VIETNAM NATIONAL UNIVERSITY OF HO CHI MINH CITY
THE INTERNATIONAL UNIVERSITY
SCHOOL OF COMPUTER SCIENCE AND ENGINEERING

A fog-computing architecture for Indoor Air Quality Monitoring

By
Truong Nhat Minh Quang

A thesis submitted to the School of Computer Science and Engineering
in partial fulfillment of the requirements for the degree of
Bachelor of Science in Computer Science

Ho Chi Minh City, Vietnam
2022

A fog-computing architecture for Indoor Air Quality Monitoring

APPROVED BY:

THESIS COMMITTEE

Assoc. Prof. Vo Thi Luu Phuong,
Dr. Le Duy Tan,
Dr. Nguyen Thi Thanh Sang,
Dr. Le Hai Duong,

Acknowledgement

I would like to express my gratitude to Dr.Le Duy Tan and the principal Tran Tien Khoa for providing me with the opportunity to work on and complete this thesis on the amazing topic of "A fog-computing architecture for Indoor Air Quality Monitoring" They also assisted me in completing key studies. Mr. Le Duy Tan, in particular, has been extremely encouraging and understanding, which has really aided me in completing this thesis. It also helped me to get closer to my scientific interests. From that point on, I felt more confident in my abilities to continue working on this project, as well as other Internet of things projects in general.

Thank you, Truong Nhat Minh Quang.

Contents

List of Figures

List of Tables

Abstract

Air pollution has emerged as a significant concern in the twenty-first century, posing threats to both the environment and public health. Recent studies have delved extensively into air pollution and air quality monitoring, yet the field continues to grapple with unresolved challenges. This study presents a Fog-computing architecture tailored for indoor air quality monitoring, focusing on multiple pollution parameters.Our Internet of Things (IoT) system is designed to collect and monitor various pollutants, including PM2.5, CO2, CO, temperature, and humidity. This is achieved through the integration of STM32f429I-DISC1, NodeMCU, and various low-cost sensors. The collected data is then transmitted to ThingsBoard, serving as our fog-computing platform, utilizing the Raspberry Pi for monitoring tasks on the administrative side. The integration of fog computing allows for efficient and decentralized data processing, enhancing the system's responsiveness. Furthermore, the implementation of deep learning models adds a layer of sophistication, enabling the creation of a real-time online interface for monitoring. This interface not only visualizes the current air quality but also provides forecasts, offering valuable insights for clients. By leveraging this comprehensive approach, our system aims to address the existing challenges in air quality monitoring and contribute to a healthier and more informed living environment.

Chapter 1

Project Overview

1.1 Introduction

The Internet of Things (IoT) constitutes a network of interconnected devices facilitating the transmission of data without direct human or computer intervention. Recent strides in hardware, software, and networking technologies have profoundly shaped the way we live and work, with the IoT playing a pivotal role. Sensors, equipped to accurately perceive their environment, transfer data, and interface with designated IoT devices, enable a comprehensive understanding of our surroundings. This capability empowers us to promptly and accurately respond to various situations. Notably, one domain significantly influenced by IoT is the control and monitoring of air quality.

In the present era, creating an Internet of Things (IoT) system comes with practical challenges, particularly when it comes to evaluating and supervising the indoor air quality. This matter is increasingly crucial for maintaining human well-being, with a direct impact on health. Monitoring the levels of pollutants within enclosed spaces is essential to cultivate a healthy indoor environment and reduce the risks of respiratory issues. The significance of this concern is heightened in regions experiencing robust economic and industrial growth, exemplified by the situation in Vietnam. In Vietnam, a country currently grappling with a severe public health crisis, the population is experiencing a rapid deterioration in health due to rising air pollution levels. As indicated by research conducted by the Global Alliance on Health and Pollution, in 2017, Vietnam stood fourth in the Western Pacific for pollution-related deaths, with an alarming toll of over 71,365 lives. Strikingly, nearly 70%[1]of these deaths were linked to air pollution, creating a substantial public health burden. The severe consequences are predominantly evident in respiratory, tracheal, and cardiovascular illnesses, impacting vulnerable demographics such as the elderly, individuals with pre-existing respiratory conditions like asthma, and those from lower-income socioeconomic backgrounds. In light of these challenges, the implementation of effective IoT-based air quality monitoring systems becomes a pivotal measure to address and mitigate the urgent repercussions of indoor air pollution, not only in Vietnam but also in analogous regions facing comparable issues.

Recent publications in the field of air quality systems often rely on unprocessed sensor data, leading to a lack of consensus on data processing methods, especially when implementing low-cost air pollution sensors. Additionally, some air quality monitoring systems use expensive sensors, requiring manual intervention and proving unsuitable for IoT automation and cost reduction compliance.

Many of these studies involve the use of Arduino microcontrollers for tracking and data collection. However, due to the prolonged data-gathering process, Arduino faces challenges such as a shortened service life and difficulty in consistent long-term operation. To overcome these limitations and enhance monitoring and forecasting capabilities for air quality, our system specifically opted for an STM32 microcontroller.

Our system employs five semiconductor sensor modules connected to the STM32 board, measuring carbon dioxide (CO2), carbon monoxide (CO), small particulate matter (PM2.5), as well as temperature and humidity. Data from these sensors is transmitted to our ThingsBoard server using the ESP8266 NodeMCU Wi-Fi module. This server offers advanced computation and data analysis capabilities, providing flexibility for extending various features compared to other enterprise IoT platforms. The use of low-cost sensors and the ability to handle more complex tasks make our system more competitive in similar scenarios than its predecessors.

1.2 Problem Statement and Proposal

In contrast to some air quality monitoring systems that incorporate data purification, the results often exhibit unreliability and inaccuracies. Moreover, the escalating volumes of data generated by our low-cost sensors make traditional raw data cleaning methods increasingly impractical. Consequently, in our pursuit of a more widely applicable data processing approach, we introduce various machine learning models and algorithms, including Logistic Regession Classifier (LR), Decision Tree Classifier (DT), Multiple Random Forest Classifier (RF), and K-Neighbors Classifier (KNN). These are employed for classifying the Indoor Air Quality Index using the parameters collected. As previously mentioned, our air quality monitoring system aims to surpass other studies by incorporating more advanced features.

Many researches has been conducted on this topic with many proposal prove to have promising result. However, their execution is still far from feasible. To begin, some research employed an insufficient number of sensors, which might lead to unwanted inaccuracy in prediction. Next, many system disregard the use of machine learning into it's system, which is very counter-intuitive given that data gathered may be in the hundreds or millions, and having an employee on the duty of analyzing without the assistance of AI is impractical. It is worth noting that most architecture currently significantly favors the usage of cloud computing, while ignores its shortcoming, which is the connection. Cloud computing has a poor reaction time, and even a little internet outage might cause entire data loss during transmission. Furthermore, The population of Ho Chi Minh City is rising at an alarming rate [2], which consequently increases the number of cars on the road, the volume of products produced in industry, the amount of energy used, and the pollutants. As a result, a reliable IAQ monitoring and forecasting system that is both affordable and long-lasting is required.

To address these issue, this research propose a fog-computing architecture for IAQ monitoring and forecasting. Furthermore, on the aspect of edge devices, this project will be using more industry-friendly MCUs; such as, the STM32f429I-DISC1, and the NodeMCU. The Raspberry Pi, the fog-computing device, will be used link between the two levels, handling data filtering, light computational tasks, and data visualization. Finally, through the implementation of Deep learning we can forecast the data collected and create a website to display these findings.

1.3 Scope and objectives

The research described focuses primarily on indoor air quality (IAQ) within households, aiming to address the presence of pollutants such as dust, carbon monoxide, and carbon dioxide that pose significant health risks to families. A key aspect of this investigation involves deploying sensors as part of an air quality monitoring system to quantitatively measure these contaminants. The study adopts methodologies to address data gaps that may arise during data collection, with a core objective of leveraging machine learning techniques to develop air quality forecasting capabilities for online and mobile applications.

This research contributes to the field in several notable ways:

1. **Designing and Implementing Real-Time Monitoring System**: The study involves the design and creation of a real-time air quality monitoring system incorporating fog computing. Fog computing enhances data processing capabilities at the edge of the network, enabling more efficient and timely analysis of sensor data related to air quality.

2. **Enhancing Data Stability and System Optimization**: Recent methodologies are employed to improve the stability and reliability of collected air quality data. By optimizing data collection techniques, the study aims to ensure the accuracy and consistency of measurements over time.

3. **Utilizing Machine Learning for Data Analysis**: The research involves acquiring and training data using various machine learning models and algorithms. This approach enables the development of predictive models that can forecast air quality trends based on historical data, contributing to proactive air quality management.

4. **Developing Web and Mobile Applications**: To make air quality data accessible and actionable, the study includes the creation of web and mobile applications. These applications provide visualizations of real-time air quality metrics and incorporate forecasting functionalities, empowering users to make informed decisions about indoor environmental conditions.

5. **Implementing IoT Platform with Open-Source Technologies**: The construction and implementation of an Internet of Things (IoT) platform using open-source technologies further democratizes access to air quality monitoring solutions. Open-source platforms foster collaboration and innovation within the IoT community, enabling the development of scalable and interoperable systems.

Overall, this research demonstrates a comprehensive approach to addressing indoor air quality challenges through the integration of advanced sensor technologies, machine learning techniques, and IoT infrastructure. By focusing on real-time monitoring, data stability, predictive analytics, and user-friendly applications, the study aims to advance our understanding of indoor air quality dynamics and empower individuals and communities to take proactive measures towards healthier indoor environments.

Chapter 2

Background and Related Works

The rapid population growth of the world economy [22] [32] has contributed to increased vehicular traffic, industrial activities, energy consumption, and other factors affecting air quality. This surge in population has significant implications for indoor air quality, thereby impacting respiratory health.

In Vietnam, air pollution has emerged as a significant contributor to respiratory and other pulmonary diseases, as highlighted in the study by Dhondt et al. in 2011 [8]. The escalating concerns surrounding this issue are exacerbated by the increase of vehicles on the roads, intensifying the severity of air pollution. The escalating presence of pollutants not only in the outdoor environment but also within enclosed spaces has heightened the gravity of the situation, particularly with regards to indoor air quality. The confluence of increased vehicular emissions and the state of indoor air pollution has created a pressing and serious public health challenge in Vietnam.

Indoor air pollution, primarily stemming from the household use of solid fuels for cooking and heating, is a pervasive issue globally and poses significant health risks, especially in developing countries. Solid fuels, such as wood, charcoal, and coal, are extensively used in these regions, affecting approximately 3 billion people. The prevalence of solid fuel use aligns closely with socioeconomic development, and it is notably associated with poverty.

In developing countries, where solid fuels are often burned inefficiently in open fires, high levels of indoor and local air pollution are observed, surpassing international standards for ambient air quality. Particulate matter, carbon monoxide, and various toxins present in biomass smoke contribute to health risks, particularly for homemakers and young children. The exposure to hazardous substances from indoor air pollution significantly contributes to the global burden of disease in developing countries.

The use of biomass stoves, a source of indoor pollution in industrialized countries as well, is typically more efficient and vented. However, it still produces pollutant concentrations above standard limits, impacting health and contributing to outdoor pollution. The situation underscores the urgent need for reliable measures to address indoor air pollution, especially in developing nations, to mitigate its detrimental effects on respiratory health and overall well-being.

Several pollutants and toxic chemicals have been identified as having adverse effects on human health, particularly in the realm of indoor air quality, as elucidated in the study conducted by Tran et al. in 2020 [29]. The indoor environment is reported to be highly contaminated, raising concerns about the potential impact on the well-being of individuals exposed to such pollutants.

Given these circumstances, there is an urgent need to implement reliable and unbiased automated systems for monitoring indoor air quality. Insights from these systems can inform strategies to mitigate the adverse effects on respiratory health by adopting measures that enhance indoor air quality and promote overall well-being.

Several research in this subject have been undertaken to monitor and forecast the indoor air quality:

In alignment with the fundamental requirements of accuracy and affordability in IoT systems, the insights gleaned from [23] have been instrumental in guiding our sensor selection for this study. The cited research emphasizes the significance of sensors that not only meet high accuracy standards but are also economically viable. Incorporating sensors that fulfill both these criteria is paramount to ensuring the effectiveness and accessibility of the IoT system under investigation.

Furthermore, the study recommends the integration of machine learning techniques for the analysis and prediction of the collected data, highlighting the potential for advanced data processing and predictive modeling. By incorporating machine learning into our methodology, we aim to leverage the capabilities of these algorithms to enhance the interpretability of the collected data and provide valuable insights for future applications. This holistic approach, combining accurate and affordable sensors with machine learning analytics, forms the foundation of our study for robust and insightful results.

J. Jo et al. [15] developed an IoT-based infrastructure for monitoring air quality, as well as a web site and a mobile application for users to examine air quality. A device known as "Smart-Air" is being developed, which has several sensors that collect data on air quality such as CO2, CO, smoke, dust, and temperature/humidity. Through a Long Term Evolution (LTE) modem, the collected data is transferred to the website for further processing.

Similarly, S. McGrath et al. [18] propose another method for monitoring air quality that records four air pollutants: ozone, PM 2.5, NO2, and CO. To display the state of air quality, a firebase web application and a mobile application are being developed. H. P. L. de Medeiros et al. [19] discuss a plan to build an IoT-based system for air quality monitoring that comprises PMSA003, MICS-6814, and MQ-131, which record air pollutants such as Ozone, PM 2.5 and PM 10, CO, NO2, and Ammonia. The devices illustrated here provide the hardware architecture for monitoring air quality, but there is no data processing or analysis.

D. Zhu et al. [34] proposed a machine learning-based air quality forecasting system capable of making predictions for up to 24 hours. This study focused on three air pollutants: SO2, PM 2.5, and ozone. The data was collected from two separate meteorological stations in Illinois' two distinct suburb residential zones. The authors developed revised models for forecasting IAQ using advanced regularization methods to improve performance in this paper.

J. Saini et al. [24] built an IoT-based monitoring system for measuring crucial IAQ parameters, and the prediction system was enhanced with Keras optimizers. The system makes use of low-cost sensors, but only PM 10 and PM 2.5 are employed to analyze IAQ.

Acknowledging the critical role of system design in our study, particular emphasis has been placed on the integration of machine learning techniques to classify air quality indices, enabling informed decisions regarding the environmental conditions. The studies conducted by Wei et al. in 2019 [31] and Zheng et al. in 2018 [33] have been instrumental sources of inspiration, providing valuable insights into the application of machine learning for air quality assessment.

Drawing from these studies, our approach involves leveraging machine learning algorithms to categorize air quality indices, distinguishing between favorable and unfavorable conditions. The methodologies and techniques proposed in [31] and [33] serve as guiding frameworks for the implementation of robust classification models, thereby enhancing the accuracy and efficiency of our system in evaluating and characterizing air quality levels. This integration of machine learning principles represents a strategic step towards achieving a more sophisticated and insightful analysis of the collected data in the context of our study.

Drawing upon the insights gleaned from the exploration of nodeMCU utilization within an Internet of Things (IoT) framework, as outlined in the research articles by Aziz et al. (2018) [3] and Chanthakit et al. (2018) [6], we are poised to formulate an innovative and cost-effective system. By leveraging the knowledge derived from these studies, our intention is to design a solution that is not only accessible to students but also conducive to small businesses. Through this approach, we aim to empower a broader demographic with the ability to develop and implement IoT systems [20], thereby fostering technological learning and encouraging entrepreneurial initiatives in a more affordable manner.

The collaboration between CeADAR and Vietnam National University - Ho Chi Minh City resulted in the development of the HealthyAir app for Android and iOS platforms, designed to monitor air quality. However, the app lacks features such as a CO concentration meter and air quality predictions. Ho Chi Minh City faces challenges with only six monitoring stations, and similar institutions in developing nations, including Vietnam, have seen partial maintenance, ultimately leading to their shutdown in 2020. The data collected from these systems was sporadic, unreliable, and lacked proper evaluation.

Despite numerous studies focusing on outdoor air quality, there remains a significant research gap in addressing the critical issue of indoor air pollution, particularly in developing countries like Vietnam. While outdoor air quality assessments are vital, the pervasive use of solid fuels for cooking and heating in households presents a substantial health risk due to indoor pollution. The combustion of solid fuels releases hazardous pollutants, impacting the respiratory health of homemakers and young children. Despite its evident importance, indoor air quality in developing countries has received comparatively less attention in research endeavors. This research aims to contribute to bridging this gap by investigating the specific dynamics of indoor air pollution, particularly the sources, concentrations, and health impacts, in order to develop effective strategies for mitigating the adverse effects on public health in the context of developing nations like Vietnam.

Furthermore, these reasons motivate us for the development of infrastructure to address the longstanding indoor air quality challenges in Ho Chi Minh City. Recognizing the enduring nature of these issues, we propose the creation of a system capable of assessing various indoor air quality indicators and delivering users more precise and relevant data for their immediate surroundings. Beyond simple classification of data into healthy or unhealthy categories, this system integrates predictive features to furnish insightful information on air quality conditions. Furthermore, it includes a cautionary mechanism to alert users in the event any indoor air parameter exceeds safty levels. To bring this envisioned indoor air quality system to fruition, we introduced real-time monitoring with fog computing, leveraging WiFi technology for the identification and quantification of indoor air contaminants. Subsequently, the collected data seamlessly uploads to an online platform for visualization and in-depth analysis.

Chapter 3

System design

3.1 Indoor Air Quality Parameters selection

Building upon the discoveries outlined in the articles referenced by [16] and [28], we introduce a system for monitoring indoor air quality. This system focuses on overseeing key parameters, namely CO, CO2, PM2.5, Temperature, and Humidity. The proposed system is crafted for monitoring air pollutants, with a focus on those that predominantly affect indoor air quality. These selected pollutants are not only prevalent but also easily traceable, making them optimal candidates for in-depth study and monitoring within the realm of air pollution. The following air pollutants have been found for the proposed indoor air quality monitoring:

- **Carbon Dioxide** (CO2): CO2 is regarded as a form of air pollution primarily stemming from heightened road traffic. Elevated concentrations of carbon dioxide can lead to various health concerns, including headaches, dizziness, shortness of breath, increased heart rate, and fatigue. This is particularly relevant in indoor environments where insufficient ventilation may contribute to the buildup of CO2 and impact overall air quality [17].

- **Carbon monoxide** (CO) [4]: The main contributors to elevated indoor CO levels are gas leaks and vehicle exhaust. Prolonged exposure to vehicle emissions can lead to CO poisoning and cardiovascular problems, particularly risky for elders and children. However, gas leaks pose even graver threats. If not properly addressed, they can result in health issues, as mentioned earlier, and, more critically, potential fires or explosions, as frequently reported in the media.

- **Particulate Matter** (PM2.5)[32]: The detrimental impacts of elevated PM2.5 levels on indoor air quality include respiratory issues like coughing and shortness of breath. Prolonged exposure may worsen existing conditions such as asthma and contribute to cardiovascular problems. Additionally, these microscopic particles can lead to eye irritation, decreased lung function, and an increased susceptibility to respiratory infections. Hence, it is imperative to implement measures to lower indoor PM2.5 concentrations for the well-being of indoor occupants.

- **Temperature and Humidity** [4]: High indoor temperatures and humidity levels can lead to discomfort, respiratory issues, dehydration, and an increased risk of heat-related illnesses. These conditions create an environment conducive to the growth of allergens like mold and dust mites, potentially triggering respiratory problems. Additionally, the warmth and humidity foster bacterial growth, impacting indoor air quality and posing health risks. To address these concerns, it is crucial to maintain proper ventilation, control indoor humidity, and implement cooling measures, such as air conditioning and fans, to ensure a healthier and more comfortable indoor environment.

3.2 System Design

The suggested systems comprise an Internet of Things (IoT) node, consisting of a sensors gathering component, a communication component for data transfer, and an upper computer component responsible for storing and managing website content. These essential components constitute the foundational building blocks of the proposed system. The comprehensive architecture diagram depicting this structure is illustrated in Figure 1.

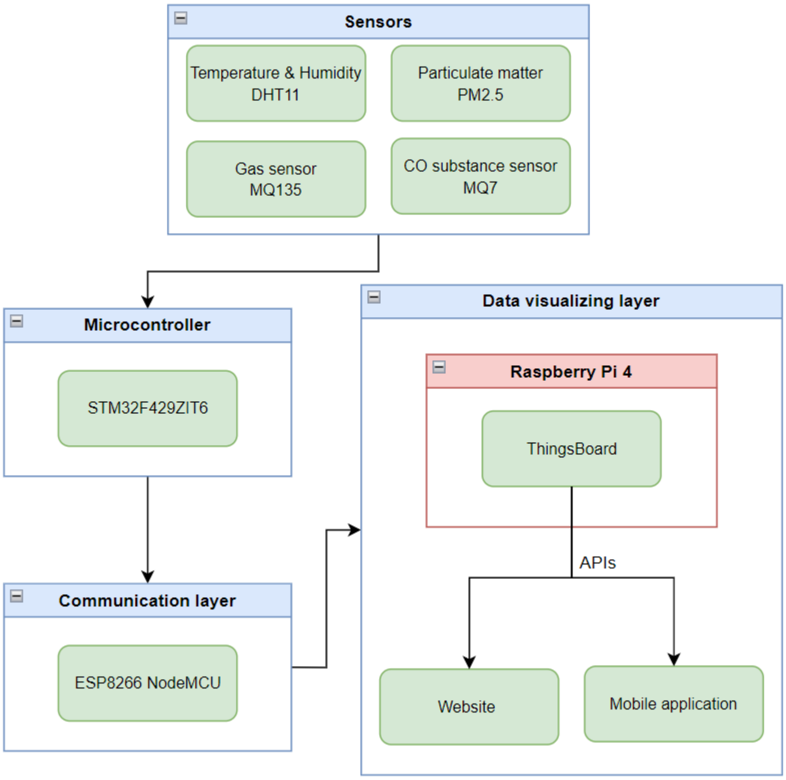

Figure 1: The general architecture of the IOT system

The operational framework of the air quality monitoring system outlined in this study involves a sequence of distinct stages, each contributing to the efficient collection, transmission, and analysis of sensor data for indoor air quality assessment. The system's workflow can be delineated into three primary stages:

1. **WiFi Connection Establishment**: The system initiates by establishing a WiFi connection, leveraging the ESP8266 WiFi NodeMCU module. This component facilitates wireless communication and network connectivity, enabling the system to transmit data seamlessly over the internet.

2. **Connection to ThingsBoard API**: Once the WiFi connection is established, the system interfaces with the ThingsBoard API. ThingsBoard, hosted on a Raspberry Pi 4, serves as the central data repository and visualization platform. The integration with ThingsBoard allows for efficient data management and real-time monitoring capabilities.

3. **Data Collection and Transmission**: The STM32F429ZIT6 microcontroller, equipped with both analog and digital inputs, plays a pivotal role in sensor data collection. At predefined intervals (e.g., every five minutes), the microcontroller gathers data from connected sensors, structures the information, and transmits it to the ESP8266 WiFi NodeMCU for onward transmission to the ThingsBoard server.

4. **Automated System Access**: The system incorporates a program uploaded into the WiFi module to automate WiFi connection and server access. This program stores and manages necessary authorization information, ensuring seamless and secure data transmission without manual intervention.

5. **Data Visualization and Analysis**: Upon receiving sensor data, ThingsBoard visualizes and analyzes the information in real-time. The platform provides graphical representations, historical trends, and predictive analytics related to air quality parameters. The use of ThingsBoard on a Raspberry Pi 4 not only facilitates cost-effective data hosting compared to cloud platforms but also supports customized applications tailored to air quality monitoring and prediction.

This operational framework underscores the system's efficiency in integrating hardware components (microcontroller, WiFi module) with cloud-based IoT platforms (ThingsBoard) to enable robust and scalable indoor air quality monitoring. By automating data acquisition, transmission, and analysis processes, the system empowers users to make informed decisions and take proactive measures to ensure healthier indoor environments. Additionally, the emphasis on cost reduction and program efficiency reflects a strategic approach to deploying IoT solutions for air quality management in residential settings.

3.3 Hardwares Design

In the context of air quality data collection, the utilization of five distinct sensors, each with its own unique input and output pin configuration, represents a multifaceted approach to monitoring indoor environmental conditions. While the setup process typically requires configuring the output pin of these sensors, the complexity and variation in pin configurations necessitate careful consideration and implementation to ensure accurate and efficient data collection.

The selection and integration of these sensors into the monitoring system exemplify a strategic decision aimed at capturing a comprehensive range of air quality metrics. These sensors may include components for measuring particulate matter (such as dust), carbon monoxide (CO), carbon dioxide (CO2), volatile organic compounds (VOCs), and other relevant pollutants commonly found indoors. Each sensor's specific pin configuration, which may involve analog or digital input and output, reflects the diversity of measurement parameters and methods required for robust air quality assessment.

To streamline the setup process and optimize sensor performance, considerations such as calibration, signal processing, and data interpretation are crucial. While configuring the output pin is typically essential for establishing communication and acquiring sensor readings, the study highlights that certain sensors may not require specific input pin settings, simplifying the integration process in some instances.

Moreover, the deployment of multiple sensors within the monitoring system underscores the importance of data fusion and sensor fusion techniques. By combining information from various sensors, the system can generate comprehensive insights into indoor air quality dynamics, including pollutant concentrations, trends, and potential health risks.

In addition to sensor hardware considerations, software development plays a pivotal role in transforming raw sensor data into actionable information. This includes implementing algorithms for data preprocessing, calibration, and fusion, as well as designing user interfaces for real-time data visualization and analysis. The study's emphasis on leveraging machine learning techniques for data analysis and forecasting further enhances the system's capabilities, enabling predictive insights into air quality trends and potential interventions.

Overall, the integration of multiple sensors with unique pin configurations represents a sophisticated yet essential aspect of air quality monitoring. Through systematic setup, calibration, and data processing, the study aims to develop a robust monitoring framework capable of providing accurate, timely, and actionable information to promote healthier indoor environments and improve overall quality of life.

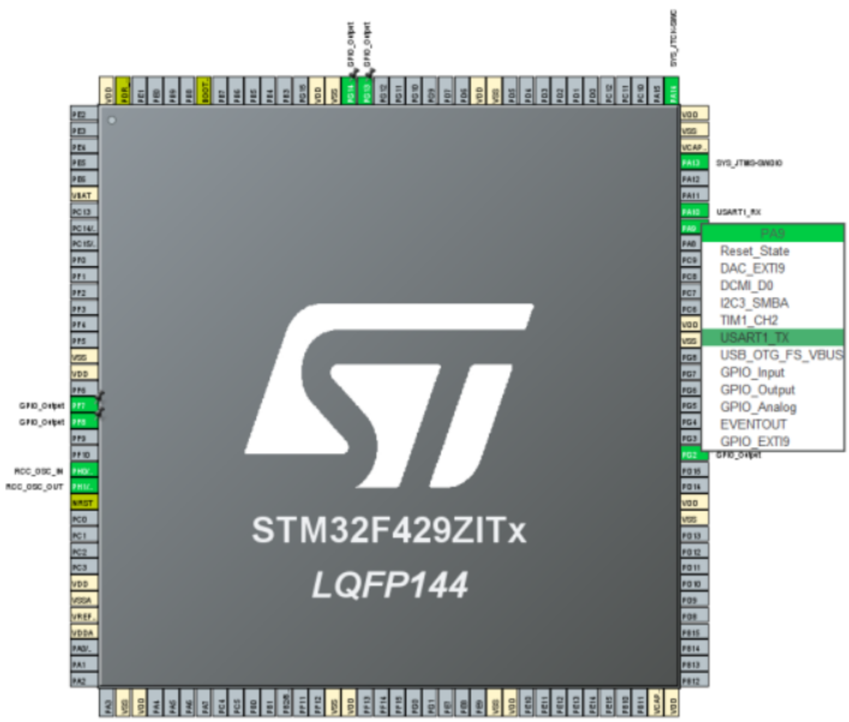

Figure 2: The pin Configuration of the STM32F429ZIT6 from the CubeIDE

The STM32CubeIDE offers a user-friendly interface that virtually represents its input-output pins and the STM32 microprocessor, as shown in Figure 2. Furthermore, it simplifies code generation by selecting the pin after each configuration. Given the involvement of multiple low-cost sensors from different manufacturers, each sensor operates independently based on the configuration specified by its vendor. This study provides detailed insights into the processing mechanics of each sensor's data.

3.3.1 Temperature and Humidity

The DHT11 temperature and humidity sensor, designed for ease of use, incorporates a built-in 5.1k resistor. Utilizing a 1-wire connection interface, the module ensures precise data through embedded signal preprocessing, eliminating the need for additional calculations. Operating on a 5-volt DC voltage, the sensor employs a moisture-retaining substrate as a dielectric for humidity data collection. Capacitance value changes, corresponding to humidity variations, are measured and transformed into digital form by the integrated circuit (IC). For temperature measurement, the sensor uses a thermistor with a negative temperature coefficient, where rising temperature results in decreased resistance. The sensor's output pin connects to the microprocessor's PF9 data pin, with its Vcc and GND pins linking to the STM32's 5-volt power and ground pins, respectively.

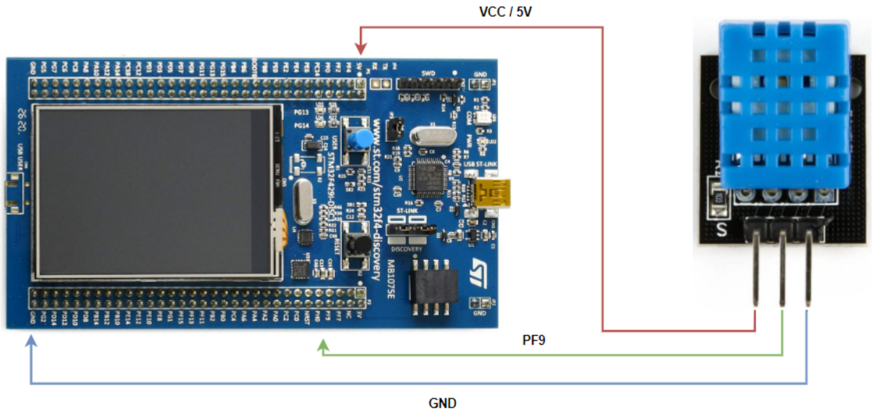

Figure 3: DHT11 and STM32 circuit design

3.3.2 Carbon Monoxide

The MQ7 sensor, a cost-effective semiconductor device meticulously engineered for the purpose of carbon monoxide detection, operates efficiently within a concentration range spanning from 20 to 2000 parts per million (ppm). This sensor distinguishes itself by employing SnO2 as its sensing material. Notably, the sensor's conductivity in pure air is relatively poor; however, it compensates for this limitation by demonstrating a rapid response time and remarkable sensitivity to carbon monoxide.

One of the distinguishing features of the MQ7 sensor lies in its provision of both analog and digital output signals. To harness these outputs effectively, the results from the sensor necessitate a conversion into a digital format, enabling subsequent measurement and calculation processes. In the context of connectivity with the STM32 microcontroller, the MQ7 sensor relies on three main wires for effective communication.

In a manner reminiscent of its predecessor, the MQ7 establishes a connection by linking its VCC and GND to the 5V and GND pins of the STM32, respectively. Additionally, the sensor's output is intricately linked to the PA2 pin of the STM32, thereby facilitating a seamless integration into the broader system for comprehensive carbon monoxide monitoring and analysis.

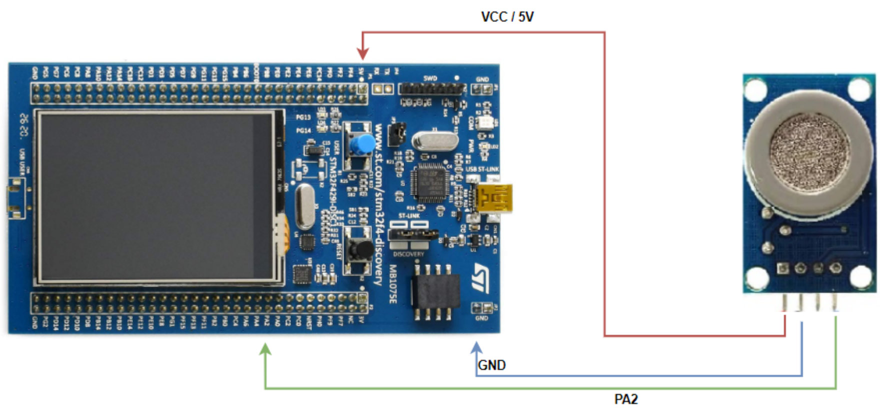

Figure 4: MQ7 and STM32 circuit design

3.3.3 Carbon Dioxide

The MQ135 sensor holds significant relevance in air quality monitoring equipment deployed within indoor environments and workplaces. Its versatility is attributed to its capability to detect a range of harmful substances, including NH3, NOx, alcohol, benzene, smoke, and CO2. Given its proficiency in identifying these pollutants, which pose potential health risks, the MQ135 is aptly labeled as an air quality sensor. Characterized by its sensitivity and swift responsiveness, this sensor is a key component in many smoke and gas detectors, often featuring thermostats to fine-tune sensitivity settings.

Much like its counterpart, the MQ7 sensor, the MQ135 sensor generates an analog output within the range of 0 to 1023. To effectively interpret and utilize this output, it becomes imperative to convert the analog signal into a digital format, particularly when expressing the results in parts per million (ppm) – a measure of the concentration of the detected substance. Given that different environments exhibit varying levels of pollution, the MQ135's ppm values provide insights into the specific concentrations of substances such as CO2. For instance, well-ventilated indoor spaces are anticipated to maintain a typical CO2 concentration of around 400 ppm. However, it is essential to acknowledge that indoor CO2 levels can fluctuate based on factors like ventilation, occupancy, and the nature of activities taking place within the enclosed space.

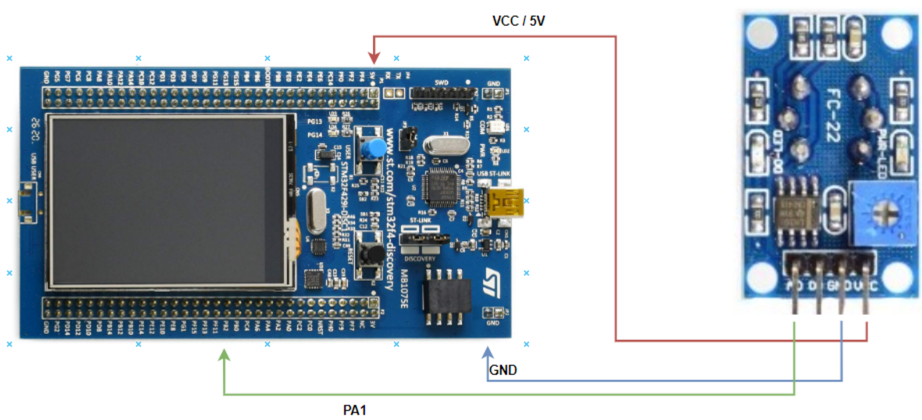

Figure 5: MQ135(back) and STM32 circuit design

As evident in the configuration, the output analog pin of the MQ135 is linked to the

PA1 of the STM32, positioned immediately adjacent to the PA2 pin (to its left). Both of these sensors share similar characteristics in that they produce analog values, necessitating conversion from the analog voltage form to a more interpretable measurement in parts per million (PPM) which will be covered in later part of the thesis.

3.3.4 Particulate Matter 2.5

The GP2Y1010AU0F stands as a specialized dust sensor module meticulously crafted to identify and quantify the concentration of airborne particles, with a specific focus on fine particles, notably those categorized as PM2.5. This innovative module operates by leveraging the combined capabilities of an infrared LED and a phototransistor, working in tandem to precisely gauge the extent of scattered light resulting from the passage of particles through the sensor apparatus. The utilization of this intricate combination of technologies underscores the sensor's nuanced ability to provide accurate measurements and insights into the presence of fine airborne particulate matter, contributing to a comprehensive understanding of air quality.

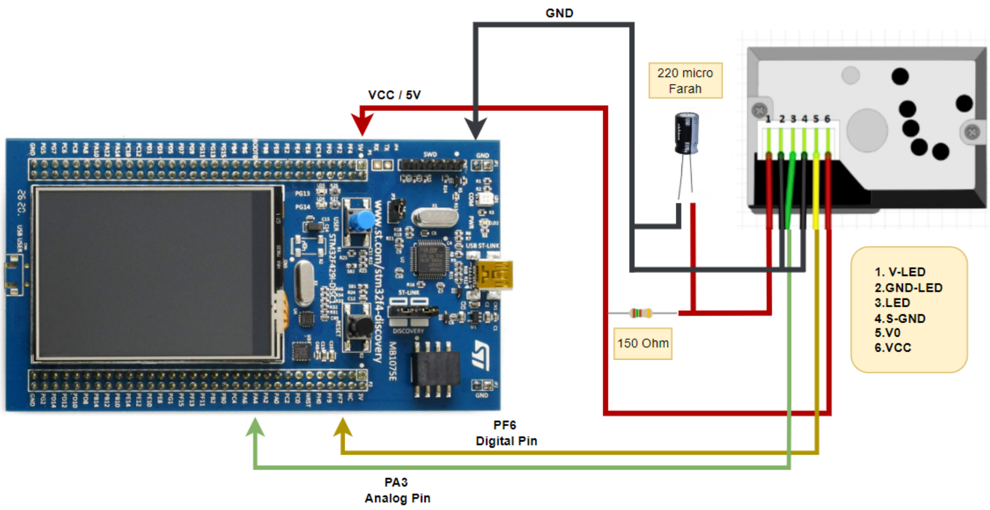

Figure 6: PM2.5 and STM32 circuit design

This sensor is widely employed in air quality monitoring devices to evaluate the concentration of particulate matter in the air. It generates a voltage output that is directly proportional to the particle concentration, facilitating precise measurements. The GP2Y1010AU0F is recognized for its compact design, user-friendly nature, and applicability to indoor air quality assessment and pollution monitoring, making it particularly relevant for research focused on monitoring indoor environments.

3.3.5 Linear Relation between the Voltage levels and the Measured values

The data collected from MQ135 and MQ7 sensors is presented in the form of analog signals, denoting voltage within the specified range of 0 to 5 Volts (1). As indicated by the user manuals of these sensors, MQ135 encompasses a detection range extending from 0 to 1023 parts per million (PPM), while MQ7 covers a range from 20 to 2000 PPM (2). Introducing variables, with x representing the voltage level and y representing the PPM level, reveals the existence of a linear relationship between these two parameters.

Expressed mathematically, the voltage level (x) and PPM level (y) exhibit a direct and proportional connection. The linear equation that captures this relationship can be derived based on the information provided in (1) and (2). This equation becomes a valuable tool for interpreting and extrapolating PPM levels from voltage readings, elucidating the concentration of substances detected by these sensors as the voltage level increases.

$$y(\text{PPM}) \sim x(\text{Voltage}) \tag{3.1}$$

or:

$$y = ax + b \tag{3.2}$$

Based on the information in (2), we can solve this system of linear equations for the MQ7:

$$20 = a \cdot 0 + b \tag{3.3}$$

$$2000 = a \cdot 5 + b \tag{3.4}$$

To find the values of a and b for this linear system, you can use the given points $(x_1, y_1) = (0, 20)$ and $(x_2, y_2) = (5, 2000)$.

Plug in the values for the points into the equation:

1. For $(x_1, y_1) = (0, 20)$:

$$20 = a \cdot 0 + b$$

This simplifies to $b = 20$.

2. For $(x_2, y_2) = (5, 2000)$:

$$2000 = a \cdot 5 + 20$$

Subtract 20 from both sides:

$$1980 = 5a$$

Divide both sides by 5:

$$a = 396$$

So, the values of a and b are $a = 396$ and $b = 20$. Thus the linear equation for the MQ7 is:

$$y = 396x + 20 \tag{3.5}$$

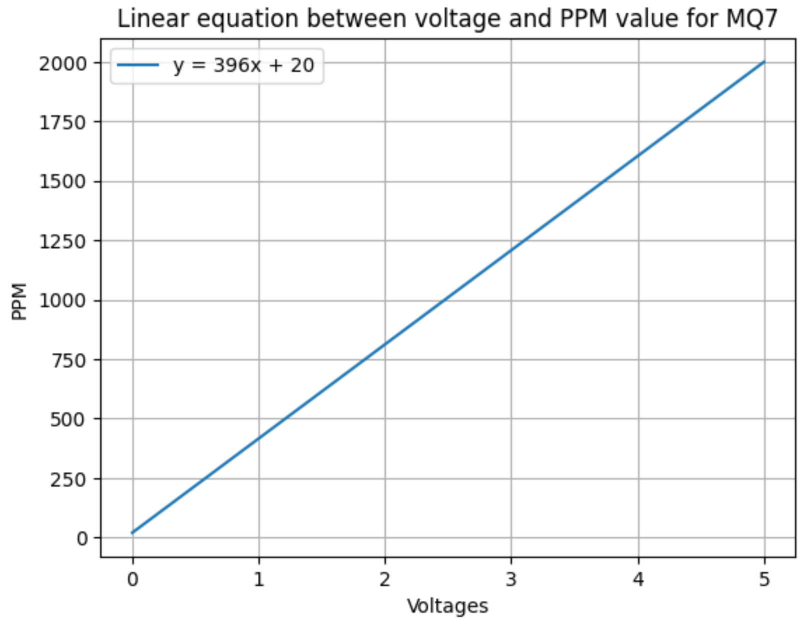

Figure 7: Linear relation of MQ7

Similarly, we can also conclude the linear equation for MQ135 is:

$$y = 202.6x + 10 \tag{3.6}$$

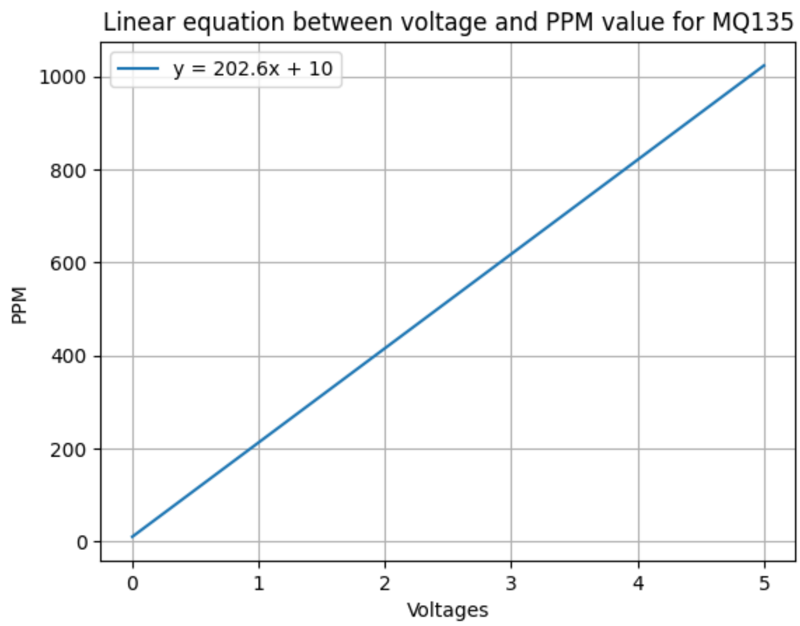

Figure 8: Linear relation of MQ135

3.4 Communication layer

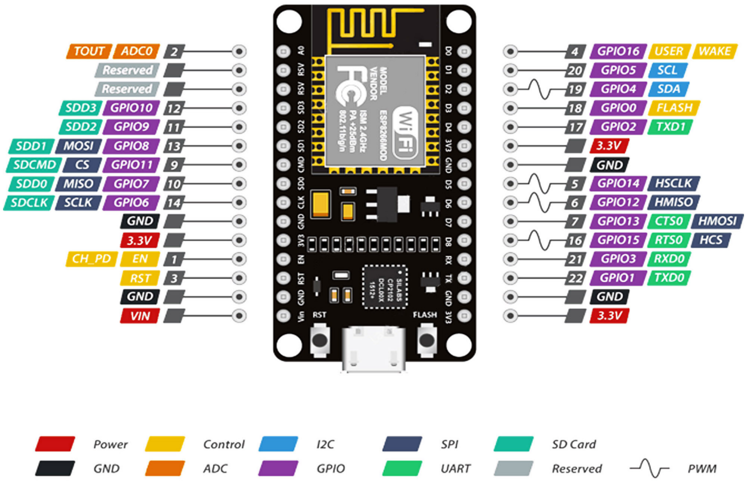

Figure 9: The architecture of the ESP8266 NodeMCU [2]

NodeMCU, an open source LUA-based firmware, was created for the ESP8266 wifi chip. NodeMCU firmware is included with the ESP8266 Development board/kit, also known as the NodeMCU Development board, in order to explore ESP8266 chip capability [18]. In the context of this project, the NodeMCU is predominantly chosen for its internet connectivity capabilities and ease of integration with the Arduino IDE. Below is an outline of how the device is connected within the system.

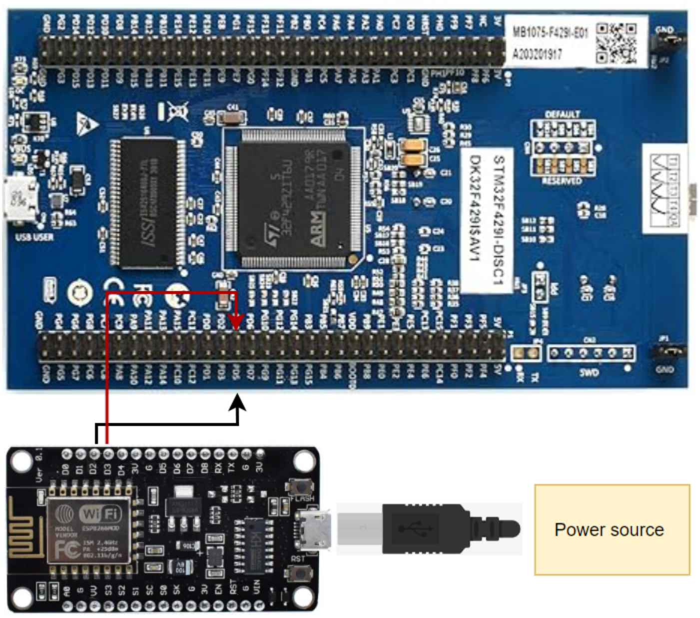

Figure 10: ESP8266 and STM32 (back) Circuit Design

The NodeMCU's utilization is driven by its proficiency in establishing internet connectivity, making it well-suited for applications where seamless communication with online platforms or networks is crucial. Its compatibility with the Arduino IDE further simplifies the development process, allowing for a more user-friendly and accessible programming environment. This combination of features makes the NodeMCU a valuable component in the project's architecture, contributing to the overall efficiency and functionality of the system.

The ESP8266 and the STM32 are intricately linked through the UART (Universal Asynchronous Receiver-Transmitter) protocol, employing PD5 and PD4 as the specified UART pins. To provide further detail, these pins are intricately connected to D2 and D3 on the ESP8266, establishing a dedicated communication channel between the two devices. This configuration facilitates a smooth exchange of data between the ESP8266 and the STM32, fostering effective coordination and integration of their individual functionalities. The utilization of UART ensures dependable and asynchronous serial communication, allowing for the streamlined and efficient transmission of information between these interconnected devices.

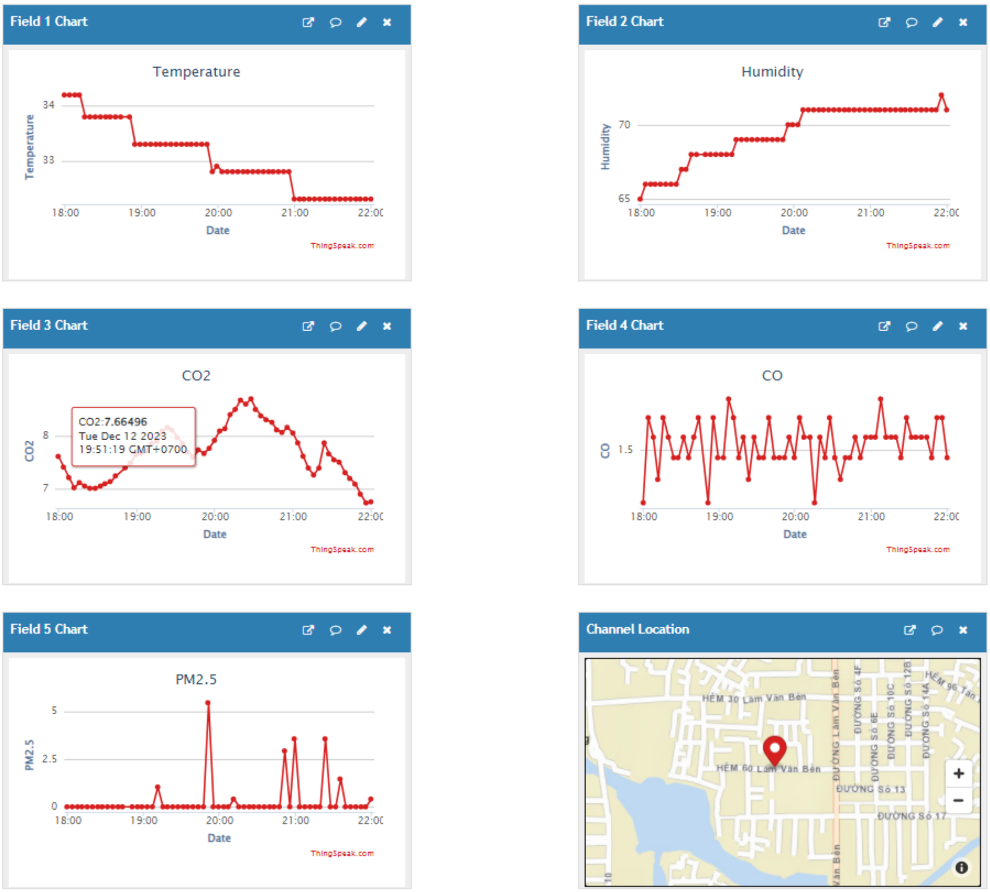

Figure 11: The representation of collected data on The ThingSpeak server

3.5 Calculation of an composite index Indoor Air-Quality Index

The indoor environmental standards are based on the current maintenance standards and recommended standards and are divided into five grades based on effects on humans [20–22], as shown in Table 3 below

Index	A (0–20)	B (21–40)	C (41–60)	D (61–80)	E (81–100)
Sort	Good	Moderate	Unhealthy for sensitive groups	Bad	Hazardous

Table 3.1: Scale of 1–100 points and color-coded index interval

The delineation of categories labeled as "unhealthy for sensitive groups" in the index is determined based on health effects measured against concentrations, such as standard epidemiological data and indoor environmental standards. These categories are computed by taking into account factors such as background concentration, exposure frequency, and compliance within each indoor space. This study aligns with the approach seen in both domestic and international environmental index evaluations, where a higher score signifies poorer air quality. Consequently, using a scale ranging from 1 to 100 points, it was concluded that a higher score indicates worse indoor air quality and a heightened risk to human health.

In Section 2.4.1, the PM2.5 index is computed, while Section 2.4.2 details the calculation process for the CO2 index. Additionally, a new section, Section 2.4.3, is introduced to outline the calculation of the CO index. In the IAQI index section (Section 2.4.4), the comprehensive indoor environment index is finalized, considering the PM2.5, CO2, and CO indices. Section 2.4.5 then provides a detailed classification from A to E, offering a comprehensive breakdown of the results.

3.5.1 Calculating the IAQI Index for CO, CO2, and PM2.5

Component	**Good**	Moderate	**Unhealthy for sensitive groups**	**Bad**	**Hazardous**
	A (0–20)	B (21–40)	C (41–60)	D (61–80)	E (81–100)
PM2.5 (mg/m^3)	0–0.015	0.016–0.035	0.036–0.05	0.051–0.08	>0.081
CO (ppm)	0–9	9–35	35–200	200–400	>400
CO2 (ppm)	0–450	451–700	701–1000	1001–3000	>3000

Table 3.2: Scale of 1–100 points and color-coded index interval

Concerning PM2.5, the indices were evaluated in consideration of their impact on human health. For the *"unhealthy for sensitive groups"* index section, which is the primary focus of this study, an integrated index range was established within a concentration of $120\,mg/m^3$. This concentration corresponds to a 10% increase in the prevalence of acute chronic obstructive pulmonary disease (COPD) in the general public. Additionally, based on the United States Environmental Protection Agency (US EPA) Air Quality Index (AQI) threshold of $80\,mg/m^3$, which is associated with a 10% increase in the prevalence of COPD in children, an integrated index section was set.

The classification of "bad" and "hazardous" labels was determined by adhering to particulate matter warnings and alarm standards applicable to the standard domestic atmospheric conditions. It's noteworthy that there is no available data on the effects of concentrations exceeding the upper limit for "hazardous." The upper limit for "hazardous" is estimated to be 600 mg/m^3, aligning with the upper limit defined by the US EPA AQI.

For CO2, the delineation between "good ($0–450\,ppm$)" and "moderate ($451–700\,ppm$)" ranges was established by referencing relative risk data obtained from epidemiological studies. The threshold for "moderate ($451 - 700\,ppm$)" corresponds to the level at which wheezing may occur. The classification of "unhealthy for sensitive groups (701–1000 ppm)" was determined based on a combination of effects baseline among relative risks, the long-term recommended levels by the American Society of Heating, Refrigeration, and Air-Conditioning Engineers (ASHRAE) (set at 700 ppm), and the cut-off level stipulated by indoor environmental standards (1000 ppm). Additionally, the categorization from "bad (1001 – 3000 ppm)" to "hazardous (>3000 ppm)" relied on high-concentration human impact data sourced from ASHRAE data.

In addition to the existing pollutants mentioned above, the inclusion of Carbon Monoxide (CO) is extremely crucial, especially in developing countries like Vietnam, where the prevalent use of traditional cooking fuels such as gas, charcoal, and wood raises concerns about elevated CO levels. This choice stems from the necessity to address health risks associated with the combustion of these fuels, emphasizing the significance of monitoring CO levels in indoor environments heavily reliant on such materials.

Thus for the context of Carbon Monoxide (CO) concentration levels indoors, our study aligns with the guidelines established by the World Health Organization (WHO). The WHO provides specific recommendations for indoor CO concentrations, serving as a pivotal reference for our investigation. Mirroring the structure used for CO2, we classify CO levels into distinct ranges: "good ($0-9, ppm$)," "safe ($10-24, ppm$)," "moderate ($25-49, ppm$)," "unhealthy for sensitive groups ($50-99, ppm$)," and further escalate the classification to "bad ($100-199, ppm$)" and "very bad ($>200, ppm$)." This adaptation ensures our assessment aligns with the WHO's insights, offering a nuanced understanding of indoor air quality concerning CO concentrations.

Table 4 shows the comprehensive presentation (proposal) of the index sections of good, moderate, unhealthy for sensitive groups, bad, and hazardous levels via assessment items (PM2.5, CO2) of the indoor environment comprehensive IAQI index; the colors for the classifications are given in Table 5. After dividing the five classes equally, using a scale of 0 to 100 points, it was determined that the higher the score, the lower the IAQI, which considers comfort, such as temperature and humidity, and the safety of indoor harmful substances. The index is shown in blue when the indoor environment is good for the sensitive groups, green when the level is good for the general public, and yellow when the level could have a slight effect on the sensitive groups. The discomfort and health effects for sensitive groups are indicated in orange, and the level of discomfort and health effects for all occupants are marked in red

In order to accurately compute the IAQI index for CO, CO2, and PM2.5, it is essential to evaluate which pollutant exerts the greatest influence on indoor air quality. This assessment will guide the assignment of appropriate weights to each parameter in the calculation process.

The research presented in both [29] and [25] underscores the substantial impact of three key parameters—namely, PM2.5, CO, and CO2—on the overall assessment of indoor air quality. While acknowledging the significance of these variables, determining a definitive ranking among them poses a challenge due to their interdependence on diverse influencing factors. Despite the inherent complexities, we contend that there are discernible variations in terms of the severity of their effects. The ensuing table outlines our proposed estimations for the weights assigned to each pollutant within the specific context of this study:

Pollutants	Concentration levels (mg/m^3)	Organization	Estimate Weight
CO	10	WHO	0.3
CO2	225	WHO	0.2
PM2.5	0.45	USEPA	0.5

Table 3.3: IAQI guidelines for major indoor air pollutants over the period of 8hrs and estimated weight

Certainly, PM2.5, with its minute particle size, poses a considerable threat to human health. These fine particles can penetrate deep into the respiratory system, reaching the lungs and even entering the bloodstream. Prolonged exposure to elevated levels of PM2.5 is associated with various respiratory and cardiovascular issues, including aggravated asthma, bronchitis, and an increased risk of heart attacks. The detrimental impact on lung function and the cardiovascular system makes PM2.5 a critical parameter in assessing indoor air quality. Therefore, even a small concentration over the Time-Weighted Average (TWA - 8 hours) can contribute to adverse health effects, highlighting the importance of monitoring and controlling PM2.5 levels in indoor environments.

The subsequent noteworthy pollutant that plays a crucial role in indoor air quality calculations is CO. It demands only a specific duration of exposure over the Time-Weighted Average (TWA) to be classified as unhealthy, and it may indicate the presence of gas leakage, a prominent factor contributing to domestic fires. Concluding the trio of pollutants, CO_2. In this specific scenario, CO_2 would need concentrations over 22 times that of CO and 500 times that of PM2.5 to reach a similar level of negative impact on human health.

There are 2 main reasons for why this would be the case:

- People exhale carbon dioxide as a natural part of the respiratory process. In enclosed spaces with inadequate ventilation, the exhaled CO_2 can accumulate over time [25].
- The use of certain appliances that burn fossil fuels indoors, such as gas stoves or heaters, can release CO_2. Poorly ventilated combustion processes can contribute to elevated indoor CO_2 levels [22].

Thus, from the above data, we can form a calculation for the Average IAQI:

$$0.2 \cdot CO_{2\ \text{index}} + 0.3 \cdot CO_{\text{index}} + 0.5 \cdot PM_{2.5\ \text{index}} \tag{3.7}$$

So for determining the IAQI:

Ratings	Classification	Score	Color
A	Good	0-20	Blue
B	Moderate	21-40	Green
C	Unhealthy for sensitive groups	41-60	Yellow
D	Bad	61-80	Orange
E	Hazardous	81-100	Red

Table 3.4: Classification of IAQI index

Parameter	Weight
Temperature	0.28
Relative Humidity	0.16
Noise	0.19
Light intensity	0.22
Odour	0.15
Score	**1.00**

Table 3.6: Parameter weight of IAC

3.5.2 Indoor Air Comfort Index Calculation Criteria (IAC)

In order to calculate the comfort index required for the calculation of the IAQI composite index, the basic data on the comfort temperature and comfort humidity index using the data results of the indoor air measurement sensor were used. The total score of the indices was 100 points. [21]

The comfort index was divided into five grades considering the development of the comprehensive index of comfort and safety, which is shown in table 3.5.

Component	**Good**	**Moderate**	**Unhealthy for sensitive groups**	**Bad**	**Hazardous**
	A (0–20)	B (21–40)	C (41–60)	D (61–80)	E (81–100)
Temperature (°C)	24–26	22–24; 26–27	21–22; 27–29	20–21; 29–30	<20; >30
Humidity (%)	45–60	40–45; 60–65	35–40; 65–70	30–35; 70–85	<30; >85

Table 3.5: Parameter index of IAC based parameter interval

The study established the comfort index score range for temperature and humidity to align with both domestic and international environmental indices. A higher score indicates poorer air quality. The scoring system employs a scale from 0 to 100 points, where higher scores correspond to less comfortable environments. Blue indicates comfort for sensitive groups, green represents general public comfort, and yellow signals slight discomfort for sensitive groups. The index turns orange when the indoor environment is uncomfortable for sensitive groups and red when discomfort extends to all occupants.

The calculation of the Indoor Air Comfort (IAC) index involves the summation of individual parameter indices, each multiplied by its respective weight.. Nevertheless, in the scope of our study, factors like Noise & Light intensity can be deemed negligible, whereas Odor is treated as a pollutant in the earlier calculation of the Indoor Air Quality Index (IAQI). Thus, we can eliminate these 3 parameters leaving us with Temperature and Humidity giving us the function like so:

$$\omega_{Temperature} + \omega_{Humidity} = 1 \tag{3.8}$$

where:

$$\frac{\omega_{Temperature}}{\omega_{Humidity}} = \frac{0.28}{0.16} = 1.75 \tag{3.9}$$

so equation 3.8 can be written as:

$$1.75 \cdot \omega_{Humidity} + \omega_{Humidity} = 1 \tag{3.10}$$

thus the weight for Temperature and Humidity for our study would be:

$$\omega_{Humidity} = \frac{4}{11}; \omega_{Temperature} = \frac{7}{11} \tag{3.11}$$

The formula for the average indoor air comfort index (IAC) can be represented by the following:

$$\frac{7}{11} \cdot \text{Temperature}_{\text{index}} + \frac{4}{11} \cdot \text{Humidity}_{\text{index}} \tag{3.12}$$

Ratings	**Classification**	**Score**	**Color**
A	Very Comfortable	0-20	Blue
B	Comfortable	21-40	Green
C	Quite Comfortable	41-60	Yellow
D	Less Comfortable	61-80	Orange
E	Uncomfortable	81-100	Red

Table 3.7: Classification of IAC index

3.5.3 Development of an IAQI Composite Index

The proposed IAQI composite index, which combines the IAQI index and comfort index, was categorized into five grades, mirroring the classification used in each index for the proposed comfort index and IAQI index outlined in this study. Table 7 illustrates the classification using a scale ranging from 1 to 100 points. Higher scores were associated with lower quality IAQI and comfort indices, reflecting factors such as temperature, humidity, and safety.

The IAQI composite index for indoor gardening can be calculated by the following two factors [16]:

$$0.5 \cdot \text{IAQI}_{\text{index}} + 0.5 \cdot \text{IAC}_{\text{index}} \tag{3.13}$$

The IAQ index, which comprises three components, PM2.5, CO2, and CO, along with the comfort index consisting of temperature and humidity, is calculated on a scale of 100 points. The weights for the IAQI and the comfort index are set at 1:1 in determining the composite index.

Based on real-time and long-term data collection, we compared and analyzed the variations in PM2.5, CO2, temperature, humidity, and the newly incorporated CO levels following the installation of indoor gardens. The analysis results, grounded in epidemiological data, prevalence data, current maintenance practices, and recommendation criteria, are represented as indices. These indices aim to enhance the interpretability of the analyzed outcomes and offer insights into how current indoor air quality may impact occupant health. The indices include a comfort index reflecting temperature and humidity, an IAQI reflecting PM2.5, CO2, and the newly considered CO, and an composite IAQ index. The composite IAQ index is derived by amalgamating the comfort index and the IAQ index.

Ratings	Classification	Score	Color
A	Good	0-20	Blue
B	Moderate	21-40	Green
C	Unhealthy for sensitive groups	41-60	Yellow
D	Bad	61-80	Orange
E	Hazardous	81-100	Red

Table 3.8: Classification based on composite IAQI index

3.6 Classification Parameters

In this research, our focus is on addressing the challenge of monitoring indoor air quality while also emphasizing the need to illustrate, visualize, and predict future trends. Our primary interest lies in understanding how the parameters studied relate to the residents' actual health within the house. We aim to uncover trends within the collected data and provide predictions for future values.

To accomplish these objectives, we leverage machine learning techniques to convert the studied parameters into the Indoor Air Quality Index. This transformation enhances our comprehension of the current air quality within the client's house. Furthermore, our approach enables us to predict each parameter up to a one-hour interval into the future. This predictive capability allows for proactive measures to either prevent or enhance the air quality in their residence.

For this specific phase of the study, we opted to employ four classification models on the dataset. These models include Logistic Regression, Decision Tree Classifier, Random Forest Classifier, and KNeighbors Classifier. Each model undergoes a tuning process to ensure it provides the most accurate predictions without succumbing to over-fitting. The objective is to fine-tune the parameters of each model, optimizing their performance to achieve a balance between precision and generalization on the given dataset.

3.7 Data Visualization layer

3.7.1 ThingsBoard

In our data collection process from sensors, we utilized ThingsBoard, which is hosted on a compact server running the Linux operating system. This server, with a size roughly equivalent to that of an ATM card, acts as an intermediary for collecting data. For this project, the Raspberry Pi 4 model B was specifically chosen as the preferred hardware model. This decision was driven by its superior hardware specifications compared to its predecessors, making it highly suitable for tasks related to server functionality. The improved performance and capabilities of the Raspberry Pi 4 play a crucial role in ensuring the efficient operation of ThingsBoard, thereby guaranteeing reliable and responsive data collection from the interconnected sensors.

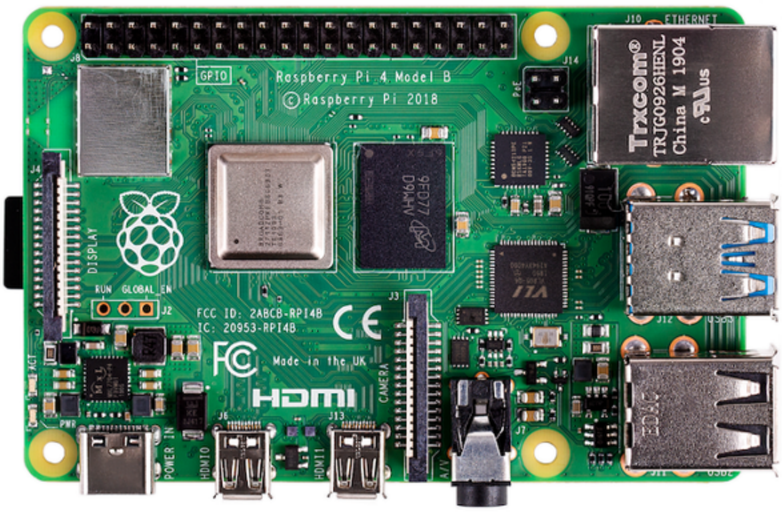

Figure 12: The Raspberry Pi 4 model B [1]

For the purpose of gathering data from sensors, we employed ThingsBoard, hosted on a compact server running the Linux operating system. This server, which is approximately the size of an ATM card, serves as an intermediary for data collection. The Raspberry Pi 4 model B was selected as the preferred model for this project. This choice was made due to its enhanced hardware specifications compared to its predecessors, rendering it particularly well-suited for tasks related to server functionality. The Raspberry Pi 4's improved performance and capabilities contribute to the efficient operation of ThingsBoard, ensuring reliable and responsive data collection from the connected sensors.

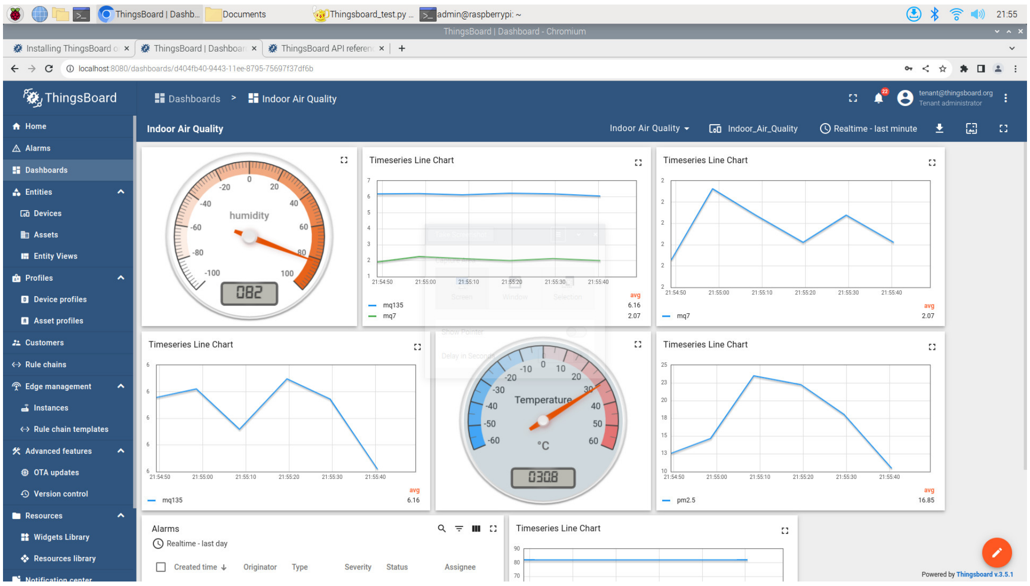

Figure 13: The main dashboard of ThingsBoard

3.7.2 Web Interface

To enhance user tracking of air quality changes, the air quality monitoring system, integrating fog technology, has implemented a web application structured on the Models-Views-Controllers (MVC) architecture. This architectural framework effectively separates the business logic layer from the display layer, contributing to improved application development and testing. MVC has gained widespread popularity due to its versatility in creating web and mobile applications. The system adheres to a Client-Server model, aligning with open-source principles to serve and contribute to the community. This approach allows users and developers to comprehend, reuse, evolve, and further develop its features.

The web interface's front-end component leverages ReactJS, a JavaScript-based framework. The user interface is designed with sections dedicated to displaying the latest data, location information, comprehensive data tracking, and predicting future air quality trends. The system issues alerts when obtained findings exceed safe thresholds, complemented by trend arrows based on recent sensor and data-derived information. While currently, the location section features only one operational monitoring station, there are plans for future expansion with the addition of more stations.

Real-time line charts within the web interface employ the GET method and HTTP protocol to fetch data from the ThingsBoard server API. This retrieval mechanism ensures accurate and interactive data visualization using the Highcharts library. The integration of these technologies allows users to have a dynamic and informed experience, staying abreast of air quality conditions and potential risks.

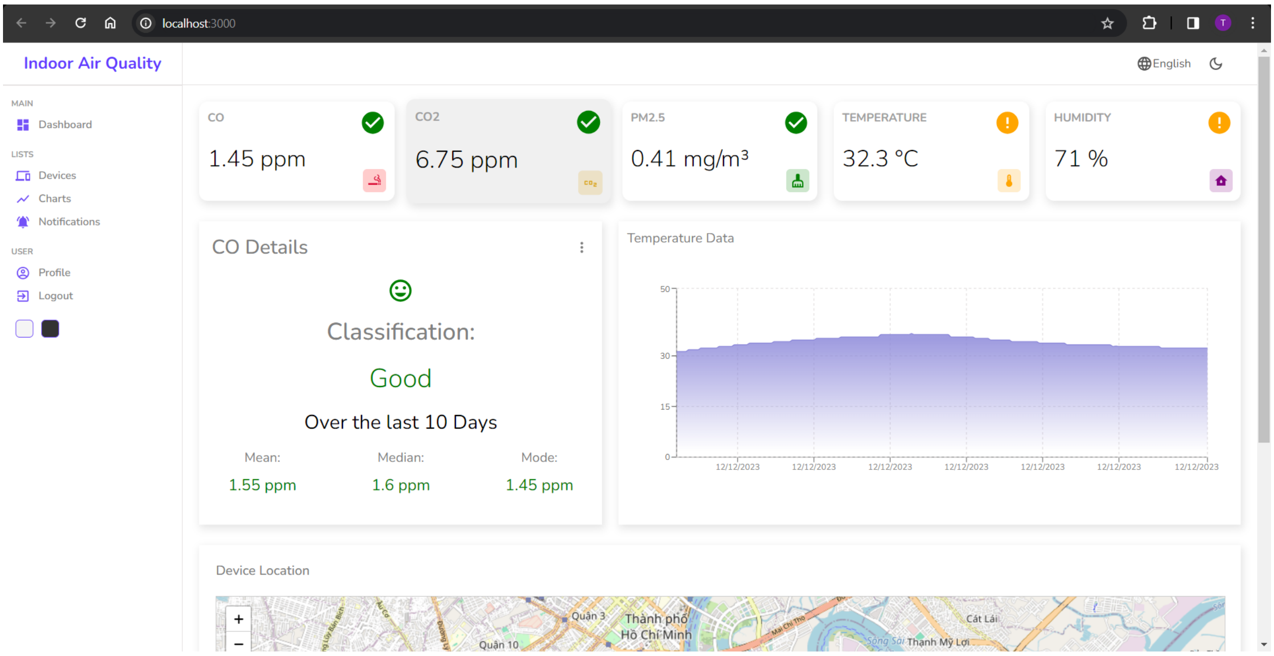

Figure 14: The main dashboard of the website

3.7.3 Mobile App

We have successfully created a mobile app that empowers users to easily assess and monitor air quality from any location. Leveraging the Flutter framework, developed by Google for mobile applications, we have engineered a high-performance application. This framework not only facilitates seamless deployment on multiple platforms such as the web, Android, iOS, and Windows desktop but also streamlines the development of cross-platform applications. Moreover, the app's back-end is managed by Firebase, with a specific focus on user authentication. This ensures that the mobile app remains dedicated to providing a user-friendly interface for conveniently displaying client house data. This holistic approach combines the versatility of Flutter with the robust back-end capabilities of Firebase to deliver a comprehensive and efficient mobile air quality monitoring solution.

In the development of our mobile app, our primary objective is to establish a user system that dynamically showcases user-specific parameters, leveraging the registered device API key. Through this personalized approach, users will gain access to a tailored interface reflecting the parameters relevant to their specific device.

The app will feature intuitive visualizations for each parameter, providing a clear and comprehensive overview of the data. These visualizations will be enhanced with the inclusion of respective threshold levels, allowing users to easily interpret and assess the current status of each parameter in comparison to predefined safety or optimal levels.

By incorporating this user-centric design, our mobile app aims to deliver a user-friendly experience, ensuring that individuals can effortlessly monitor and comprehend the air quality data associated with their registered device. This approach not only enhances user engagement but also contributes to a more informed and empowered user base in managing their indoor air quality.

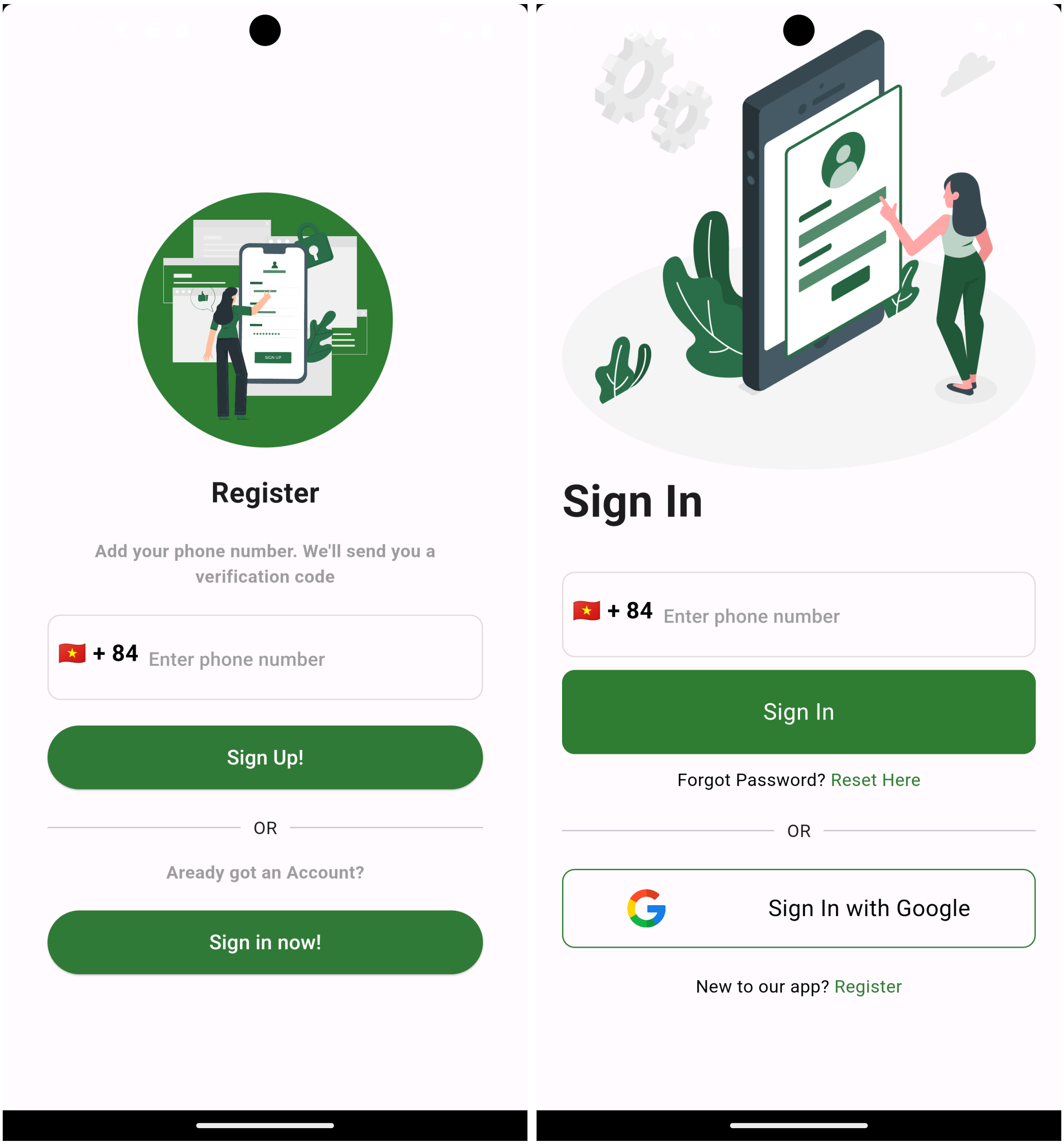

Figure 15: The Login and Register screen of the Mobile app

Chapter 4

Result of Implementations

4.1 Hardware Design

Our air quality monitoring system, employing fog computing in Ho Chi Minh City's District 7, logged data on temperature, humidity, gases, and PM 2.5 concentrations. The dataset is available on GitHub. Distrcit 7, justifying the installation of the system depicted in Figure 16. Deep learning models in this study will be trained using the dataset collected from these sensors. For the application of the deep learning models, specifically, 70% of the dataset was allocated as the training set, while the remaining 30% served as the test set.

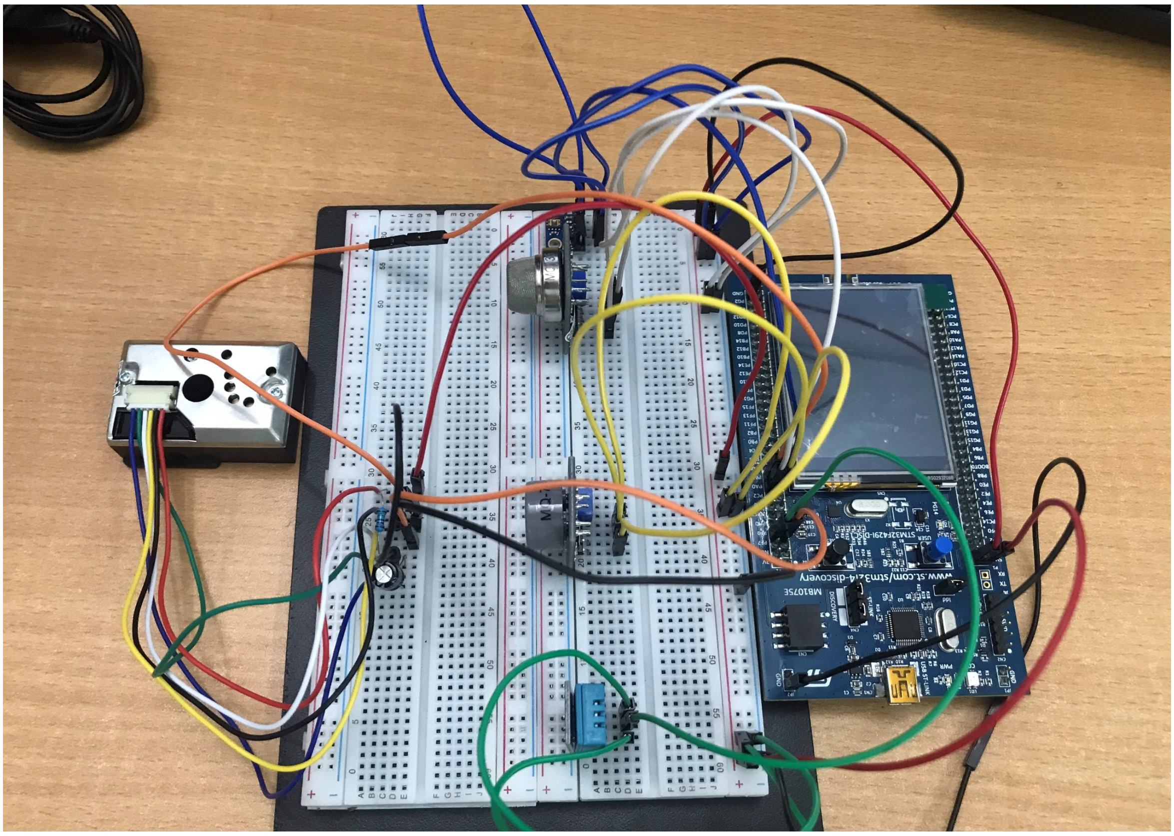

Figure 16: The hardware design

Additionally, rather than using a services database, we decided to invest in a Raspberry Pi to host the Thingsboard server. This will help us with database space issues, future maintenance, and will also give us complete control over how we want to configure our server to meet our needs.

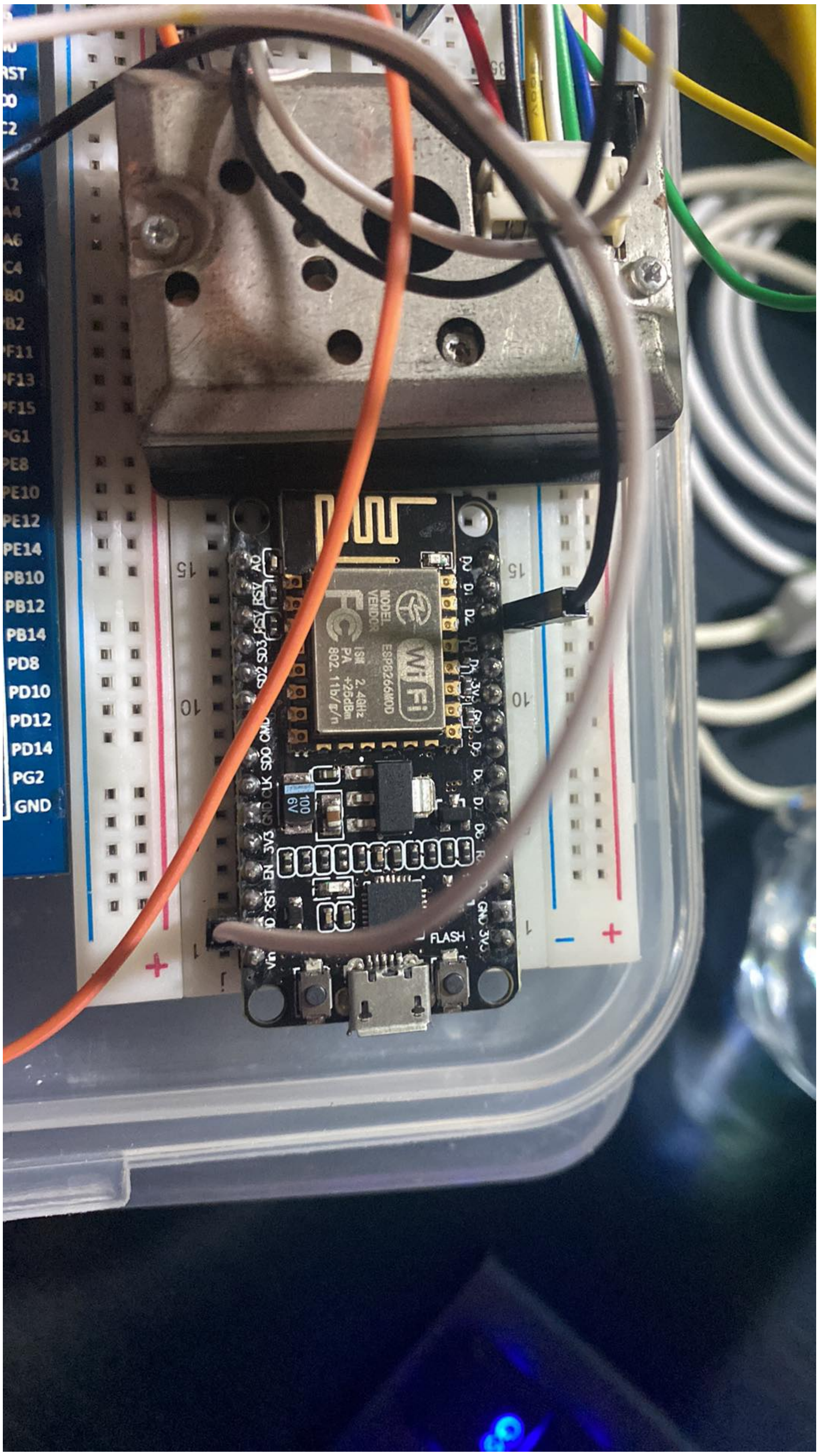

Figure 17: The NodeMCU

4.2 ThingsBoard

The interface displays visualization for each piece of data that was gathered via computerized timepieces and gauges. These data, which are used for data visualization and future weather predictions, are uploaded to Thingsboard using MQTT protocols.

The dashboard is made as shown below. System admin, tenant, and client are three separate views with different uses. However, the server's layout is the same; the information that is shown is only different. The system administrator has access to the tenant's data and can see and update it. Although they can customize the dashboard to their preferences and view other tenant account information, the client side and tenant side can only access the dashboard that depicts their individual data.

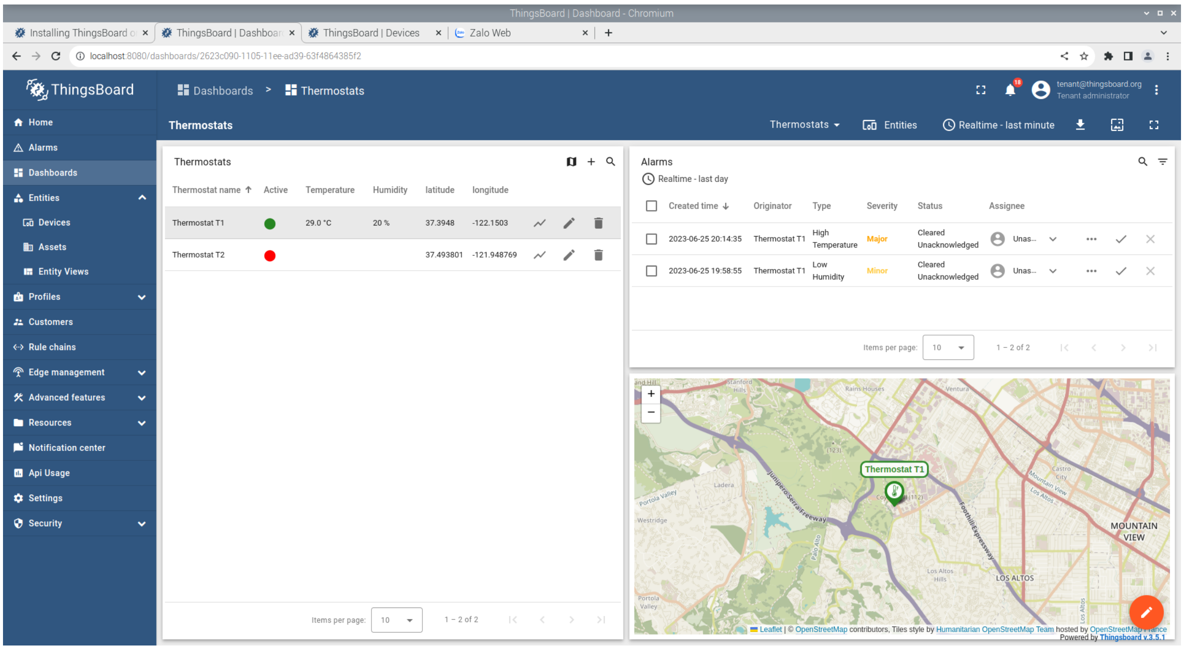

Figure 18: The main dashboard interface

The user account information is visibly presented in a designated column positioned on the left side of the page. This section incorporates three fundamental components, namely: the central display interface, serving as the primary platform for viewing the content of each module; the navigation menu, offering an organized and accessible pathway to various sections; and the main display interface, which serves as a focal point for presenting essential information and features related to the user account. This structured layout ensures a user-friendly and coherent experience by integrating key elements for seamless navigation and interaction.

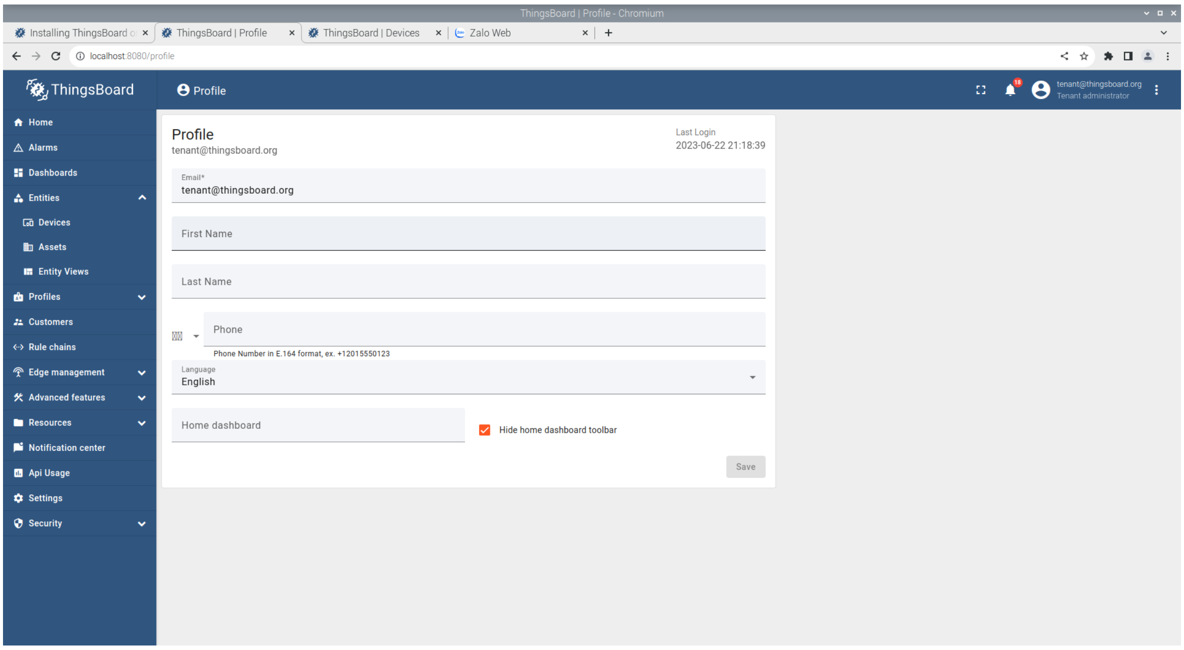

Figure 19: The account information

Tenants are likely to navigate through the dashboard tabs as they interact with the system, as this is where the data is presented. Initially, there will be a list of dashboards available, each tailored to the specific data acquired by sensors. For instance, there could be a dashboard dedicated to heat and humidity sensor data. To provide a more condensed view of all this information, an overview dashboard will be available, potentially linking to other dashboards for a more detailed perspective if needed. Additionally, for enhanced analysis, the same dashboard will feature a map indicating the system's location.

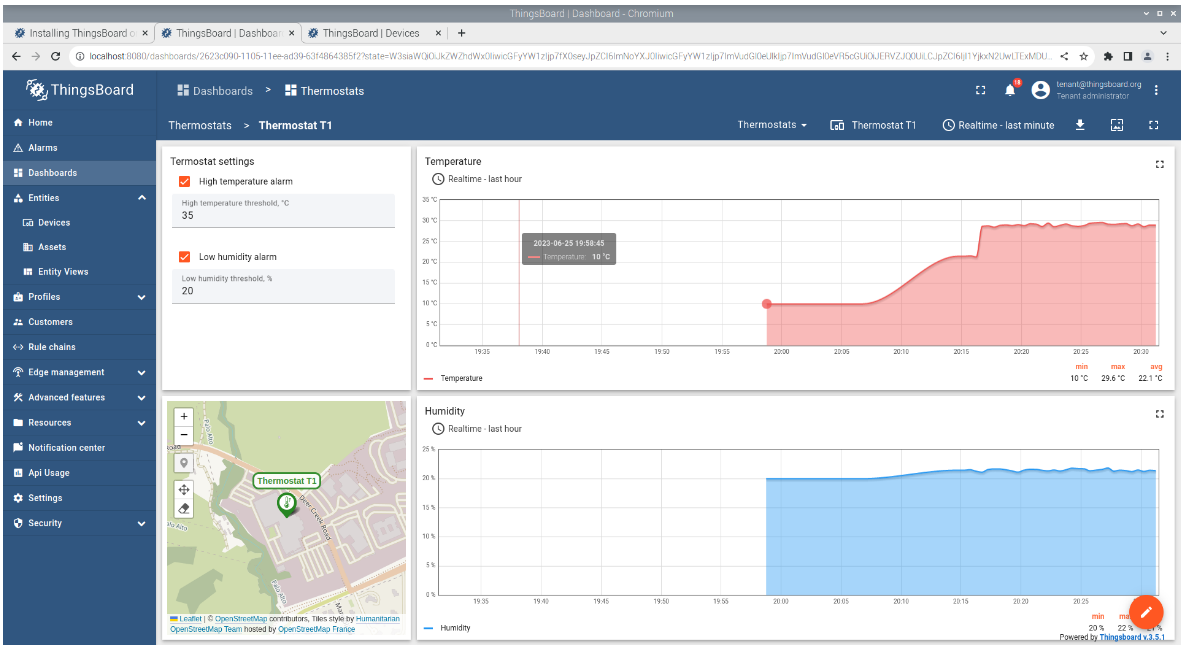

Figure 20: The dashboard for the thermostat

4.3 The Website

The following illustration represents the web interface, specifically designed to facilitate interaction between the administrator and the Indoor Air Quality (IAQ) monitoring system. This interface serves as a crucial point of engagement, allowing the administrator to effectively control and manage various aspects of the IAQ monitoring system. Through this intuitive web platform, the administrator can access and oversee the system's functionalities, ensuring seamless and efficient monitoring of indoor air quality parameters.

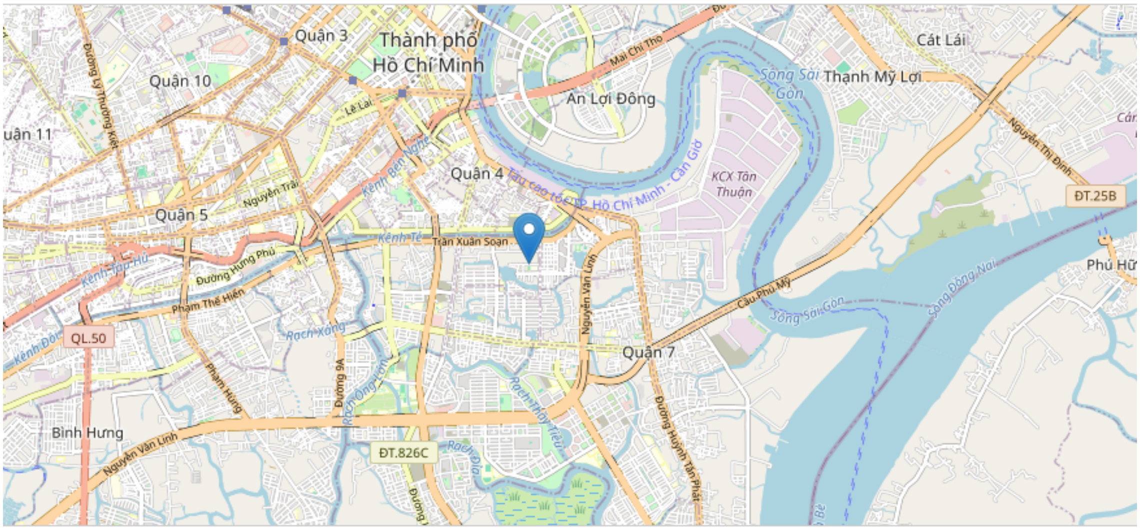

Figure 21: Map module using leaflet on the website

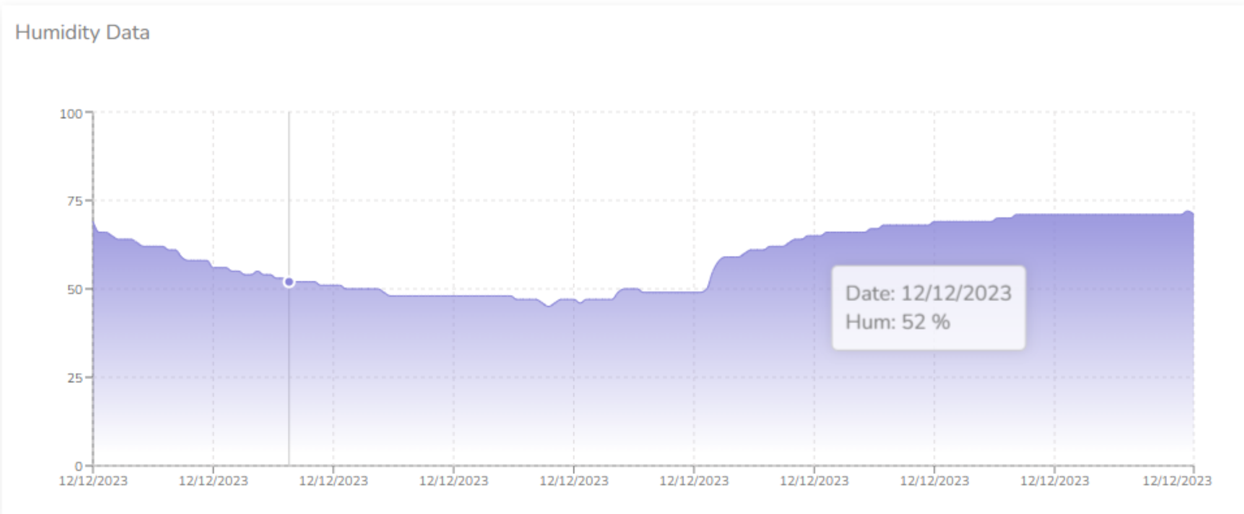

Figure 22: The chart to display the levels of each parameter

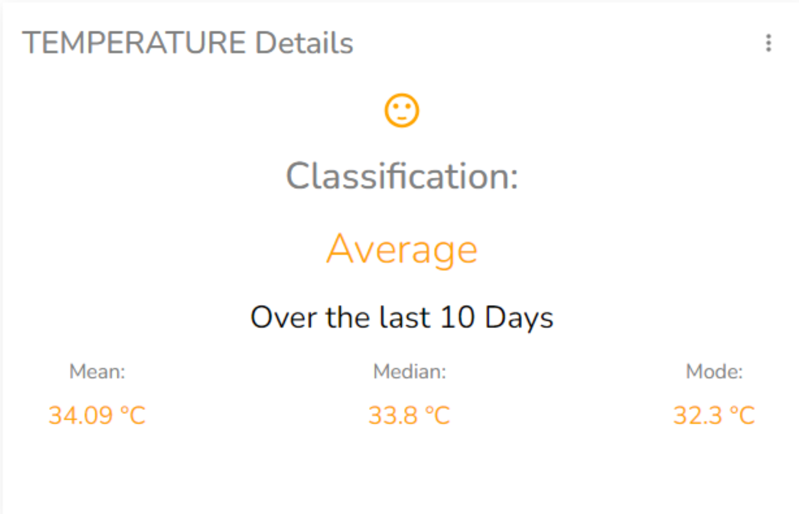

Figure 23: The Classification module

Regarding the dashboard, a dedicated website has been generated to facilitate a client in controlling and managing data entries. This website provides comprehensive insights into the submitted data, showcasing the timestamp of each entry, the levels of individual substances, and the corresponding location from which the data was collected.

Key features of the dashboard include:

- **Timestamp Display**: The website prominently showcases the timestamp for every data entry, providing a chronological view of when the information was submitted on the chart format. This feature empowers administrators to effectively track changes over time, facilitating a comprehensive understanding of the evolution of data entries.

- **Substance Levels**: The levels of each substance, including PM2.5, CO2, CO, temperature, and humidity, are presented in a clear and organized manner. This presentation format enables administrators to swiftly assess the air quality metrics, facilitating the prompt identification of any abnormal trends or patterns in the data.

- **Location Information**: The dashboard incorporates details about the location from which the data was collected. This geographical context enhances administrators' understanding by providing a spatial perspective on variations in air quality across different areas.

The website is designed to be user-friendly, empowering administrators with the tools needed to efficiently control and manage the air quality data. This centralized hub facilitates informed decision-making, allowing administrators to take timely actions based on the real-time information provided by the IoT system.

Figure 24: Dark theme of the website

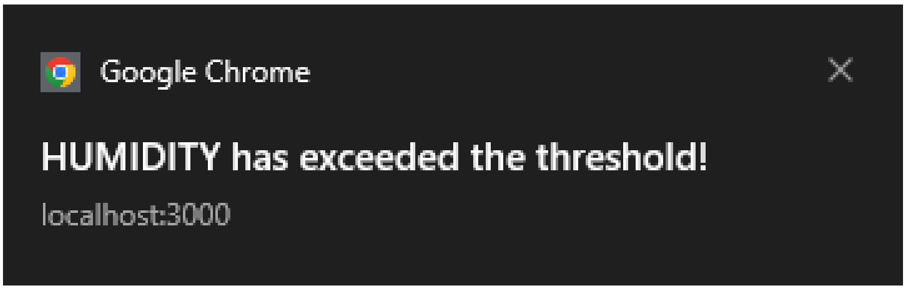

Figure 25: The notification

In addition to the aforementioned functionalities, we have incorporated additional features to enhance the interactive user experience of the website. This includes a notification system that alerts users when any parameters surpass a predefined threshold. Moreover, we have implemented a dark theme for the website, offering users an alternative visual setting for a more personalized and comfortable browsing experience.

all in all, we believe that all these implementations can help users keep track and have a better understanding on their indoor air quality. furthermore, there are still more improvements can be made like implementing the

4.4 The Mobile App

In the development of the mobile application, Firebase is employed for uncomplicated user authentication. Additionally, a straightforward API call to the server is implemented to retrieve the necessary data. The retrieved data is then seamlessly visualized as charts within the mobile app, providing users with an enhanced and intuitive understanding of the various parameters being monitored. This streamlined integration of Firebase and the server API contributes to a user-friendly experience, allowing for secure authentication and clear visualization of data charts for comprehensive parameter analysis.

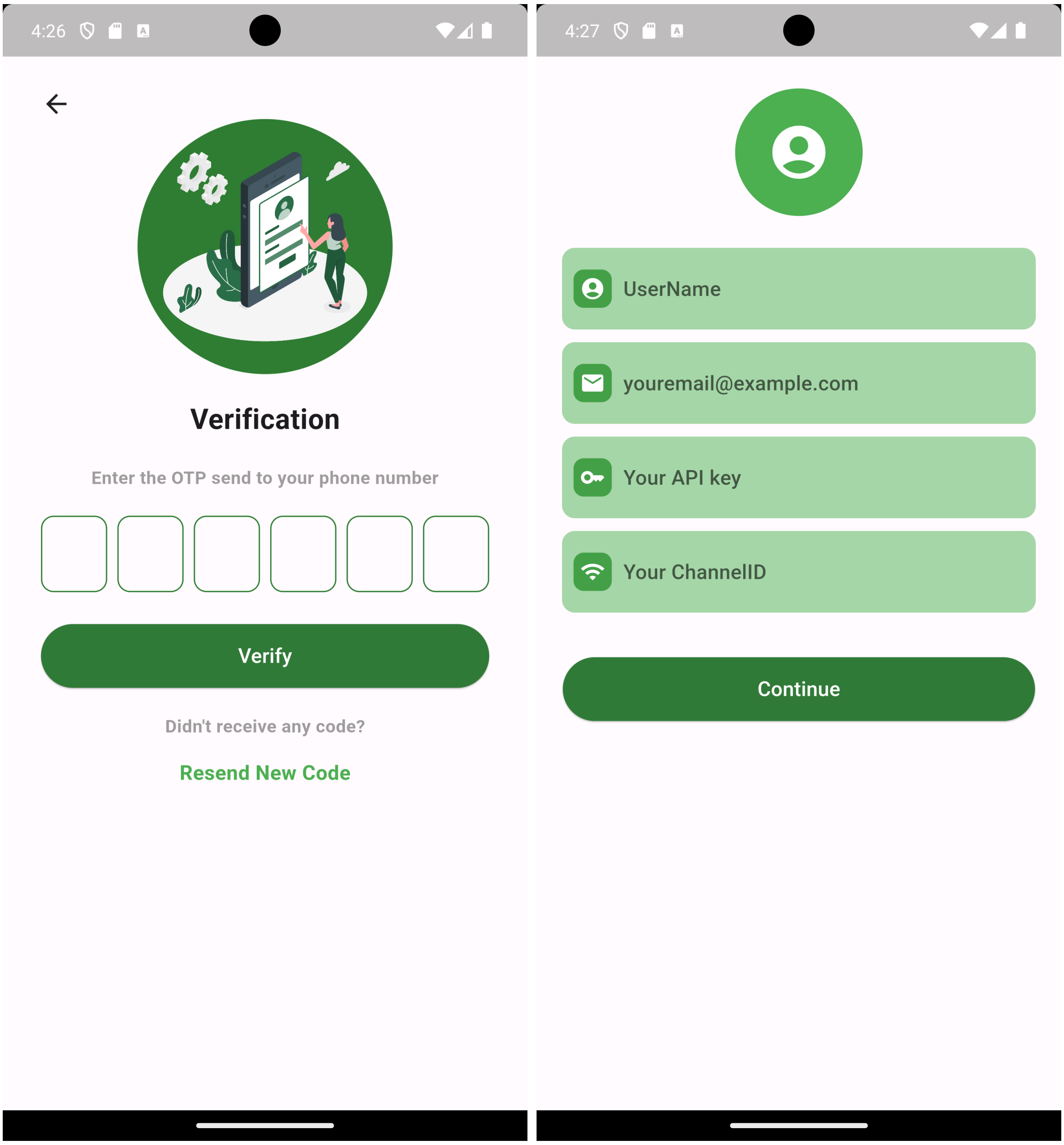

Figure 26: OTP authentication and user information input screen

In Figure 26, we have implemented a straightforward authentication process using a phone number and OTP (One-Time Password). Upon successful authentication of the OTP code, users are seamlessly redirected to the user information screen. At this juncture, users can input their personal information and their respective device API key for further engagement with the system.

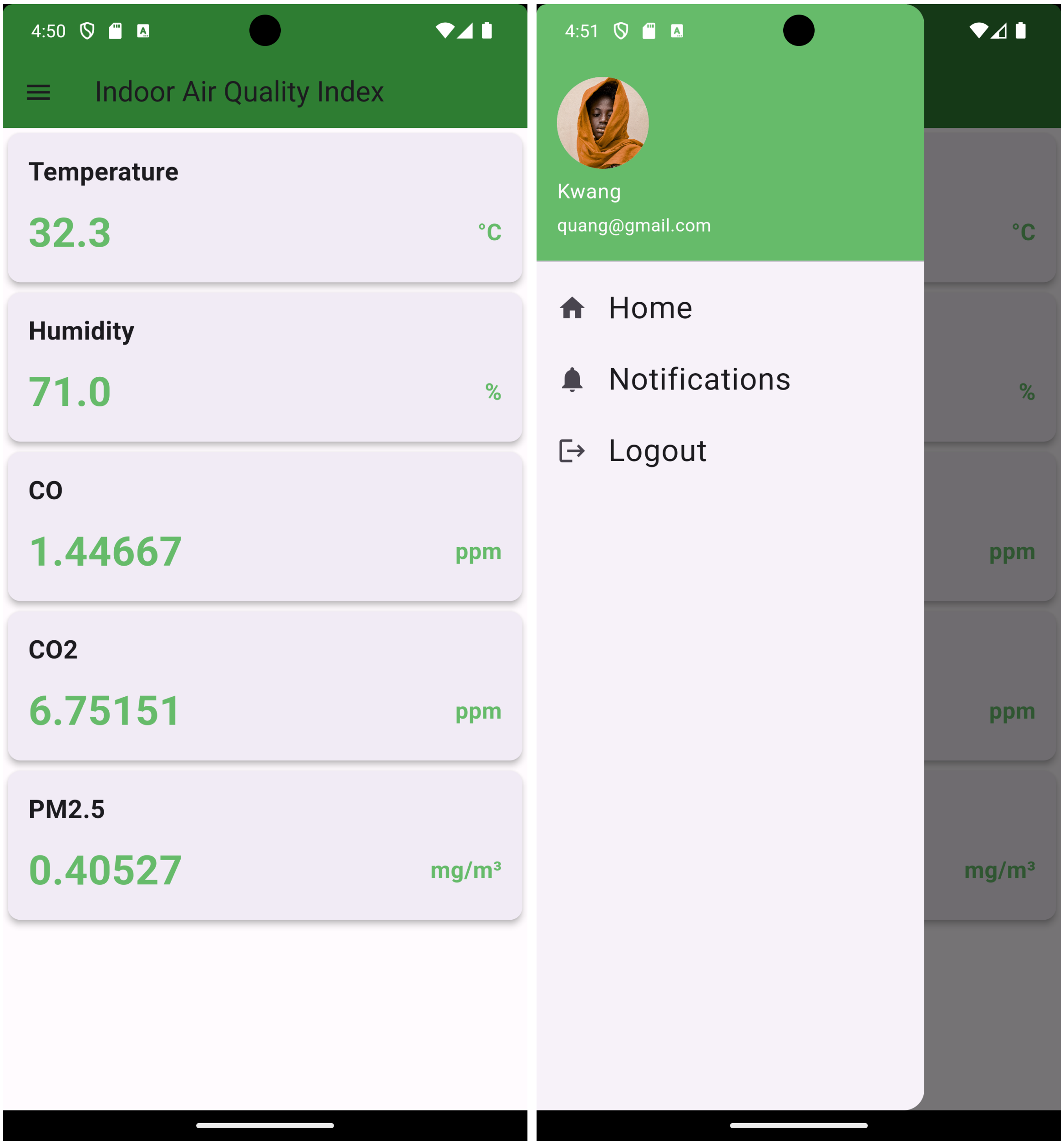

Figure 27: Main screen of the mobile app

Following this, users undergo a seamless transition to the main screen of the app, where they are presented with a prominent display showcasing the most recent values for each parameter. Furthermore, an intelligently integrated drawer feature is included, granting users convenient access to a range of functionalities. Within this drawer, users can explore and manage their account's personal information and utilize a logout function. This thoughtful and user-centric design significantly elevates the overall user experience, ensuring effortless navigation and swift access to essential features embedded within the mobile application.

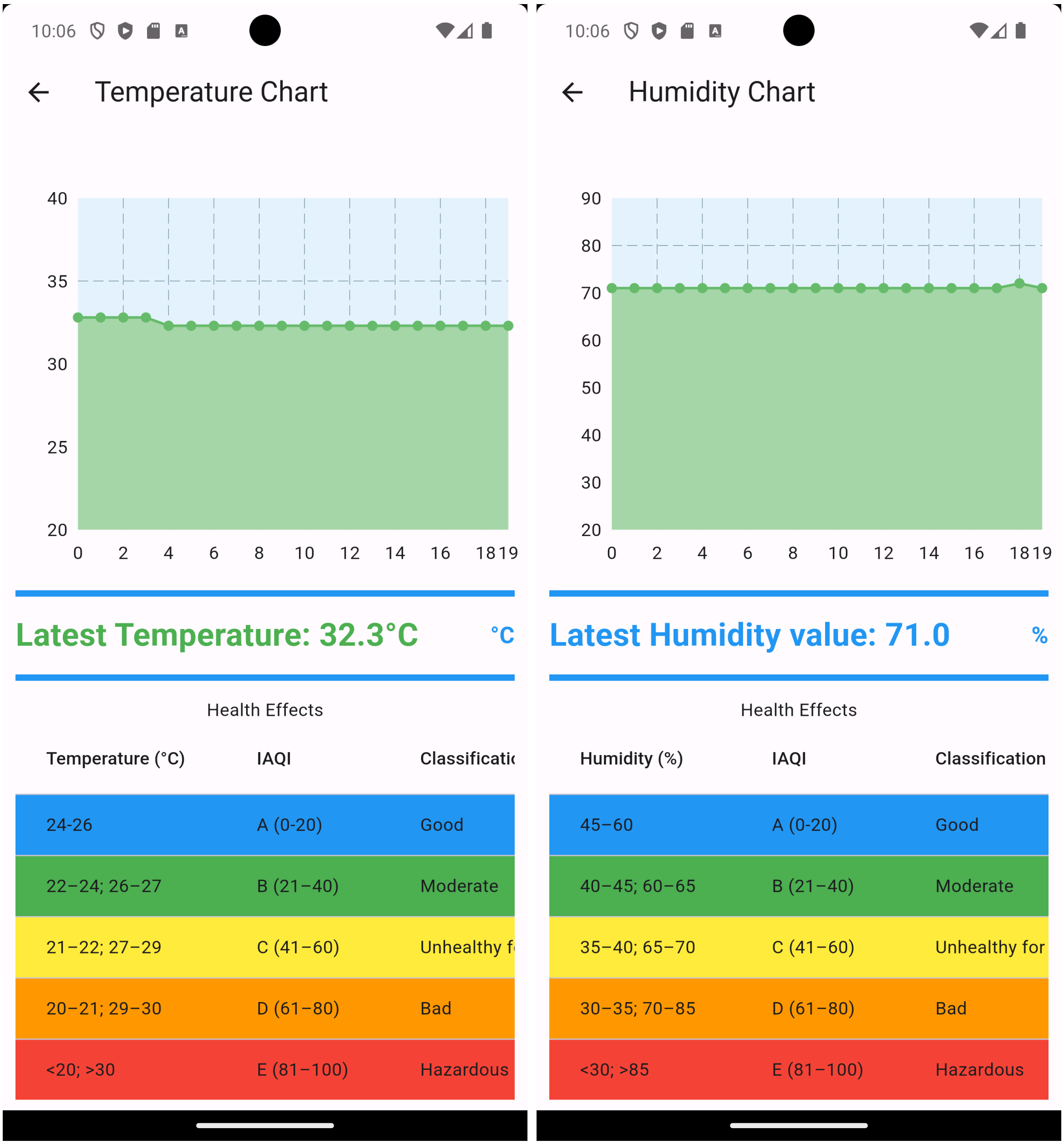

Figure 28: The data visualization of temperature and humidity

As illustrated in Figure 22, the screen features both a chart and a table, providing visual representations of the thresholds for Indoor Air Quality. The classification is presented in a spectrum from good to hazardous, with a color-coded scheme ranging from blue to red. This design enhances the user's ability to quickly interpret and assess the air quality levels based on the provided visual cues.

4.5 Classification Results

In our study focused on indoor air quality classification, we extensively utilized and evaluated four distinct classification models with the primary objective of identifying the most effective model to integrate into our website for optimal performance. Through a systematic comparison process, we rigorously assessed each model's performance based on various metrics including accuracy, kappa scores, and overall predictive capability. Notably, our findings consistently highlighted the Decision Tree Classifier and the Random

Forest Classifier as standout performers, consistently demonstrating the highest levels of accuracy and robustness across our testing scenarios. Consequently, based on these compelling outcomes, we have judiciously decided to leverage both the Decision Tree Classifier and the Random Forest Classifier for the core elements of our study. This strategic selection is further substantiated and visually elucidated in the accompanying figure, which graphically illustrates our model evaluation process and the compelling rationale behind our model choice for deployment within our web-based framework.

Model	**Train Accuracy**	**Test Accuracy**	**Kappa Score**
Logistic Regression	0.997772	0.997905	0.119764
Decision Tree Classifier	1.000000	0.999952	0.989223
Random Forest Classifier	1.000000	0.999905	0.978213
K-Neighbors Classifier	0.997748	0.997809	0.041577

Table 4.1: The comparison between models

In our comprehensive investigation into indoor air quality classification, we embarked on a detailed analysis employing a suite of four distinct classification models. The overarching goal was to discern the most effective model that could be seamlessly integrated into our website infrastructure to optimize its performance in real-world applications. To achieve this, we meticulously compared and evaluated the performance of each model across a range of key metrics, including accuracy, kappa scores, precision, recall, and F1-score, ensuring a thorough and multifaceted assessment of their capabilities.

Among the array of models tested, it became evident through rigorous testing and validation that both the Decision Tree Classifier and the Random Forest Classifier consistently emerged as top performers. These models demonstrated exceptional accuracy rates and robust predictive power across multiple testing scenarios, reaffirming their efficacy in the context of indoor air quality classification. Consequently, based on these compelling results and the need for reliable performance in practical settings, we strategically opted to leverage the combined strengths of the Decision Tree Classifier and the Random Forest Classifier for our study's core implementation.

The process of model selection was not arbitrary but rather informed by empirical evidence derived from our systematic evaluation, which is vividly represented in the accompanying figure. This graphical representation serves as a visual testament to the rigorous methodology applied in our study, illustrating the comparative performance of each model and highlighting the distinct advantages of the Decision Tree Classifier and the Random Forest Classifier. By leveraging these insights, we aim to optimize the accuracy and reliability of our indoor air quality classification system, ensuring its efficacy and usability within our web-based platform.

Furthermore, our decision to integrate these specific models into our study underscores a commitment to harnessing advanced machine learning techniques for real-world applications, particularly in the critical domain of indoor air quality monitoring and assessment. Moving forward, this strategic approach will facilitate the seamless deployment of our classification system, empowering users with actionable insights and contributing to advancements in air quality management and public health awareness.23.

Chapter 5

Discussion and Evaluation

5.1 Comparison

The existing IAQ monitoring systems on the market indeed fulfill a critical need, but they often face challenges related to database management, response time, and data visualization. These issues can be attributed to limitations in the database infrastructure and hardware compatibility, leading to gaps in data collection and system performance.

To address these shortcomings, our approach focuses on leveraging advanced hardware components to enhance data acquisition and system efficiency. Instead of relying solely on Arduino as the main microcontroller, we advocate for using the STM32 microcontroller along with the ESP8266 NodeMCU for WiFi connectivity. This strategic combination offers several advantages over traditional setups:

1. **Improved Performance and Response Time**: The STM32 microcontroller is more powerful and efficient compared to Arduino, offering greater processing capabilities and faster response times. This upgrade enhances the speed and reliability of data acquisition, enabling real-time monitoring of air quality parameters.

2. **Enhanced Connectivity and Compatibility**: By incorporating the ESP8266 NodeMCU for WiFi connectivity, our system ensures seamless integration with wireless networks. This not only facilitates data transmission to the cloud or central database but also supports remote monitoring and control functionalities.

3. **Reduced Maintenance Requirements**: The use of STM32 and ESP8266 NodeMCU not only improves performance but also contributes to lower maintenance demands over extended deployment periods. The robust hardware design minimizes the risk of system failures and reduces the need for frequent interventions or hardware replacements.

4. **Enhanced Customer Experience**: Ultimately, our approach aims to elevate the overall customer experience by delivering reliable, real-time data with minimal latency. The combination of powerful microcontrollers and efficient connectivity translates into a user-friendly interface and enhanced data visualization capabilities, empowering users to make informed decisions about indoor air quality.

In summary, our proposed hardware configuration represents a significant advancement in IAQ monitoring technology, addressing critical pain points related to system performance and usability. By harnessing the capabilities of STM32 and ESP8266 NodeMCU, we are poised to deliver a robust and efficient solution that sets new standards for data acquisition, processing, and visualization in the realm of indoor air quality monitoring.

The strategic choice of employing Raspberry Pi as the hosting device for our IoT system's database, specifically utilizing ThingsBoard, has indeed conferred substantial

Functions	Our Proposal	Arduair: Air Quality Monitoring [7]	Research on Indoor Air Monitoring System Based on STM32 [12]	Air Quality Monitoring and Forecasting System [28]
Dashboard	x	x		
Low cost	x	x		x
IoT Platform	x			x
Performance	x	x	x	x
Publication	x			
Deep learning model	x	x	x	x

Table 5.1: Comparison between models

benefits and flexibility to our architecture. By integrating Raspberry Pi, a versatile and cost-effective single-board computer, as the backbone of our database infrastructure, we have effectively empowered our system with enhanced control and customization capabilities.

One of the key advantages of leveraging Raspberry Pi for hosting ThingsBoard is the ability to tailor the database environment to suit our specific requirements. Unlike off-the-shelf solutions or cloud-based platforms that may have limited customization options, hosting ThingsBoard on Raspberry Pi enables us to optimize database settings, configurations, and performance parameters to align precisely with our IoT system's needs. This level of control fosters greater efficiency and responsiveness in data management, ensuring reliable operation and real-time responsiveness.

Furthermore, by utilizing Raspberry Pi, we have achieved improved connectivity and integration within our IoT ecosystem. Raspberry Pi's compatibility with a wide range of IoT devices and protocols facilitates seamless communication and data exchange between connected sensors, actuators, and the ThingsBoard database. This interoperability enhances the overall robustness and reliability of our system, enabling comprehensive data collection, analysis, and visualization.

Another noteworthy benefit of Raspberry Pi as the hosting device is its cost-effectiveness and accessibility. Raspberry Pi offers a cost-efficient alternative to traditional server infrastructure, making it an ideal choice for research and development projects with budget constraints. Moreover, its compact form factor and low power consumption make it suitable for deployment in diverse environments, including resource-constrained settings where space and energy efficiency are paramount considerations.

In summary, our decision to utilize Raspberry Pi for hosting the ThingsBoard database underscores our commitment to building a robust, scalable, and adaptable IoT system architecture. This strategic investment in database infrastructure not only optimizes system performance but also lays a solid foundation for future expansions and enhancements. Moving forward, we remain dedicated to refining and leveraging this architecture to unlock new possibilities in IoT-enabled environmental monitoring and beyond.

5.2 Evaluation

Expanding on the topic of IoT-enabled air quality monitoring systems, it becomes evident that these innovations have far-reaching implications for environmental sustainability and public health. The successful integration of IoT devices, coupled with fog computing, exemplifies a paradigm shift in real-time data collection and analysis, paving the way for a cleaner and healthier future.

The IoT-based Real-time Air Quality Monitoring System described in this study represents a significant leap forward in environmental technology. By leveraging inexpensive yet high-quality sensors and cloud-based platforms like ThingsBoard, this system exemplifies how IoT can democratize access to critical environmental data. Notably, the user-friendly interface with interactive charts provides stakeholders with actionable insights into air quality trends and potential health risks.

One notable aspect of this project is the emphasis on predictive capabilities through the integration of machine learning algorithms. Unlike conventional monitoring systems that simply report current air quality, this IoT system can anticipate trends and issue alerts for poor air quality conditions. This proactive approach empowers individuals and communities to take timely actions to mitigate health risks associated with air pollution.

Moreover, the emphasis on open-source development and community collaboration is commendable. By making the project's code accessible and encouraging replication and improvement, the research team fosters innovation and knowledge-sharing within the developer community. This open approach not only accelerates technological advancements but also promotes transparency and accountability in environmental monitoring initiatives.

Looking ahead, the project's roadmap highlights several areas for further enhancement. Refining machine learning algorithms to improve accuracy and performance, as well as optimizing communication protocols to reduce implementation costs, are critical objectives for future iterations. These advancements will not only strengthen the system's capabilities but also enhance its scalability and applicability across diverse environmental monitoring scenarios.

In conclusion, the IoT-enabled Real-time Air Quality Monitoring System with Fog Computing represents a cornerstone in modern environmental technology. Its success underscores the transformative potential of IoT and machine learning in addressing complex challenges such as air pollution. As the project continues to evolve and inspire further innovations, it holds the promise of creating a more sustainable and healthier world for generations to come.

Chapter 6

Conclusion

In conclusion, the advent of the Internet of Things (IoT) has instigated a profound revolution in the realm of indoor air quality monitoring, ushering in a paradigm shift marked by enhanced accessibility, affordability, and precision attributable to the integration of IoT-enabled devices. This research introduces a groundbreaking IoT-based Real-time Indoor Air Quality Monitoring System seamlessly incorporating Fog Computing. This system exhibits the remarkable capability of autonomously classifying the Indoor Air Quality Index based on meticulously collected parameters. The intricately designed web interface, characterized by user-friendliness, unfolds a comprehensive panorama featuring recent updates, historical data, anticipated values, and interactive charts. This surpasses conventional methods, setting new standards for effectiveness in indoor air quality monitoring.

The system's seamless integration with state-of-the-art IoT technologies not only facilitates remote monitoring but also expedites data transfer, thus significantly elevating the efficacy of real-time data collection within indoor environments. The device's impressive accuracy, achieved through the strategic use of cost-effective yet superior sensors, instills confidence in users who can rely on it to anticipate potential health risks. Beyond offering a replicable procedure for enhancing indoor air quality, this research serves as a catalyst for ongoing improvements, advocating for the continual betterment of indoor environments through the implementation of advanced technologies.

It is important to note that the primary objective of this study is to address existing challenges encountered by conventional air quality systems, with a specific focus on enhancing the cloud database and machine learning algorithms. The research is geared towards improving machine learning algorithms and connection protocols within the project design. In terms of machine learning, additional testing across various models will be conducted to identify a more reliable approach, including the exploration of additional layers to enhance accuracy. For connection protocols, further deployment and testing are necessary to establish a more robust and rapid device connection. These advancements signify the potential of this study as a technology that can significantly benefit individuals and communities alike.

Bibliography

[1] `https://www.raspberrypi.com/products/raspberry-pi-4-model-b/`. [Accessed 29-04-2024].

[2] *[Thông tin] k thut NodeMCU ESP8266 — mecsu.vn.* `https://mecsu.vn/ho-tro-ky-thuat/thong-tin-ky-thuat-nodemcu-esp8266.R1q`. [Accessed 29-04-2024].

[3] Dlnya Abdulahad Aziz. "Webserver based smart monitoring system using ESP8266 node MCU module". In: *International Journal of Scientific & Engineering Research* 9.6 (2018), pp. 801–808.

[4] Jessie P Buckley, Jonathan M Samet, and David B Richardson. "Commentary: Does air pollution confound studies of temperature?" In: *Epidemiology* 25.2 (2014), pp. 242–245.

[5] Mauro Castelli et al. "A machine learning approach to predict air quality in California". In: *Complexity* 2020 (2020).

[6] Somphop Chanthakit and Choopan Rattanapoka. "Mqtt based air quality monitoring system using node MCU and node-red". In: *2018 Seventh ICT International Student Project Conference (ICT-ISPC)*. IEEE. 2018, pp. 1–5.

[7] Vikhyat Chaudhry. "Arduair: air quality monitoring". In: *International Journal of Environmental Engineering and Management* 4.6 (2013), pp. 639–646.

[8] Stijn Dhondt et al. "Environmental health impacts of mobility and transport in Hai Phong, Vietnam". In: *Stochastic Environmental Research and Risk Assessment* 25 (2011), pp. 363–376.

[9] Vitali Fioletov, James B Kerr, and Angus Fergusson. "The UV index: definition, distribution and factors affecting it". In: *Canadian journal of public health* 101 (2010), pp. I5–I9.

[10] Kalpesh Gupta, G Gokul Krishna, and T Anjali. "An IoT based system for domestic air quality monitoring and cooking gas leak detection for a safer home". In: *2020 International Conference on Communication and Signal Processing (ICCSP)*. IEEE. 2020, pp. 0705–0709.

[11] Sarath Guttikunda and Puja Jawahar. "Can we vacuum our air pollution problem using smog towers?" In: *Atmosphere* 11.9 (2020), p. 922.

[12] Zewei He and Xuxian Ruan. "Research on indoor air monitoring system based on STM32". In: *Academic Journal of Engineering and Technology Science* 5.7 (2022), pp. 46–52.

[13] Matthew Henschke, Xinzhou Wei, and Xiaowen Zhang. "Data visualization for wireless sensor networks using ThingsBoard". In: *2020 29th Wireless and Optical Communications Conference (WOCC)*. IEEE. 2020, pp. 1–6.

[14] J Horner. "Carbon monoxide: the invisible killer". In: *The journal of the Royal Society for the Promotion of Health* 118.3 (1998), pp. 141–145.

[15] JunHo Jo et al. "Development of an IoT-based indoor air quality monitoring platform". In: *Journal of Sensors* 2020 (2020), pp. 1–14.

[16] Ho-Hyun Kim et al. "Evaluation of IAQ Management Using an IoT-Based Indoor Garden". In: *International Journal of Environmental Research and Public Health* 17 (Mar. 2020), p. 1867. DOI: 10.3390/ijerph17061867.

[17] Darkwah Williams Kweku et al. "Greenhouse effect: greenhouse gases and their impact on global warming". In: *Journal of Scientific research and reports* 17.6 (2018), pp. 1–9.

[18] Sean McGrath et al. "IoT personal air quality monitor". In: *2020 31st Irish signals and systems conference (ISSC)*. IEEE. 2020, pp. 1–4.

[19] Helton Pierre Lucena de Medeiros and Gustavo Girão. "An iot-based air quality monitoring platform". In: *2020 IEEE international smart cities conference (ISC2)*. IEEE. 2020, pp. 1–6.

[20] Yogendra Singh Parihar et al. "Internet of things and nodemcu". In: *journal of emerging technologies and innovative research* 6.6 (2019), p. 1085.

[21] N Pasaribu and AS Yuwono. "Formulation of indoor air comfort index for office building [A case study in sub-district office]". In: *IOP Conference Series: Earth and Environmental Science*. Vol. 871. 1. IOP Publishing. 2021, p. 012030.

[22] Rogelio Perez Padilla, Astrid Schilmann, and Horacio Riojas-Rodriguez. "Respiratory health effects of indoor air pollution". In: *The international journal of tuberculosis and lung disease : the official journal of the International Union against Tuberculosis and Lung Disease* 14 (Sept. 2010), pp. 1079–86.

[23] Kinnera Bharath Kumar Sai, Somula Rama Subbareddy, and Ashish Kumar Luhach. "IOT based air quality monitoring system using MQ135 and MQ7 with machine learning analysis". In: *Scalable Computing: Practice and Experience* 20.4 (2019), pp. 599–606.

[24] Jagriti Saini, Maitreyee Dutta, and Goncalo Marques. "Indoor air quality prediction using optimizers: A comparative study". In: *Journal of Intelligent & Fuzzy Systems* 39.5 (2020), pp. 7053–7069.

[25] Luigi Schibuola and Chiara Tambani. "Indoor environmental quality classification of school environments by monitoring PM and CO2 concentration levels". In: *Atmospheric Pollution Research* 11 (Feb. 2020), pp. 332–342. DOI: 10.1016/j.apr.2019.11.006.

[26] Thomas Lee Scott and Amna Eleyan. "CoAP based IoT data transfer from a Raspberry Pi to Cloud". In: *2019 International Symposium on Networks, Computers and Communications (ISNCC)*. IEEE. 2019, pp. 1–6.

[27] V Suma, Ramya R Shekar, and Kumar A Akshay. "Gas leakage detection based on IOT". In: *2019 3rd International conference on Electronics, Communication and Aerospace Technology (ICECA)*. IEEE. 2019, pp. 1312–1315.

[28] Quynh Tran et al. "Air Quality Monitoring and Forecasting System using IoT and Machine Learning Techniques". In: July 2022. DOI: 10.1109/GTSD54989.2022.9988756.

[29] Vinh Van Tran, Duckshin Park, and Young-Chul Lee. “Indoor Air Pollution, Related Human Diseases, and Recent Trends in the Control and Improvement of Indoor Air Quality”. In: *International Journal of Environmental Research and Public Health* 17 (Apr. 2020), p. 2927. DOI: 10.3390/ijerph17082927. URL: https://www.mdpi.com/1660-4601/17/8/2927.

[30] Tuyen Phong Truong, Duy Thanh Nguyen, and Phong Vu Truong. “Design and deployment of an IoT-based air quality monitoring system”. In: *International Journal of Environmental Science and Development* 12.5 (2021), pp. 139–145.

[31] Wenjuan Wei et al. “Machine learning and statistical models for predicting indoor air quality”. In: *Indoor Air* 29.5 (2019), pp. 704–726.

[32] Yu-Fei Xing et al. “The impact of PM2. 5 on the human respiratory system”. In: *Journal of thoracic disease* 8.1 (2016), E69.

[33] Hong Zheng et al. “A multiple kernel learning approach for air quality prediction”. In: *Advances in Meteorology* 2018 (2018).

[34] Dixian Zhu et al. “A machine learning approach for air quality prediction: Model regularization and optimization”. In: *Big data and cognitive computing* 2.1 (2018), p. 5.