Developed to address the specific challenges of IoT applications, LoRaWAN offers a balance between long-range communication, low power consumption, and cost-effectiveness.

At its core, LoRaWAN utilizes the LoRa (Long Range) modulation technique, which enables devices to transmit small packets of data over extended distances while consuming minimal power. This makes it well-suited for IoT applications where devices may be deployed in remote or challenging environments, such as agriculture, smart cities, industrial monitoring, and more.

One of the defining features of LoRaWAN is its star-of-stars network architecture. In this setup, end-devices (sensors or actuators) communicate with gateways that act as intermediaries between the devices and the network server. The gateways, in turn, relay the data from the devices to the central network server, which manages the entire communication process. This architecture allows for efficient use of bandwidth and provides flexibility in scaling the network to accommodate a large number of devices.

LoRaWAN operates in various frequency bands, including sub-GHz bands (868 MHz in Europe, 915 MHz in North America, and others in different regions). The choice of frequency bands contributes to the network's long-range capabilities, as lower frequencies generally offer better propagation characteristics, allowing signals to penetrate obstacles and cover larger distances.

In terms of communication classes, LoRaWAN defines three classes that cater to different use cases. Class A devices are the most power-efficient, as they have specific receive windows for communication, allowing them to consume minimal power. Class B devices have scheduled receive windows, providing a balance between power consumption and responsiveness. Class C devices, on the other hand, have continuous receive windows, making them suitable for applications that require nearly real-time communication.

Security is a critical aspect of any communication protocol, and LoRaWAN is no exception. It incorporates various security features, including device authentication, end-to-end encryption, and message integrity checks, to ensure the confidentiality and integrity of data transmitted over the network. These security measures are essential to protect sensitive information and prevent unauthorized access to IoT devices.

The open and collaborative nature of the LoRa Alliance, a consortium of companies supporting the development of the LoRaWAN standard, has played a pivotal role in its widespread adoption. The LoRa Alliance works to ensure interoperability between different vendors' devices and promotes the global deployment of LoRaWAN networks.

5.1.6 CLOUD SERVER

A cloud server is a virtualized computing resource provided by cloud service providers, transforming the traditional paradigm of physical servers into a dynamic and flexible digital environment. These servers leverage virtualization technology, allowing multiple instances to run on a single physical server. Cloud servers offer a wide array of advantages, notably scalability, enabling users to effortlessly adjust computing resources in response to changing demands. This scalability is particularly beneficial for businesses experiencing variable workloads, as they can easily scale up during peak periods and scale down during quieter times. Additionally, cloud servers provide unparalleled flexibility, allowing users to choose server configurations that align with their specific requirements. Whether it's a general-purpose server for everyday computing tasks or a specialized server optimized for memory-intensive applications, cloud providers offer a diverse range of options to cater to different workloads.
Another key feature of cloud servers is their global accessibility via the internet. This accessibility

promotes remote collaboration and facilitates access to resources from virtually anywhere in the world. This level of flexibility is crucial in today's interconnected and digital business landscape, enabling organizations to embrace remote work and streamline operations. Moreover, cloud servers operate on a cost-efficient pay-as-you-go model, eliminating the need for substantial upfront investments in hardware. This economic model is particularly advantageous for startups and small businesses, allowing them to access cutting-edge computing resources without the burden of high initial capital expenditures. Combined with managed services, security measures, and geographic redundancy, cloud servers represent a transformative approach to computing infrastructure, empowering businesses with reliable, scalable, and secure solutions that adapt to their evolving needs.

5.1.7 MCP3008

The MCP3008, a 10-bit analog-to-digital converter (ADC), plays a pivotal role in the IoT door project by facilitating the conversion of analog signals from sensors, such as door sensors or environmental sensors, into digital data. Integrated seamlessly with the microcontroller, such as the Raspberry Pi, via the Serial Peripheral Interface (SPI), the MCP3008 enables the acquisition of precise analog measurements essential for monitoring the state of the door or environmental conditions. Leveraging software libraries compatible with the chosen microcontroller platform, developers can effortlessly interface with the MCP3008, retrieving sensor data with accuracy and efficiency. The acquired analog sensor readings undergo processing, including unit conversion and threshold detection, to derive actionable insights for the IoT door system. By seamlessly integrating the MCP3008, the project gains enhanced capabilities in sensor data acquisition and processing, fostering comprehensive monitoring and control functionalities tailored to the project's requirements.

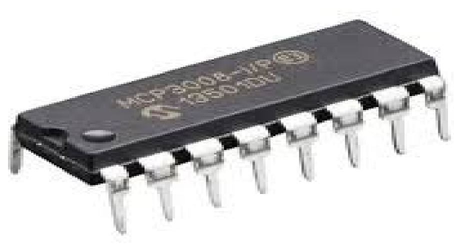


Fig 5.5 MCP3008

5.2 SOFTWARE REQUIREMENTS

5.2.1 PUTTY

PuTTY, a versatile terminal emulator software, serves as an essential tool for establishing remote connections to Raspberry Pi devices. With its support for SSH, Telnet, and serial connections, PuTTY enables users to interact with Raspberry Pi's command line interface from a remote location. This functionality is particularly valuable in headless setups where direct physical access to the Raspberry Pi is unavailable. PuTTY's SSH support ensures secure communication between the client (personal computer) and the Raspberry Pi server, safeguarding data integrity and confidentiality. Additionally, PuTTY offers features such as session management, flexible configuration options, port forwarding, and X11 forwarding, enhancing its usability and adaptability to diverse usage scenarios. Installation of PuTTY is straightforward, requiring no complex setup process, and once configured with the Raspberry Pi's IP address and SSH credentials, users can seamlessly initiate connections and perform various tasks remotely, including software development, system administration, and IoT projects.

5.2.2 Raspberry Pi Imager

Raspberry Pi Imager is a user-friendly software tool designed to simplify the process of installing operating systems onto Raspberry Pi microSD cards. Developed by the Raspberry Pi Foundation, this tool offers a straightforward interface that allows users to select from a variety of supported operating systems, including Raspbian, Ubuntu, and others, and write them directly onto an SD card with ease. The Imager eliminates the need for manual downloading and imaging of disk images, streamlining the setup process for both beginners and experienced users. Its intuitive design guides users through the installation steps, making it accessible even to those with limited technical expertise. Additionally, Raspberry Pi Imager is compatible with Windows, macOS, and Linux operating systems, ensuring accessibility across different platforms. Overall, Raspberry Pi Imager is an invaluable tool for Raspberry Pi enthusiasts, hobbyists, and educators, simplifying the process of setting up and experimenting with Raspberry Pi projects.

5.2.3 Geany IDE

Geany is a lightweight and versatile Integrated Development Environment (IDE) primarily designed for programmers who prefer simplicity and efficiency. It offers a minimalistic interface while still providing essential features for coding in various programming languages, including but not limited to C, C++, Python, Java, and HTML/CSS. Geany's key features include syntax highlighting, code folding, auto-completion, and a built-in terminal for executing commands directly from the IDE. It supports project management, allowing users to organize their files and folders effectively. Geany also offers customization options, allowing users to tailor the IDE's appearance and functionality to their preferences. Its lightweight nature makes it suitable for both beginners and experienced developers who value speed and simplicity in their development workflow.

5.2.4 FastAPI

FastAPI can be utilized to create a local server for the IoT door project, providing a robust and efficient framework for handling HTTP requests and responses. Leveraging the simplicity and performance of FastAPI, developers can quickly define endpoints to interact with the IoT door system. By writing Python code with FastAPI decorators, such as **@app.get** or **@app.post**, developers can define routes and corresponding functions to handle various actions, such as retrieving sensor data, controlling the actuator, or receiving commands from remote clients. FastAPI's automatic data validation and serialization, powered by Pydantic, ensure that incoming requests are validated against predefined models, reducing the risk of errors and enhancing security. Additionally, FastAPI's support for asynchronous programming enables efficient handling of concurrent requests, making it suitable for real-time applications like the IoT door system. With FastAPI, developers can create a local server tailored to the specific requirements of the IoT door project, providing seamless communication.

CHAPTER 6
RESULT AND DISCUSSION

6.1 SAMPLE DATA SHEET

DATE	TIME	CO	PM	TEM	HUM
08-05-2024	11:02:44	0.354	28.9	33.5	67.6
08-05-2024	11:02:44	0.754	28.9	33.5	67.6
08-05-2024	11:02:44	0.754	28.9	33.5	67.6
08-05-2024	11:02:44	0.753	27.9	33.5	67.6
08-05-2024	11:02:44	0.756	27.9	33.5	67.6
08-05-2024	11:02:44	0.751	27.9	33.5	67.6
08-05-2024	11:02:44	0.753	27.9	33.5	67.6
08-05-2024	11:02:14	0.753	28.1	33.5	67.6
08-05-2024	11:03:14	0.753	28.1	33.6	67.8
08-05-2024	11:03:14	0.0753	28.1	33.6	67.8
08-05-2024	11:03:14	0.0753	26.9	33.6	67.8
08-05-2024	11:03:14	0.0753	26.9	33.5	67.6
08-05-2024	11:03:14	0.0753	26.9	33.5	67.6
08-05-2024	11:03:14	0.0753	22.5	33.5	67.6
08-05-2024	12:02:19	0.0747	22.5	33.5	67.6
08-05-2024	12:02:19	0.0747	22.5	33.5	67.6
08-05-2024	12:02:19	0.0747	22.5	33.6	67.6
08-05-2024	12:02:19	0.0747	23.5	33.6	67.6
08-05-2024	12:02:19	0.0747	23.6	33.6	67.6
08-05-2024	12:02:19	0.0747	23.9	33.6	67.6
08-05-2024	12:02:19	0.0747	23.7	33.6	67.6
08-05-2024	12:02:19	0.0747	23.4	33.6	67.6
08-05-2024	12:02:19	0.0747	23.4	33.6	67.6

Fig 6.1 Sample Data Sheet

6.2 GRAPHICAL REPRESENTATION

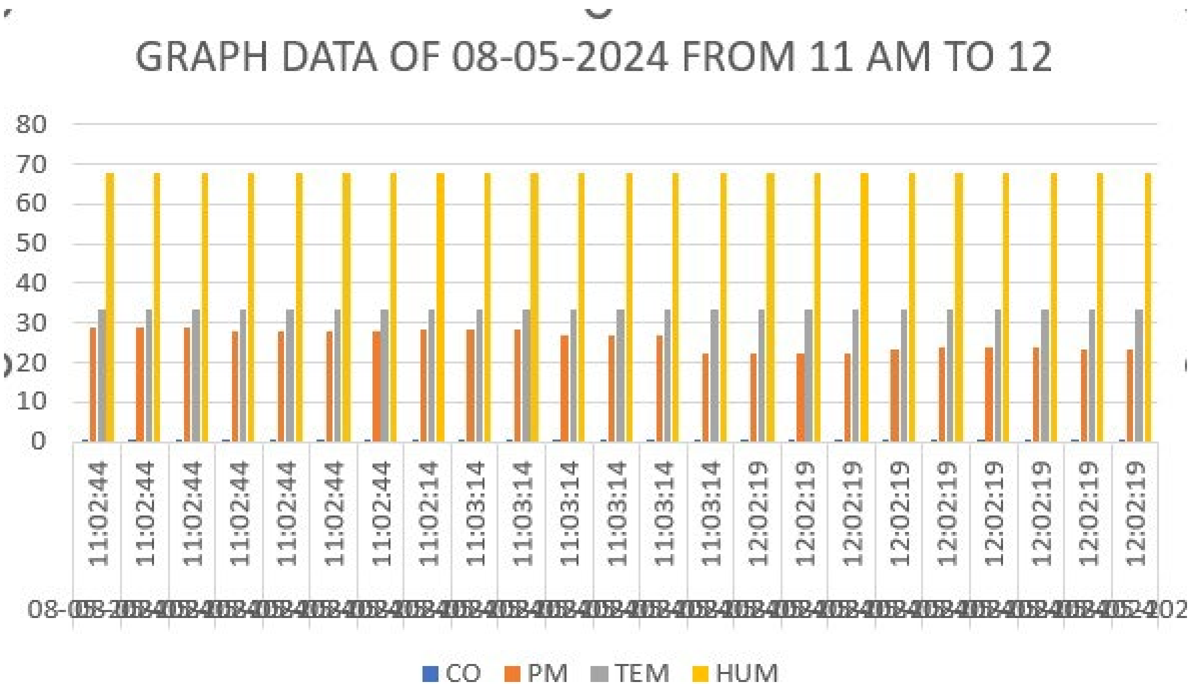


Fig 6.2 Analysis of Data

Hence, from this result the average of carbon monoxide is 0.50481

Average of particulate matter is 21.129629

Average of humidity is 36.0333

Average of temperature is 67.6

CHAPTER 7
FUTURE SCOPE

The proposed air quality monitoring and prediction project is not only a response to current environmental concerns but also represents a beacon of innovation with significant potential for future development and impact. At its core, the project harnesses the power of cutting-edge technologies, notably the Internet of Things (IoT) and advanced machine learning algorithms, to revolutionize how we understand and address air pollution. By leveraging these tools, the project aims to redefine the landscape of air quality monitoring and prediction, offering actionable insights that can mitigate the adverse effects of pollution on public health and the environment. However, the true potential of this endeavor lies not only in its immediate applications but also in its capacity for evolution and expansion in the years to come.

Looking ahead, the project stands to benefit from ongoing advancements in technology. As IoT devices become more sophisticated and cost-effective, there is an opportunity to further enhance the system's capabilities. For instance, the integration of emerging innovations such as edge computing and 5G networks could enable faster and more efficient data collection and analysis. This could lead to more accurate predictions and timely interventions, ultimately improving the system's effectiveness in safeguarding public health.

Moreover, the expansion of sensor networks presents a promising avenue for future development. While the project initially focuses on deploying sensors in urban areas, there is potential to extend coverage to previously underserved regions, including rural areas and developing countries. By doing so, the project can provide a more comprehensive understanding of air pollution dynamics on a global scale, helping to address environmental justice issues and ensure that all communities have access to clean air.

Integration with smart cities initiatives represents another avenue for future growth. As cities around the world increasingly embrace smart technologies to improve urban infrastructure and sustainability, there is an opportunity to integrate air quality data into existing city planning and management systems. This would enable data-driven decision-making to enhance environmental sustainability and livability, ultimately leading to healthier and more resilient communities.

Furthermore, the project's data and insights hold promise for healthcare applications. Poor air quality is a significant risk factor for respiratory illnesses, particularly for vulnerable populations such as children, the elderly, and individuals with pre-existing health conditions. By providing real-time information on air pollution levels, the project can help healthcare providers and policymakers develop targeted interventions and early warning systems to protect public health.

Climate change mitigation represents another area where the project can make a significant impact. Many air pollutants are also greenhouse gases, contributing to global warming and climate change. By monitoring and analyzing emission sources and trends, the project can provide valuable data to inform policies and initiatives aimed at reducing greenhouse gas emissions. In doing so, the project can contribute to broader efforts to address climate change and create a more sustainable future for generations to come.

Moreover, the project serves as a platform for public engagement and education. By making air quality information accessible to the general public through user-friendly interfaces and intuitive visualizations, the project empowers individuals to take meaningful actions to reduce air pollution and promote environmental stewardship. This can foster a sense of community ownership and collective responsibility for environmental protection, ultimately leading to more sustainable behaviors and lifestyles.

Finally, collaboration with international organizations and research institutions is essential for maximizing the project's impact. By sharing knowledge, best practices, and resources, stakeholders can work together to address common challenges and achieve shared goals. This collaborative approach not only enhances the effectiveness of the project but also fosters a sense of solidarity and cooperation in the global fight against air pollution and climate change. The future scope of the air quality monitoring and prediction project is vast and multifaceted, encompassing technological innovation, expanded deployment, integration with smart city initiatives, healthcare applications, climate change mitigation, public engagement, and international collaboration. By continuing to evolve and adapt to emerging challenges and opportunities, the project has the potential to make a significant contribution to environmental sustainability and public health in the years to come.

CHAPTER 8
CONCLUSION

The integration of machine learning and IoT technologies in air pollution monitoring and prediction systems marks a significant leap forward in our collective efforts to address the pressing environmental challenges posed by air pollution. This innovative approach not only enhances the accuracy and efficiency of real-time monitoring but also enables proactive measures through predictive analytics. By harnessing the power of machine learning algorithms, these systems can analyze vast amounts of data from diverse sources, providing a comprehensive understanding of air quality dynamics.

The synergy between machine learning and IoT devices empowers cities and communities to make informed decisions for mitigating air pollution. With the continuous monitoring of air quality parameters such as particulate matter, ozone levels, and nitrogen dioxide, these systems offer a nuanced understanding of pollution patterns. This real-time data enables authorities to implement targeted interventions, such as traffic management strategies or emission control measures, to address specific sources of pollution. The predictive capabilities of these systems further amplify their impact by allowing stakeholders to anticipate pollution spikes and take preemptive actions, ultimately contributing to the overall improvement of air quality.

Furthermore, the scalability and adaptability of machine learning and IoT-based air pollution monitoring systems make them versatile tools for a wide range of environments. Whether deployed in urban centers, industrial complexes, or even in remote areas, these systems can be customized to suit the unique challenges and sources of pollution in each location. The flexibility of IoT devices ensures seamless integration with existing infrastructure, while machine learning algorithms continuously evolve and improve their predictive accuracy over time. This adaptability positions these systems as indispensable assets in the global fight against air pollution, offering scalable solutions that can be tailored to diverse geographical and industrial contexts.

Despite the undeniable advantages of machine learning and IoT in air pollution monitoring, it is crucial to acknowledge certain challenges and considerations. Data privacy, for instance, emerges as a paramount concern in the collection and analysis of sensitive environmental data. Striking a balance between the need for comprehensive data and individual privacy rights requires careful implementation of robust security measures and transparent data governance frameworks. Additionally, the accessibility of these advanced technologies poses challenges for communities with limited resources. Bridging the digital divide and ensuring equitable access to air quality monitoring tools must be prioritized to address environmental justice concerns and foster a collaborative, inclusive approach to combating air pollution.

The convergence of machine learning and IoT technologies in air pollution monitoring and prediction systems signifies a groundbreaking approach to environmental stewardship. These systems not only provide real-time insights into air quality but also empower communities and authorities to take proactive measures in the fight against air pollution. The adaptability and scalability of these technologies make them indispensable tools with the potential to revolutionize environmental management strategies worldwide. As we navigate the complexities of urbanization and industrialization, embracing innovative solutions such as machine learning and IoT-driven air pollution monitoring becomes imperative for building sustainable, healthier communities and safeguarding the well-being of current and future generations.

REFERENCES

[1] Savita Vivek Mohurle, Dr. Richa Purohit and Manisha Patil, A study of fuzzy clustering concept for measuring air pollution index,2018, International Journal of Advanced Science and Research.

[2] Gaganjot Kaur Kang, Jerry Zeyu Gao, Sen Chiao, Shengqiang Lu, and Gang Xie, Air Quality Prediction: Big Data and Machine Learning Approaches,2018, International Journal of Environmental Science and Development.

[3] Z. Didi and I. El Azami, "IoT, Comparative Study Between the Use of Arduino Uno, Esp32, and Raspberry Pi in Greenhouses," in Digital Technologies and Applications, 2022, vol. 455, ISBN: 978-3-031-02446-7.

[4] A. A. Aziz et al., "Portable Outdoor Air Quality Monitoring Using A Wireless Sensor Network (WSN)," 2021 4th International Conference on Computing & Information Sciences (ICCIS), Karachi, Pakistan, 2021, pp. 1-5, doi: 10.1109/ICCIS54243.2021.9676381.

[5] The article "Differences of Performance Analysis of Single Channel LoRaWAN Network for Air Pollution Monitoring System Using IoT Platform in Smart City – A Review" by N. F. E. B. N. Kamaruzaman, S. S. Sarnin and N. F. Naim discusses the performance of a single-channel LoRaWAN network for air pollution monitoring in a smart city.

[6] S. Ooko, "A Comparison of Arduino, Raspberry Pi, and ESP8266 Boards," IEEE Transactions on Emerging Topics in Computing 2022.

[7] J. Jo, B. Jo, J. Kim, S. Kim, W. Han, "Development of an IoT-Based Indoor Air Quality Monitoring Platform," in Journal of Sensors, vol. 2020, Article ID 8749764, 14 pages, 2020.

[8] T. W. Ayele and R. Mehta, "Air pollution monitoring and prediction using IoT," in Proc. Int. Conf. Inven. Commun. Comput., vol. 15, no. 12, Year, pp. 31392.

[9] Air Quality Prediction and Monitoring Using Machine Learning-Based Forecasting Approach Kumar, D., Chaudhri, S.N., Rajput, N.S. 2023 2023 International Conference on IoT, Communication and Automation Technology, ICICAT 2023.

[10] J. J. Caubel, T.E. Cados, T.W. Krichstetter, (2018) 'A New Black Crabon Sensor for Dense Air Quality Monitoring Networks', IEEE.

APPENDIX-1

Source Code

Data Reading

```
from fastapi import FastAPI
from typing import List
import serial
import board
import 38dafruit_dht
from gpiozero import MCP3008
import csv
import pynmea2
import time
import os
import string
from datetime import datetime
app = FastAPI()
serial_port = '/dev/ttyUSB0'
baud_rate = 9600
ser = serial.Serial(serial_port, baud_rate, timeout=2)
dhtDevice = 38dafruit_dht.DHT22(board.D16)
adc = MCP3008(channel=0)
def read_pm_data() -> List[float]: //start of pm sensor
    """
    Function to read PM data from the Nova PM sensor.
    Returns a list of PM2.5 data points (up to 40).
    """
    pm_data = []
    try:
      ser.write(b'\Xaa\Xb4\x04\x13\x00\x01\Xe1')
     for _ in range(30):
          data = ser.read(10)
          pm25 = int.from_bytes(data[2:4], byteorder='little') / 10.0
          pm_data.append(pm25)
```

```
    except Exception as e:
        print("Error reading PM data:", e)
    write_to_csv(pm_data, "pm_data.csv")
    return pm_data[:30]
def read_dht_data() -> List[dict]: //start of temperature sensor
dht_data = []
    try:
        for _ in range(30):
            temperature_c = dhtDevice.temperature
            humidity = dhtDevice.humidity
            dht_data.append({"temperature_c": temperature_c, "humidity": humidity})
    except RuntimeError as error:
      print(error.args[0])
    write_to_csv(dht_data, "dht_data.csv")
    return dht_data[:30]
def read_mq7_data() -> List[float]: //start of carbon monoxide sensor
    mq7_data = []
    try:
        for _ in range(30):
            analog_value = adc.value
            mq7_data.append(analog_value)
    except Exception as e:
        print("Error reading MQ7 data:", e)
    write_to_csv(mq7_data, "mq7_data.csv")
    return mq7_data[:30]
def write_to_csv(data: List, filename: str):
    mode = 'a' if os.path.exists(filename) else 'w'
    with open(filename, mode=mode, newline='') as file:
        writer = csv.writer(file)
        for item in data:
            timestamp = datetime.now().strftime("%Y-%m-%d %H:%M:%S")
            if isinstance(item, dict):
```

Data passing to ML

```
import time
```

```
import joblib
import requests
from fastapi import FastAPI
from fastapi.middleware.cors import CORSMiddleware
app = FastAPI()
origins = ["*"]
app.add_middleware(
    CORSMiddleware,
    allow_origins=["*"],
    allow_credentials=True,
    allow_methods=["GET"],
    allow_headers=["*"],
)
model = joblib.load('model.pkl')
def collect_co_level():
    response = requests.get("http://127.0.0.1:8000/mq7_data")
    co_levels = response.json()
    avg_co_level = sum(co_levels) / len(co_levels)
    return avg_co_level
def collect_loaction():
    response = request.get("http://127.0.0.1:8000/location")
    loc = response.json()
    return loc
def collect_pm_level():
    response = requests.get("http://127.0.0.1:8000/pm_data")
    pm_levels = response.json()
    avg_pm_level = sum(pm_levels) / len(pm_levels)
    return avg_pm_level
def collect_temperature_humidity():
    response = requests.get("http://127.0.0.1:8000/dht_data")
    dht_data = response.json()
    temperatures_c = [entry["temperature_c"] for entry in dht_data]
    humidity_values = [entry["humidity"] for entry in dht_data]
    avg_temperature = sum(temperatures_c) / len(temperatures_c)
```

```
    avg_humidity = sum(humidity_values) / len(humidity_values)
    return avg_temperature, avg_humidity
@app.get("/prediction")
async def get_prediction():
    avg_co_level = collect_co_level()
    avg_pm_level = collect_pm_level()
    #loc_dat = collect_loaction()
    avg_temperature, avg_humidity = collect_temperature_humidity()
    latitude = 12.3456
    longitude = 78.9123
    sensor_data = [avg_co_level, avg_pm_level, avg_temperature, avg_humidity, latitude, longitude]
    new_data = [sensor_data]
    prediction = model.predict(new_data)
    prediction = int(prediction[0])
```

Webpage

```
from fastapi import FastAPI
from typing import List
import serial
import board
import 41dafruit_dht
from gpiozero import MCP3008
import csv
import pynmea2
import time
import os
import string
from datetime import datetime
app = FastAPI()
serial_port = '/dev/ttyUSB0'
baud_rate = 9600
ser = serial.Serial(serial_port, baud_rate, timeout=2)
dhtDevice = 41dafruit_dht.DHT22(board.D16)
adc = MCP3008(channel=0)
```

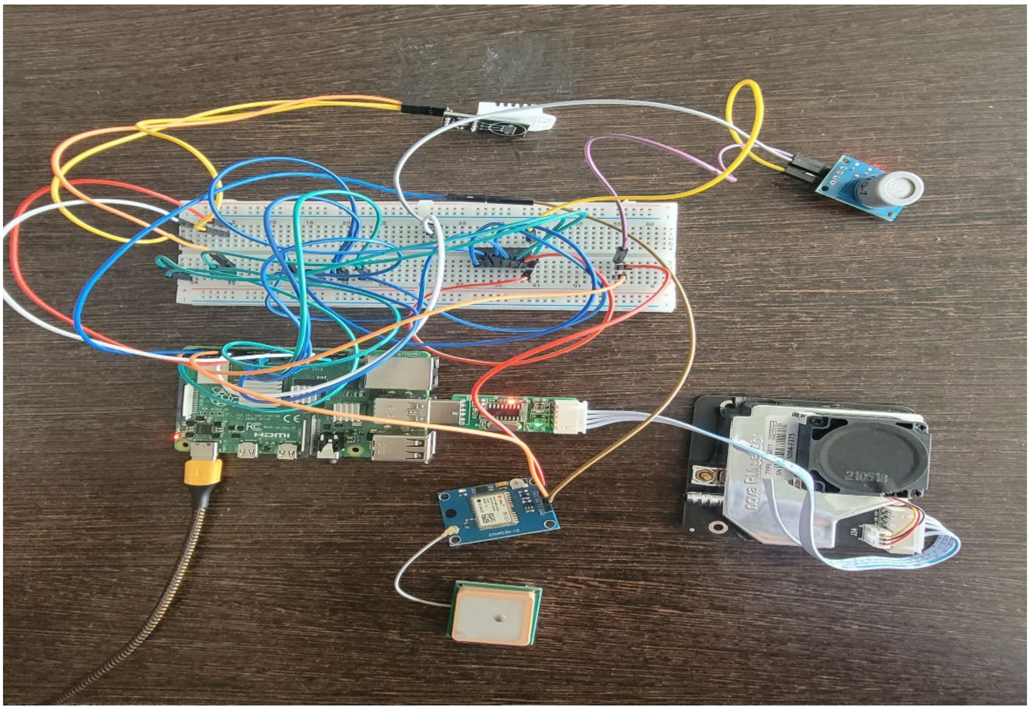
```
def read_pm_data() -> List[float]:
    """
    Function to read PM data from the Nova PM sensor.
    Returns a list of PM2.5 data points (up to 40).
    """
    pm_data = []
    try:
        ser.write(b'\Xaa\Xb4\x04\x13\x00\x01\Xe1')
        for _ in range(30):
            data = ser.read(10)
            pm25 = int.from_bytes(data[2:4], byteorder='little') / 10.0
            pm_data.append(pm25)
    except Exception as e:
        print("Error reading PM data:", e)
    write_to_csv(pm_data, "pm_data.csv")
    return pm_data[:30]
def read_dht_data() -> List[dict]:
    dht_data = []
    try:
        for _ in range(30):
            temperature_c = dhtDevice.temperature
            humidity = dhtDevice.humidity
            dht_data.append({"temperature_c": temperature_c, "humidity": humidity})
    except RuntimeError as error:
        print(error.args[0])
    write_to_csv(dht_data, "dht_data.csv")
    return dht_data[:30]
def read_mq7_data() -> List[float]:
    mq7_data = []
    try:
        for _ in range(30):
            analog_value = adc.value
  mq7_data.append(analog_value)
    except Exception as e:
```

```
        print("Error reading MQ7 data:", e)
    write_to_csv(mq7_data, "mq7_data.csv")
    return mq7_data[:30]
def write_to_csv(data: List, filename: str):
    mode = 'a' if os.path.exists(filename) else 'w'
    with open(filename, mode=mode, newline='') as file:
        writer = csv.writer(file)
        for item in data:
            timestamp = datetime.now().strftime("%Y-%m-%d %H:%M:%S")
            if isinstance(item, dict):
                item["timestamp"] = timestamp
                writer.writerow(item.values())
            else:
                writer.writerow([timestamp, item])
@app.get("/pm_data")
def get_pm_data() -> List[float]:
    return read_pm_data()
@app.get("/dht_data")
def get_dht_data() -> List[dict]
    return read_dht_data()
@app.get("/mq7_data")
def get_mq7_data() -> List[float]:
    return read_mq7_data()
@app.get("/location")
def get_location() -> str:
    port = "/dev/ttyS0"
    ser_gps = serial.Serial(port, baudrate=9600, timeout=0.5)
    dataout = pynmea2.NMEAStreamReader()
    newdata = ser_gps.readline()
    if newdata[0:6] == "$GPRMC":
        newmsg = pynmea2.parse(newdata)
        lat = newmsg.latitude
        lng = newmsg.longitude
        gps = "Latitude=" + str(lat) + " Longitude=" + str(lng)
```

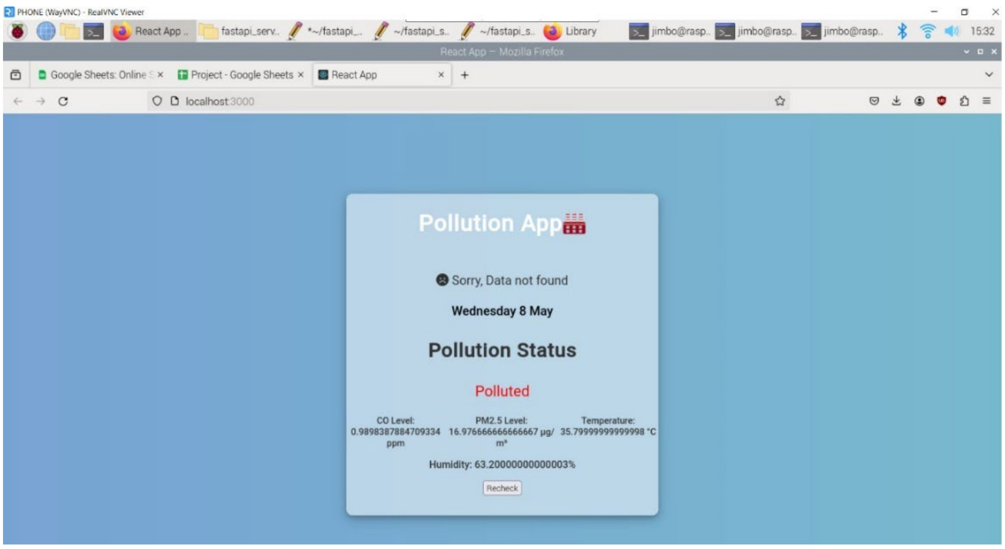
```
        return gps
    return "Location data not available"
if __name__ == "__main__":
    import uvicorn
    uvicorn.run(app, host="0.0.0.0", port=8000)
```

APPENDIX-2

IMPLEMENTED MODULE



POLLUTED OUTPUT



NON-POLLUTED OUTPUT