B.M.S. COLLEGE OF ENGINEERING

(Autonomous Institute, Affiliated to VTU)
Bull Temple Road, Basavanagudi, Bengaluru - 560019

A Capstone Project Report on

"Air Pollution Monitoring, Forecasting and Controlling System"

Submitted in partial fulfilment of the requirements for the award of degree

BACHELOR OF ENGINEERING

IN

INFORMATION SCIENCE AND ENGINEERING

By

Prateek M Gummaraju - 1BM19IS117

Ruchi Agarwal - 1BM19IS133

Samartha S - 1BM19IS219

Under the guidance of

Prof. Pallavi B.
Assistant Professor

Department of Information Science and Engineering
2022-2023

B.M.S. COLLEGE OF ENGINEERING

(Autonomous Institute, Affiliated to VTU)
Bull Temple Road, Basavanagudi,
Bengaluru – 560019

Department of Information Science and Engineering

C E R T I F I C A T E

This is to certify that the project entitled "**Air Pollution Monitoring, Forecasting and Controlling System**" is a bona-fide work carried out by **Prateek M Gummaraju (1BM19IS117), Ruchi Agarwal (1BM19IS133)** and **Samartha S (1BM19IS219)** in partial fulfilment for the award of degree of Bachelor of Engineering in **Information Science and Engineering** from **Visvesvaraya Technological University, Belgaum** during the year **2022-2023**. It is certified that all corrections/suggestions indicated for Internal Assessments have been incorporated in the report deposited in the departmental library. The project report has been approved as it satisfies the academic requirements in respect of project work prescribed for the Bachelor of Engineering Degree.

Prof. Pallavi B **Assistant Professor**	**Dr. Jayarekha P** **Professor and HOD**	**Dr. S. Muralidhara** **Principal**

Examiners

Name of the Examiner	**Signature of the Examiner**
1.	
2.	

DECLARATION

Prateek M Gummaraju (1BM19IS117), Ruchi Agarwal (1BM19IS133) and **Samartha S (1BM19IS219)** students of B.E. Information Science and Engineering, B.M.S. College of Engineering, Bangalore - 19, hereby declare that the capstone project entitled **"Air Pollution Monitoring, Forecasting and Controlling System"** is an authentic work carried out under the supervision and guidance of **Prof. Pallavi B, Assistant Professor**, Department of Information Science and Engineering, B.M.S. College of Engineering, Bangalore. We have not submitted the matter embodied to any other university or institution for the award of any other degree.

Place:

Date:

Name	USN	Signature
Prateek M Gummaraju	1BM19IS117	
Ruchi Agarwal	1BM19IS133	
Samartha S	1BM19IS219	

ACKNOWLEDGEMENT

The satisfaction that accompanies the successful completion of this Capstone Project Phase-2 would be incomplete without the mention of the people who made it possible through constant guidance and encouragement.

We would take this opportunity to express our heart-felt gratitude to **Dr. B. S. Ragini Narayan**, Chairperson, Donor Trustee, Member Secretary & Chairperson, BMSET, **Dr. P. Dayananda Pai**, Member Life Trustee, BMSET and **Dr. S. Muralidhara**, Principal, B.M.S. College of Engineering for providing the necessary infrastructure to complete this Capstone Project Phase-2.

We wish to express our deepest gratitude and thanks to **Dr. Jayarekha P**, Head of the Department, Information Science and Engineering and the Project Coordinators **Dr. Nalini M K** and **Prof. Harini S** for their constant support.

We wish to express sincere thanks to our guide **Prof. Pallavi B**, Assistant Professor, Department of Information Science and Engineering for helping us throughout and guiding us from time to time.

A warm thanks to all the faculty of the Department of Information Science and Engineering, who have helped us with their views and encouraging ideas.

Prateek M Gummaraju (1BM19IS117)

Ruchi Aggarwal (1BM19IS133)

Samartha S (1BM19IS219)

ABSTRACT

Air pollution is the presence of compounds in the atmosphere that are hazardous to human beings and the health of other living organisms, or that can impair climate and materials. Chemical compounds such as carbon monoxide, ozone, nitrogen dioxide etc, are common air pollutants and they lower the quality of air.

In many industrial and urban areas today, maintaining and monitoring air quality has become a top priority. Numerous elements, such as time, location, and uncertain variables have an impact on air quality. Due to the rising levels of air pollution, there is a great need to implement effective air quality monitoring systems that gather data on the concentration of various air pollutants and provide not only the current assessments of the level of pollution, but also predict the level of pollution in the near future. In addition to this, we find it essential to compare the pollution levels with the recommended limit given by organizations such as the World Health Organisation (WHO) and consequently provide measures to prevent the same.

With this project we aim to successfully demonstrate the possibility of a low cost, IOT based system built using Arduino and various sensors that not only monitors the current air quality, but also has the ability to forecast the future predictions and also provide suggestions on how to control the pollution and protect human beings from the harmful pollutants.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION

1.1 Overview

Air pollution is defined as the presence of compounds in the atmosphere that are hazardous to human beings and the health of other living organisms, or that can impair climate and materials. These compounds can come from various sources, including industrial and agricultural emissions, transportation, and natural sources such as wildfires. The effects of air pollution can be devastating to human health, as well as to the environment. Chemical compounds such as carbon monoxide, ozone, nitrogen dioxide etc, are common air pollutants and they lower the quality of air. The emissions from automobiles, trucks, and other vehicles as well as the pollutants from factories and power plants all contribute to air pollution. A further factor in air pollution is agriculture, notably the use of fertilisers and pesticides. Pollutants are also released into the atmosphere by natural processes like wildfires and volcanic eruptions. The health of both humans and animals can be negatively impacted by air pollution. Respiratory conditions like asthma and bronchitis can be brought on by the minute particles in contaminated air that can enter the lungs. Heart disease and lung cancer risk can grow with prolonged exposure to contaminated air. Additionally, because it damages crops and impairs sight, air pollution can have a negative impact on the ecosystem.

One of the most important environmental problems affecting our planet right now is air pollution. We must act to contain it because it poses a major risk to both human health and the environment. In order to safeguard human health, air pollution must be controlled. Numerous health conditions, including cancer, heart disease, and respiratory disorders, can be brought on by the pollutants in the air we breathe. We can prevent these and other health issues from occurring by reducing air pollution, which would save lives and enhance quality of life for millions of people. Climate change is one of the biggest environmental problems of our day, and air pollution can contribute to it. Climate change in turn has a negative effect on the ecosystem and biodiversity of our planet. Controlling air pollution is crucial for financial reasons as well. Buildings and infrastructure can suffer harm from air pollution, which can be expensive to restore. Additionally, it can lower visibility, which is bad for industries like tourism that depend on the beauty of the natural world. Air pollution can also raise health care expenses since those who breathe it are more likely to require medical attention. By reducing air pollution, we can lower these expenses and encourage economic growth. Based on this, we can say that controlling air pollution is very important to protect human and animal health, the environment, economy and the entire planet.

As we have expressed above, air pollution is a very significant challenge that our world is facing and we must take active actions to curb air pollution and improve the quality of life. Hence, the topic that we have chosen for our project is “Air pollution monitoring, forecasting and controlling systems, using IoT and Data Science”. To know the extent of air pollution in a particular geographical area, it is important for us to first measure and monitor the air quality and air pollution in that area, and hence that is the first part of our project, which is to monitor air pollution. Once we have observed the quality of air and level of air pollution, we will get an idea about how to reduce the pollution and improve the quality of air. There is no single cause or solution to air pollution. Therefore, taking action once will not help us fight against air pollution. We as humans need to put continuous efforts and take actions everyday to fight air pollution. The next part of our project is to forecast the future air pollution levels. This will help the users understand the quality of air and pollution levels in the upcoming days and will help them make plans accordingly. It will also serve as a reminder that air pollution cannot be reduced in a single day with a single action. And the last part of the project is to suggest actions and precautions that the users can take to improve the quality of air and reduce the levels of air pollution. Depending on the current quality of air and air pollution levels, the device will intelligently suggest what the users can do to protect themselves in case the pollution level is high, such as wearing a mask or not travelling to certain areas and also suggest methods to curb the pollution such as planting certain plants which will absorb certain harmful air pollutants.

Thus we are implementing a holistic approach to tackle air pollution. With this project we not only aim to reduce the levels of air pollution, but also positively impact the lives of people by helping them change their lifestyle in such a way that it will encourage them to adopt some habits that will help reduce air pollution, which will help them, their pets and their environment to be healthier.

1.2 Motivation

It is now believed that air pollution is the largest environmental health risk on Earth, with 7 million deaths globally annually being attributable to it. It has a severe impact on the lives of human beings and animals, it has a negative impact on the environment, ecosystem and biodiversity of the world, it also impacts the economy and social well-being of the world. Air pollution is a very challenging and complex problem that requires a multifaceted solution that includes the reduction in the usage of fossil fuels, promoting the use of renewable sources of energy, alternative sources of energy and clean energy sources, implementing stricter regulations that govern industrial emissions, and raising awareness amongst the masses. People all around the world must come together and adopt a better lifestyle that will help us tackle the issues of air pollution.

The main causes of the current rapid rise in air pollution, particularly in urban and industrial regions, are vehicle emissions, manufacturing facilities and industries accumulating a large number of dangerous pollutants. In many industrial and urban areas today, maintaining and monitoring air quality has become a top priority. Numerous elements, such as time, location, and uncertain variables have an impact on air quality. Due to the rising levels of air pollution, there is a great need to implement effective air quality monitoring systems that gather data on the concentration of various air pollutants and provide not only the current assessments of the level of pollution, but also predict the level of pollution in the near future. In addition to this, we find it essential to compare the pollution levels with the recommended limit given by organisations such as the World Health Organisation (WHO) and consequently provide measures to prevent air pollution levels from breaching the limits.

Air pollution is not a new problem. It has been affecting the world for a long time. But its importance has only been recognised in recent years due to the humongous increase in the levels of air pollution and its severe impacts. Over the years, few people and organisations have tried implementing air pollution monitoring systems. But these systems did not receive the necessary traction and they are used in only very few parts of the world. Based on our research, we have come to understand that there are 3 main reasons which are preventing the global usage and implementation of these systems. They are: High costs, bulky equipment and low awareness.

With this project, we hope to successfully demonstrate the viability of a low-cost, Internet of Things based system constructed using Arduino and a variety of sensors that can not only monitor the current air quality but also predict the future and offer recommendations on how to reduce pollution and protect people from dangerous pollutants. We hope this system will help regulate and bring down the levels of air pollution and improve the quality of air, which in turn will improve the quality of lives of humans.

1.3 Objective

Our broad objective is to monitor the air quality and the air pollution levels in a given geographical area, then forecast the future air quality and level of air pollution in that area and if the air pollution is excessive, then suggest methods in which the level of pollution can be brought down to the safe limit. We also hope that the users become more aware of the problem at hand and they start adopting better practices that will help fight the severe harmful impacts of air pollution. We aim to achieve the above by building an IoT device that is cost effective, compact and highly accurate.

- Build an IoT system to monitor Air Quality:
    - a) Monitor Concentrations of various air pollutants
    - b) Measure AQI
- Build a model to predict the Air Quality in the near future, using the output of the IoT system as input.
- Build a web application to display the real time findings.
- Provide analytics on the website, such as the maximum recommended limit and the current concentration.
- Provide measures to improve Air Quality.
- Provide precautionary measures.

1.4 Scope

The scope of this project is to successfully build an air pollution monitoring, forecasting and controlling system that is accurate, cost-effective, compact and user-friendly. We aim to leverage the advantages of sophisticated technologies like the Internet of Things and Data Science.

There are three modules in this project: firstly, we start with the data collection and storage. The data that is collected via the sensors is stored in our cloud database. We will then run efficient machine learning and deep learning algorithms that will accurately forecast the future prediction levels of air pollution, which is the second module. Lastly we will display the results in our web application.

1.5 Existing System

There is no one stop solution for eliminating air pollution. With the rising trends of population expansion, urbanisation, industrialization, and increased affluence, the number of potential sources of air pollution grows quickly every year. Because air pollution is caused due to a mixture of different air contaminants, different methods have to be adopted to control the release of the different pollutants into the atmosphere. Catalytic converters are used to control vehicular exhaust, scrubbers are used to control the release of harmful pollutants from industrial emissions, chlorofluorocarbons and hydrofluorocarbons usage is being reduced and alternatives are used, and many such methods are being used to reduce air pollution. However, there is very little worldwide development currently in the approach that we are following: a device which monitors, forecasts and controls air pollution with a holistic approach. The only way people can get accurate updates on air pollution is by reading the Government report at the end of each month. This gives a day by day AQI breakdown for 7 major areas in Bangalore. All other websites provide a singular AQI for the whole city for any given day, which through our research, we have found out to be inaccurate. Therefore, we think it is essential to have local devices. Based on our research, we have come to understand that there are 3 main reasons which are preventing the commercial usage and implementation of these systems. They are: High costs, bulky equipment and low awareness.

1.6 Proposed System

With our project, we hope to successfully demonstrate the viability of a low-cost, Internet of Things based system constructed using Arduino and a variety of sensors that can not only monitor the current air quality but also predict the future and offer recommendations on how to reduce pollution and protect people from dangerous pollutants. We hope this system will help regulate and bring down the levels of air pollution and improve the quality of air, which in turn will improve the quality of lives of humans.

We are suggesting an IoT device because IoT has tremendous potential in this field. IoT devices make it possible to gather and analyse data in real-time, which will help to optimise operations and enhance decision-making, which results in improved efficiency and accuracy. IoT helps with increased connectivity by enabling seamless communication and cooperation by connecting objects, machinery, and people. Another significant advantage of using IoT devices is that these devices can track the functionality and wear of machinery, enabling preventive maintenance which will cut down on downtime. IoT devices also offer enhanced security and safety. They provide real time monitoring and alerting which will reduce the likelihood of mishaps. And lastly, IoT also has economic advantages. It helps to reduce operational costs through improved efficiency and automation.

The forecasting of future levels of air pollution will be done using machine learning and deep learning techniques. We will implement various algorithms such as linear regression, support vector machines, time series forecasting, random forest etc. And then we will compare the accuracy of these models and choose the best performing model. The advantages and applications of machine learning and deep learning are very important. There is an enormous amount of research being conducted in these fields because of their applications and benefits. ML and DL algorithms can be used to analyze huge amounts of data and make predictions. They have the ability to process huge datasets and provide high levels of accuracy. ML and DL techniques can be used to automate tasks that are repetitive in nature, perform such tasks at an increased speed and also reduce the error rate. These algorithms have the ability to identify underlying trends and patterns in huge amounts of data and can extract useful information and also make predictions. With the continued research and development in this field, these technologies will continue to grow and evolve, and will become more important.

Our system will have 3 modules:

1.6.1 Monitoring

An IoT device will be built using Arduino as the microcontroller. We will use the MQ135 Air Quality sensor, MQ131 Ozone sensor, MQ7 Carbon Monoxide sensor, DHT 11 sensor and PM2.5 dust sensor. These sensors will measure the current concentrations of the various pollutants and will output the current AQI. The values of the concentrations of the gasses is to be stored in a database, which will be used to forecast the future levels of AQI.

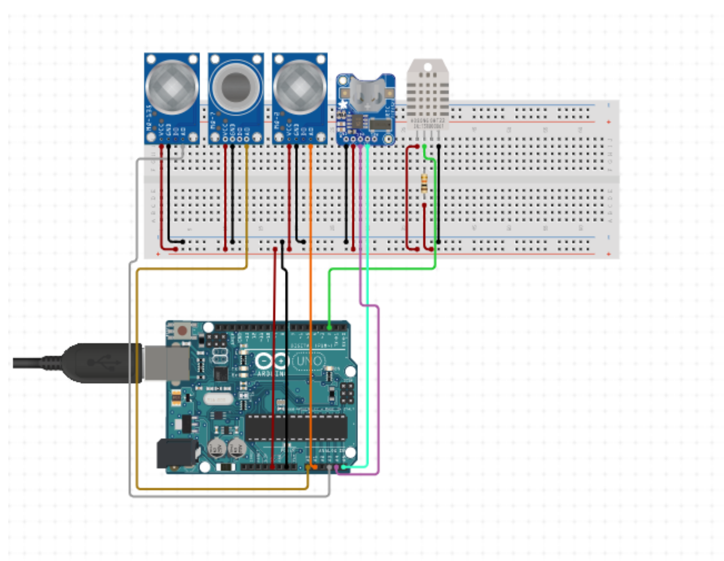

Fig 1.6.1 IoT implementation of the proposed system

1.6.2 Forecasting

Using the data collected in the previous module, an ML model is to be trained to predict the future levels of AQI. Based on the literature survey, the AutoRegression model provided the best results.

1.6.3 Controlling

AQI is categorized into 6 categories based on the value. The categories are:

AQI Value	Category
0 to 50	Good
51 to 100	Moderate
101 to 150	Unhealthy for sensitive groups
151 to 200	Unhealthy
201 to 300	Very Unhealthy
301 and higher	Hazardous

Table 1.6.1 Categorization of AQI

Depending on the category of the current AQI level, suggestions are to be provided to improve the category to ‘Good’.

Hence, using IoT and ML and DL technologies, we hope to build a cost effective and compact device that will sense the current air quality and pollution level, predict the future pollution levels and also offer suggestions on how to safeguard the user’s health incase the pollution level is significantly high and lastly suggest methods to reduce air pollution. We hope this project will help create awareness and is suitable to be adopted by organisations all over the world. We wish to join the movement to curb air pollution and make the world a better place to live in.

CHAPTER-2

PROBLEM STATEMENT

2.1 Problem Statement

With 7 million deaths worldwide attributed to air pollution each year, it is currently thought to be the greatest environmental health danger on the planet. Air pollution causes and aggravates a wide range of diseases like asthma, cancer, heart disease etc. Air pollution is increasing rapidly today, especially in urban and industrial areas primarily due to vehicular emissions, manufacturing facilities and industries accumulating a large number of dangerous pollutants. The lack of effective and accessible monitoring and control systems has hindered progress in addressing this issue. Due to the rising levels of air pollution and its adverse effects, there is a great need to implement effective air quality monitoring systems that are accurate and cost effective.

Using this as our motivation, our problem statement is: “To build an intelligent and automated device that has the ability to monitor, forecast and also control air pollution by leveraging booming technologies like IoT and Data Science”.

2.2 Motivation

It is now believed that air pollution is the largest environmental health risk on Earth, with 7 million deaths globally annually being attributable to it. It has a severe impact on the lives of human beings and animals, it has a negative impact on the environment, ecosystem and biodiversity of the world, it also impacts the economy and social well-being of the world. Air pollution is a very challenging and complex problem that requires a multifaceted solution that includes the reduction in the usage of fossil fuels, promoting the use of renewable sources of energy, alternative sources of energy and clean energy sources, implementing stricter regulations that govern industrial emissions, and raising awareness amongst the masses. People all around the world must come together and adopt a better lifestyle that will help us tackle the issues of air pollution.

The main causes of the current rapid rise in air pollution, particularly in urban and industrial regions, are vehicle emissions, manufacturing facilities and industries accumulating a large number of dangerous pollutants. In many industrial and urban areas today, maintaining and monitoring air quality has become a top priority. Numerous elements, such as time, location, and uncertain variables have an impact on air quality. Due to the rising levels of air pollution, there is a great need to implement effective air quality monitoring systems that gather data on the concentration of various air pollutants and provide not only the current assessments of the level of pollution, but also predict the level of pollution in the near future. In addition to this, we find it essential to compare the pollution levels with the recommended limit given by organizations such as the World Health Organisation (WHO) and consequently provide measures to prevent air pollution levels from breaching the limits.

Air pollution is not a new problem. It has been affecting the world for a long time. But its importance has only been recognised in recent years due to the humongous increase in the levels of air pollution and its severe impacts. Over the years, few people and organizations have tried implementing air pollution monitoring systems. But these systems did not receive the necessary traction and they are used in only very few parts of the world. Based on our research, we have come to understand that there are 3 main reasons which are preventing the global usage and implementation of these systems. They are: High costs, bulky equipment and low awareness.

With this project, we hope to successfully demonstrate the viability of a low-cost, Internet of Things based system constructed using Arduino and a variety of sensors that can not only monitor the current air quality but also predict the future and offer recommendations on how to reduce pollution and protect people from dangerous pollutants. We hope this system will help regulate and bring down the levels of air pollution and improve the quality of air, which in turn will improve the quality of lives of humans.

2.3 Objectives

Our broad objective is to monitor the air quality and the air pollution levels in a given geographical area, then forecast the future air quality and level of air pollution in that area and if the air pollution is excessive, then suggest methods in which the level of pollution can be brought down to the safe limit. We also hope that the users become more aware of the problem at hand and they start adopting better practices that will help fight the severe harmful impacts of air pollution. We aim to achieve the above by building an IoT device that is cost effective, compact and highly accurate.

- Build an IoT system to monitor Air Quality:
  - a) Monitor Concentrations of different air pollutants
  - b) Measure AQI
- Build a model to predict the Air Quality in the near future, using the output of the IoT system as input.
- Build a web application to display the real time findings.
- Provide analytics on the website, such as the maximum recommended limit and the current concentration.
- Provide measures to improve Air Quality.
- Provide precautionary measures.

CHAPTER-3

DETAILED SURVEY

[2018] "Air pollution monitoring and prediction using IoT", Temesegan Walelign Ayele, R. Mehta

The system uses IoT with Long Short Term Memory which is a type of RNN to detect air contaminants. IoT is basically connecting various electronic devices with each other over a network. We can connect almost any device this way, including smartphones, home electronics, lighting, and wearable technology. IoT refers to all internet connected devices that collect, send, and track information from their surrounding environment using built-in sensors, processors, and communication tools. These "smart" or "connected" devices can occasionally interact with other similar devices and follow up on the information they receive from one another.

The strategy works well and yields mediocre outcomes. The results are not properly documented, and there is no finished product (IoT gadget). Improvement possible. The system has not been tested with real time data, nor does it have results published. It offers a Bird's eye view of the architecture and design of the system and the potential usages and applications of it.

Air sensors are used by the system to detect and provide this information to the arduino. The data is then stored by the arduino on the web server. It uses the LSTM algorithm for prediction. It accurately minimises training cycles and offers quick convergence.

The paper also gives an overview of the potential harmful effects of prolonged exposure to these pollutants (Sulphur Dioxide, Nitrogen Dioxide, Ozone, etc.)

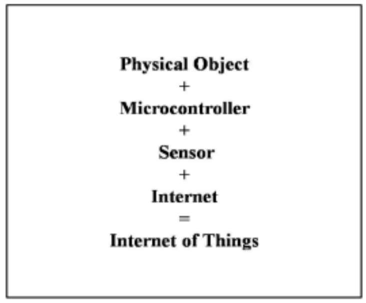

Fig 3.1 Basic Components of IoT

[2021] "Air Pollution Monitoring and Prediction Using IoT and Machine Learning", Ssneha Balasubramanian et al.

In this paper, the authors propose an Internet of Things (IoT) system that makes use of a number of sensors, including the MQ135 and MQ131 sensors as well as the PM2.5 sensor, to record the current concentration levels of different pollutants and predicts the future values of the Air Quality Index using various ML and DL models. The authors also go on to discuss the literature survey they have employed, and how they arrived at building the system.

The Air Quality Index is predicted for the next 5 hours and they use Machine Learning models, including Linear Regression, time series, SVR, and stacking ensemble models. The Ensemble model has the lowest Root Mean Squared Error value amongst all the models.

Their proposed system architecture is to connect the sensors to Aduino Uno, which is connected to the Cloud using Node MCU. Then the prediction is made using the Machine Learning model.

This study forecasts the AQI value for the next 5 hours using a variety of machine learning models, including Linear Regression, SVR, time series models, Stacking Ensemble, etc. The Stacking Ensemble model is the chosen model selected for use since it performs the best when the RMSE of all the models is compared.

[2022] "Air Quality Monitoring And Prediction Using IOT And Machine Learning Approaches", T.S. Kitchilan, M.D.B.C.K. Abeyratne

In this work, sensors built on the Arduino platform were used to create a device for measuring air quality. The sensors were then calibrated to increase accuracy. Random Forest algorithm was used to forecast the air quality index values .They utilized data from the air quality sensor over a period of three months.

The study mainly focussed on the minute details of developing and deploying a portable and easy on the pocket device to capture and measure the parameters that determined the Air Quality.

The authors of this paper have made use of the MQ-131 sensor for Ozone 3, MQ-7 sensor for Carbon Monoxide, Sharp GP2Y1010AU0F for PM 2.5, DHT-11 for temperature and humidity.

Parameter	Sensor	No. of readings	Mean % accuracy	Mean % error
O3	MQ-131	770	90.54	9.46
CO	MQ-7	760	91.89	8.11
PM2.5	Sharp GP2Y1010AU0F	810	89.77	10.23
Temperature	DHT-11	800	91.55	8.45
Relative Humidity	DHT-11	800	89.59	10.41

Fig 3.2 Accuracy of sensors

After the sensors were calibrated, the produced Internet of Things gadget performed with a noticeably higher level of precision. In terms of prediction accuracy, Random Forest Trees surpassed linear regression, decision trees, and neural networks, while runtime efficiency was much enhanced by the use of split data techniques and a Bayesian method for parameter tuning.

The gadget underwent testing and validation, and the results showed an overall accuracy of about 91%. The produced air quality system's accuracy has increased by about 6% to 11%, demonstrating the beneficial effects of the study's calibration techniques.

This work has shown the potential for creating air quality monitors that are affordable, portable, and accurate, allowing for effective real-time monitoring of air quality.

This study showed that non-linear regression approaches are better at forecasting air quality measurements by proving that the non-linear models performed better. It also demonstrates how an accurate air quality forecast model can be successfully implemented using the Bayesian technique.

Additionally, it has been demonstrated that the ensemble machine learning approach outperforms more traditional gradient boosting and regression machine learning techniques in terms of accuracy and efficiency when used to forecast the Air Quality Index.

[2019] "Air Quality Prediction using Machine Learning Algorithms", Sachin Bhimrao Bhoite et al.

This paper talks about how severe the issue of air pollution has become and states that one of the biggest challenges that the government faces is managing the air quality levels in urban and industrial areas. The burning of fossil fuels, traffic patterns, and industrial variables all have a big impact on air pollution. Due to the rising levels of air pollution, it is essential to implement models that will track data on the concentrations levels of air pollutants like sulphur dioxide, nitrous oxide, etc. Their dataset is divided into Train, Validation, and Test in the ratio 0.8: 0.2: 0.2. A time series model was developed once the data was analysed. Auto Regressive models and ARIMA were employed.

This paper discusses the various Machine Learning models that can be employed to solve the problem of predicting the quality of air. However, this research does not talk about the collection of data, rather uses pre collected data to build the models. It provides a good comparison between the various models.

The dataset was obtained from Kaggle. The Mean Squared Error reported by the model is 166.358. The paper only does it for the Maharashtra region, not for the entire nation.

The autoregression time series model uses observations from prior time steps as input to a regression equation and predicts the value at the next time step. It is a rather simple concept that can yield accurate and precise predictions for a number of time series problems. A class of statistical models for assessing and predicting time series data is known as an ARIMA model. With the addition of integration, ARIMA is an extension of the more straightforward Auto Regressive Moving Average.

[2020] "Adaptive machine learning strategies for network calibration of IoT smart air quality monitoring devices", Saverio De Vito, et al.

This paper talks about AQMS which is Air Quality Multi Sensors Systems. They are IoT devices that make use of inexpensive chemical microsensor arrays that have recently demonstrated their ability to produce reasonably accurate quantitative estimates of air pollutant levels. In order to solve the issue of geographical sparseness impacting the current network of AQ Regulatory Monitoring Systems, the easy accessibility of the proposed components allows the implementation of air quality monitoring networks.

Urban authorities and monitoring agencies have recently become interested in the short- to mid-term performances of field data driven calibration models due to seasonal variations in the probability distribution of priors, observables, and hidden context factors.

Different combinations of updates periodicity and number of incoming GT labelled samples were investigated, with a focus on network calibration and applying online machine learning components. The results demonstrated the potential for a field deployed AQMS receiving regular or opportunistic input from high accuracy labelled data sources to significantly improve the performance achievable over more than one year.

Results also show that the number of labelled samples and update intervals are important factors in influencing final performances. To be more precise, the most uncommon updates need a higher percentage of tagged data to meet the same predetermined performance targets.

[2018] "Air Quality Prediction: Big Data and Machine Learning Approaches", Gaganjot Kaur.

Air quality is impacted by many different factors including the location, the time and other hazy circumstances. The big data analytics approach is becoming more popular now due to the research and development in big data applications, the availability of sensing networks and availability of sensor data. The authors of this paper have compared different big data and ML approaches for forecasting the levels of AQI.

This study analyses the findings of studies that have been conducted on decision trees, deep learning, and other AI-based techniques for evaluating air quality. It also sheds light on some of the difficulties and the requirements for further study.

The immense amount of research and development in fields such as machine learning, big data and IoT has made it desirable for future smart cities to have real-time air quality monitoring and evaluation. The paper presents the results of the author's literature analysis, and contrasts contemporary studies on the assessment of air quality using ML and big data.

[2020] "Real Time Localised Air Quality Monitoring and Prediction Through Mobile and Fixed IoT Sensing Network", Dan Zhang et al.

The traditional strategy of installing stationary sensors does not effectively offer a comprehensive view of air pollution in people's immediate surroundings since the nearest sensors may be kilometres away. This study combines both stationary and mobile IoT sensors that are mounted on vehicles that patrol the area to predict the pattern of air quality in that area. With this method, the complete range of how air quality varies in close-by areas may be examined.

This study shows that the method can measure and forecast air quality accurately utilising data from the actual world and several machine learning methods. For a smart city application, this evaluation's outcome for good air quality monitoring and prediction is encouraging.

In this experiment, the authors deployed 3 fixed IoT sensors and 3 mobile IoT sensors fitted onto cars which were roaming around in a given locality.

A communication System was built to enable the different sensors to send back data to a central server through Voice over Long-Term Evolution(VoLTE). The collected data was Geo-tagged and stored in a Database. This data was used to predict the Air Quality metrics for the next day and compared with the ground truth to see how the models were performing. This paper also mentions the importance of being able to model the difference between weekends and weekdays. A simpler way of doing that would be to add traffic density to the dataset.

By merging stationary and mobile sensors, a novel method for forecasting the immediate air quality around people was investigated in this research. The outcomes of the experiments demonstrate how well our suggested hybrid distributed fixed and IoT sensor system predicts the quality of the air near humans. Additionally, by using public transportation systems like buses and taxis—which are already fitted with IoT sensor devices to assess various areas—the proposed system can be made practically realisable. The system's forecasted air quality data can be used in a variety of situations, such as when making outdoor activity plans.

[2020] "IoT enabled Environmental Air Pollution Monitoring and Rerouting system using Machine learning algorithms", Leeban Moses, et al.

Air quality index is a parameter that can be used to gauge air quality. The AIR quality index is calculated using a number of factors. The main gases whose concentrations contribute the most are Carbon monoxide, sulphur dioxide, nitrogen dioxide, ozone and particulate matter. In this paper, they have discussed the implementation of an IoT device that is connected to the cloud. The sensors connected to the microcontroller unit first collect the concentration levels of the various pollutants. They also measure the humidity and temperature. By utilising Lora nodes and Lora Gateway, the acquired data may be updated on a cloud platform. The Google map API used in the development of the online application allows for frequent updates to the pollution status.

The prediction study for particulate matter was carried out using the neural network Multi-Layer Perceptron and SVR learning method with the acquired time series samples.

AQI is expressed as a number between 0 and 500. The lower the AQI value, the better the quality of air. The concentration levels of various pollutants is averages and is used to calculate the air quality index (AQI). For instance, the 24-hour average concentration is used to calculate the AQI for PM2.5.

A web application is used to continuously monitor the concentration levels of the air contaminants.

The AQI can be calculated as:

$$I_{p} = [\{(I_{HI} - I_{LO}) / (B_{HI} - B_{LO})\} * (C_{p} - B_{LO})] + I_{LO}$$

B_{HI} = Break point concentration greater or equal to given conc.
B_{LO} = Break point concentration smaller or equal to given conc.
I_{HI} = AQI value corresponding to B_{HI}
I_{LO} = AQI value corresponding to B_{LO}

Fig 3.3 AQI Formula

In this paper, they have discussed the implementation of an IoT device that is connected to the cloud and uses sensors to collect the concentrations of pollutants like ozone, carbon monoxide, sulphur dioxide etc. The estimated AQI value is used to update the maps' with the new levels. A passenger can choose a route which is least affected by air pollution. The user can also schedule trips in advance by using some specialised machine learning techniques that forecast future values of AQI.

[2019] "Air pollution prediction through internet of things technology and big data analytics", Safae Sossi Alaoui et al.

This paper brings together 3 important concepts of Big Data Analytics, Internet of Things and Data Science to be able to properly predict Air Pollution.

Apache Spark is an open source processing framework. It is a quick and versatile cluster computing system. It offers high-level APIs in a variety of programming languages, such as Java, Scala, Python, and R, and features an engine that has been optimised to accommodate large execution graphs. It also supports quick application development for huge data. A broad range of sophisticated tools are available, such as Spark SQL for handling SQL queries and structured data, MLlib for machine learning (ML), GraphX for handling graphs, and Spark streaming.

Gradient-boosted Trees (GBTs) was the Machine Learning algorithm selected. Decision trees are iteratively trained by GBTs to minimise a loss function. GBTs are similar to decision trees in that they handle categorical features, can be applied to multiclass classification settings, don't require feature scaling, and can capture nonlinearities and feature interactions.

The 6 main steps for building a GBT model are:

1) Data pre-processing
2) Setting up the environment
3) Data Loading
4) Data Splitting
5) Data Modelling
6) RMSE measuring

[2012] "Developing a risk-based air quality health index", Tze Wai Wong, et al.

Based on the Canadian method, a risk based, multiple pollutant air quality health index (AQHI) reporting system was created in Hong Kong. The risks of respiratory and cardiovascular disorders linked to four air pollutants: So2, No2, O3, and particulate matter with an aerodynamic diameter less than 10 mm were determined using time series analyses (PM10). These air contaminants' combined increased risks for hospital admissions were totaled up. These pollutant concentrations' designated as short-term Air Quality Guidelines by the World Health Organization, served as the basis for the cut-off values of the accumulated excess risk, for the issue of various health warnings.

The calculations performed to get the Air Quality Index from Pollutants like So2, No2, O3, PM10, etc. were observed and can be used in our project. Health advice was given as well.
For all age groups combined, the RRs of emergency hospital admissions for respiratory and cardiovascular disorders were considerably increased for all 4 air pollutants (NO2, O3, PM10, and SO2). In the 5-year study period, the extra risk of hospital admissions linked to air pollution

ranged from 2.64% to 31.51%, with a median of 9.04% and a mean of 9.50%. Additionally, the health advice table is made public..

Table 2
Distribution of percentage excess risk (%ER) of hospital admissions for cardiovascular and respiratory diseases by health risk category and AQHI band.

Recommended health risk category	AQHI band	%ER	No. of days	Frequency (%)
Low	1	0–1.88	0	0.0
	2	>1.88–3.76	36	2.0
	3	>3.76–5.64	333	18.2
Moderate	4	>5.64–7.52	277	15.2
	5	>7.52–9.41	339	18.6
	6	>9.41–11.29	306	16.8
High	7	>11.29–12.91	194	10.6
Very high	8	>12.91–15.07	172	9.4
	9	>15.07–17.22	93	5.1
	10	>17.22–19.37	27	1.5
Serious	10+	>19.37	49	2.7
Total			1826	100.00

Fig 3.4 Distribution of %ER of hospital admission for cardiovascular and respiratory diseases

Table 2 displays the %ERs that served as the cut-off thresholds for the five risk categories. For the general public, a summed %ER of 12.91% was recorded (all ages). This became the criterion for inclusion in the "high risk" category. The 12.91% %ER was reduced to 11.29% for the vulnerable groups, which include both children under the age of five and seniors 65 years and older. 4 This served as the lowest cut-off for the category of "high risk." The lower cut-off point

for the "moderate risk" category was determined to be half of this (5.64%), and the %ER at or below 5.64% was classified as "low risk." The category of "severe risk" is associated with an ER of 19.37%.

[2021] "Air pollution monitoring system using IoT and Artificial Intelligence", Meivel S, et al.

The paper proposes a system which keeps a track of combinations of gases such as CO2, methane and dust. The proposed model uses gas sensors which are responsible for detecting the gas level of the environment. Threshold levels for each gases are decided and the signals detected from the sensors are fed into the microcontroller. ESP8266 can fetch data and load into the Internet of things. The resistance of material inside the sensor was altered, due to which the potential difference based on the gas concentration changed in the sensor which is quantified of as the output voltage. An alarm will sound if the concentration of any of the gases exceeds a predetermined level.

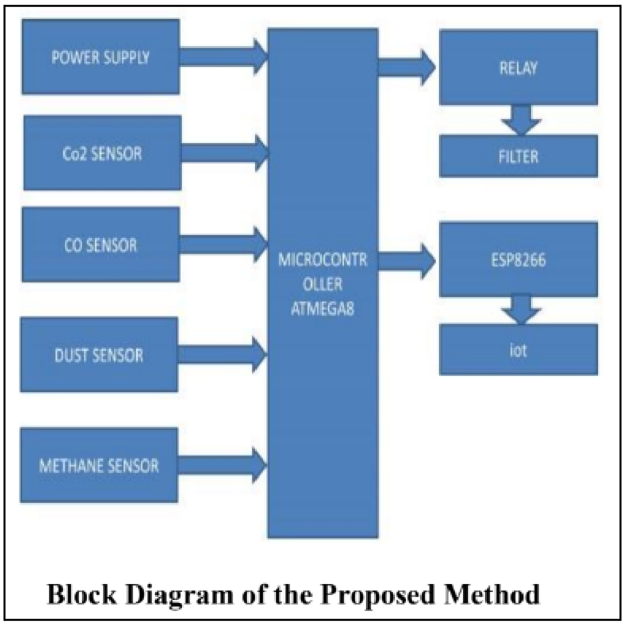

Fig 3.5 Block diagram of proposed model

The model created is a low-cost, high-fidelity air quality monitoring system. It is put in place, and tested. The system will collect data for each instant, send it through Wi-Fi, and notify the staff when the threshold level has been reached.

[2020] "Modelling air quality prediction using a deep learning approach: Method optimization and evaluation", Weilin Wang, Limin Jiao, Suli Zhao, Anbao Liu et al.

In this article, a deep learning framework is introduced which uses a neural network with a temporal sliding Long Short Term Memory extended model (TS-LSTME) to predict the air quality over every hour of the next day. Through a bidirectional long short-term memory (LSTM), which also happens to be multi-layered, that included the hourly historical PM2.5 concentration, temporal data, and meteorological data, the model incorporated the ideal time lag to realise sliding prediction. The suggested approach was used to forecast the upcoming 24-hour average PM2.5 concentration in China's most severely polluted region, Jing-Jin-Ji.

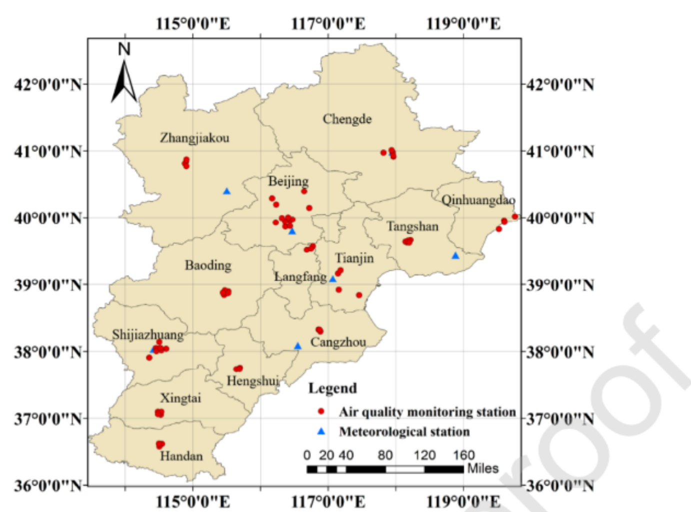

Figure 1. Distribution of air quality monitoring stations and meteorological stations over Jing-Jin-Ji area.

Fig 3.6 Distribution of air quality monitoring stations in Jing-Jin-Ji area

Additionally, the suggested model may produce highly accurate PM2.5 concentration estimates in long-term series (48 h and 72 h). The model was also put for predicting O3 concentration to the test. Other air contaminants might be subject to the proposed model's application. The suggested techniques can help early warning and regional pollution management while also enhancing public access to information services about air quality prediction.

In comparison to multiple linear regression, support vector, the conventional LSTM, and the LSTME, the suggested model performed better and had a higher correlation coefficient R

Squared. The suggested approach has great applicability and is suitable for tasks requiring long-term forecasting.

[2020] "Impact of lockdown on air quality in India during COVID-19 pandemic", Ramesh P. Singh, et al.

Using ground-based and satellite-based measurements, the air quality (particulate matter-PM2.5, Air Quality Index, and tropospheric NO2) across India is analysed in this work. Air quality index (AQI) and PM2.5 levels have significantly decreased. Tropospheric NO2 concentration also showed a lowering tendency over Delhi, Mumbai, Hyderabad, Kolkata, and Chennai during the lockdown period in 2020 compared to the same period in 2019. The air quality has greatly improved during the period of total lockdown, which gives the city administration crucial information for developing rules and regulations on how to improve air quality. The trajectory analysis was performed using HYPSPLIT model.

The figure depicts the effects of a total lockdown in India, the average PM2.5 concentrations for March 2020, the average PM2.5 concentration for March 2019, and the average PM2.5 concentration for March 2020 after the lockout period (March 22–31, 2020).

Station Name	Latitude	Longitude	Box coordinates for NO_2 data from OMI satellite
Delhi	28.59	77.18	W-76.68, S-28.07, E-77.68, N-29.07
Kolkata	22.54	88.35	W-87.86, S-22.08, E-88.86, N-23.08
Mumbai	19.06	72.86	W-72.42, S-18.55, E-73.42, N-19.55
Hyderabad	17.44	78.47	W-77.78, S-17.01, E-78.78, N-18.01
Chennai	13.05	80.25	W-79.76, S-12.56, E-80.76, N-13.56

Fig 3.7 Location of US embassies in India

The association between PM2.5 and AQI and the fact that one of the main causes of poor air quality is vehicular motion were the paper's main findings. Following an analysis of India's biggest cities, it lists all the potential causes of air pollution.

In comparison to these sizable, densely populated urban footprints of India, where US embassies are situated, the lockdown appears to demonstrate a dramatic improvement in air quality. Our findings indicate a significant decrease in air pollution during lockdown, particularly in Delhi and Kolkata, which are known to have high levels of pollution in both India and the rest of the world.

[2018] “A Review on Indoor Air Quality Monitoring using IoT at Campus Environment”, Anindya Ananda Hapsari, et al.

Systems for enhancing indoor air quality (IAQ) provide a useful way to maintain a healthy environment. This study's objectives are to examine IAQ monitoring systems, review prior research, and suggest future IoT-based IAQ monitoring investigations. The author of this systematic study summarises and analyses papers concerning IAQ utilising IoT that were pulled from three databases.

This study's objective is to evaluate an indoor air quality (IAQ) monitoring system that uses an Internet of Things (IoT) approach on a campus area by conducting a thorough evaluation of previous research. The outcomes of the articles were divided into five categories in order to further examine the issue. The categories include IAQ monitoring system, IAQ on campus, IAQ data gathering, monitoring IAQ utilising Internet of Things, and IAQ monitoring problems. This study classified a different study that included IAQ and only focused on literature. The material acquired was carefully and thoroughly examined from sources that were pertinent to the topic.

The goal of this project is to make it easier for researchers to use the Internet of Things to build a monitoring system for indoor air quality. This article discusses the systems and categorises earlier relevant studies. Sensors, protocols, and the Internet of Things were used in the discussion material on monitoring systems to combat IAQ. The MQ gas sensor, DHT, and SHT are examples of inexpensive sensors that are frequently employed and can be used in an IAQ monitoring system. The study demonstrates that several nations consider the requirement for installing monitoring equipment in indoor spaces. Evidence that many researchers study indoor air quality (IAQ) each year shows that it is still an issue worth talking about. Additionally, IAQ topics utilising the Internet of Things must be created for sustainable development.

[2017] ".Air Quality Monitoring System Based on IoT using Raspberry Pi", Somansh Kumar, et al.

An independent real-time air quality monitoring system is presented in this paper. A cutting-edge method for better managing data from various sensors may be found in the Internet of Things and cloud computing. This data is collected and transferred by the low-cost, ARM-based Raspberry Pi minicomputer. The system is tested in Delhi, where measurements are tabulated and compared with information from the regional environment control authority.

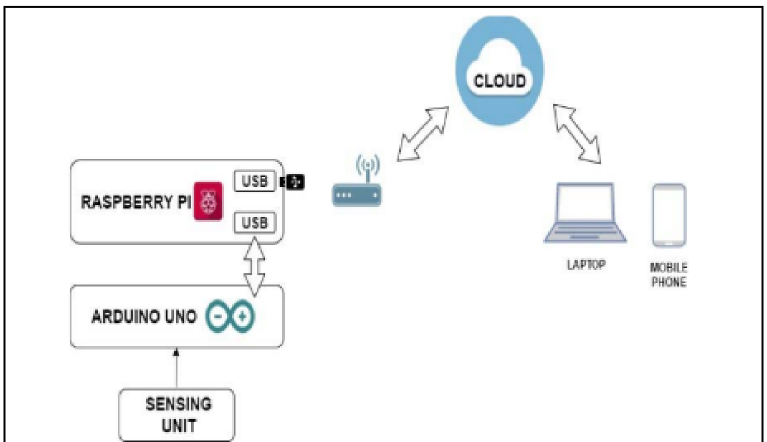

Fig. 1.Simplified diagram of Proposed System

Fig 3.8 Diagram of the proposed system

The sensor-based air quality monitoring system is very accurate, reasonably priced, and simple to operate. A PM sensor, DSM501A, is interfaced to Arduino's digital pin 5, while DHT22 and BMP180 are attached to its digital pins 3 and 4, and MQ135 and MQ9 are connected to its analogue pins 2 and 3. A USB cable connects Arduino to Raspberry Pi. A Wi-Fi adapter is used to connect Raspberry Pi to the internet, and Raspberry Pi is connected to the adapter through a USB port.

A cutting-edge method for better managing data from various sensors may be found in the Internet of Things and cloud computing. This data is collected and transferred by the low-cost, ARM-based Raspberry Pi minicomputer.

[2020] "Air Quality Index Forecasting using Auto-regression Models", Nimisha Tomar, et al.

This study's goal is to develop a daily AQI forecasting model that may be used for regional and local management of air quality. Weather-related variables like wind direction and speed, humidity, temperature, rainfall patterns, and solar radiation, as well as artificial variables like gases produced by industrial waste, pollutant gases from vehicles, and used fuel from radioactive activity and hydrogen bomb testing, have a significant impact on both the causes and the effects of air pollution.

Time series models have been shown to perform better than other models for data that is serially associated. So, the ARMA model is employed. Moving average is used here and AR stands for auto-regression. However, they do not apply to series that are not stationary. Therefore, the main challenge is to use stochastic models to stationarize and predict the time series if it is not stationary. Detrending, differencing, and other techniques can be used to change a non-stationary series into a stationary series.

The MSE was 27.00 when auto-regression was used on a time series dataset to forecast the Air Quality Index value seven days in advance. The MSE can be kept to a minimum by reducing the difference between the present day and the future day when the AQI is to be measured. By keeping the difference low, the model will increase its accuracy

[2012] “Forecasting of daily Air Quality Index in Delhi”, Anikendar Kumar, et al.

Government organisations can utilise the Air Quality Index (AQI), which can be calculated using a formula based on an extensive examination of the concentration of air pollutants, to describe the state of the air quality at a specific site. The goal of the current work is to create a forecasting model for daily AQI prediction that can serve as the foundation for decision-making procedures.

In Delhi, three statistical models—principal component regression (PCR), time series auto regressive integrated moving average (ARIMA), and a combination of both (model 3)—are used to forecast the daily AQI for each season. Pollutant concentrations are used to evaluate the performance of all three models, and it is clear that model 3 performs the best in terms of agreement with observed values when compared to model 1 and model 2. The statistical parameters also support the same.

The performance of Model 3, which combines ARIMA and PCR to tackle the problem of utilising the autocorrelation and collinearity in the variables, is satisfactory. As a result, this model 3 can be used to anticipate air quality in additional Indian urban centres. Although the model is functioning satisfactorily, it still has a lot of uncertainty.

[2017] “Daily Air Quality Index forecasting with hybrid models: A Case in China”, Suling Zhu, et al.

The current models used for forecasting or predicting pollution in the air such as AutoRegressive Integrated Moving Average, Support Vector Regression and Multiple Linear Regression are not adequate. They are unable to understand the underlying trends in the time series data . As a result, it is necessary to provide new, efficient methods for predicting air pollution indexes. To overcome the aforementioned issues and improve forecasting accuracy, the major goal of the current research is to create efficient forecasting models for regional Air Quality Indexes. In order to forecast AQI data, two hybrid models are proposed. These are EMD - SVR - Hybrid and EMD - IMFs - Hybrid.

The following are the main steps of the EMD - SVR - Hybrid model: The original AQI data is sorted using the data pre-processing technique EMD (empirical mode decomposition), yielding one group of smoother IMFs and a noise series, where the IMFs hold the crucial data (level, fluctuations, and others) from the original AQI series. The sum of the IMFs is predicted using LS - SVR, and the residual sequence of LS - SVR is predicted using S-ARIMA (seasonal ARIMA). Additionally, EMD - IMFs - Hybrid forecasts each IMF individually using statistical models before combining the findings as EMD - IMFs.

Comparing the suggested hybrid models to ARIMA, SVR, EMD - GRNN, GRNN, Wavelet - GRNN, and Wavelet - SVR, they show higher predicting accuracy. As a result, the suggested hybrid models can be employed as efficient and user-friendly tools for managing and issuing air pollution warnings.

In addition, for AQI forecasting, hybrid models with EMD and residual revision, such as ARIMA, SVR, EMD - GRNN, GRNN, Wavelet - GRNN, and Wavelet - SVR, perform better than individual models and hybrid models without residual revision.

CHAPTER-4

SURVEY SUMMARY TABLE

Sl.No	Title of the Paper	Problem Addressed	Authors Approach / Method	Results
1	“Air pollution monitoring and prediction using IoT”	Building IoT device to record pollutants and calculate the AQI, and predict future AQI using Machine Learning models	The system employs IoT in conjunction with the Machine Learning method known as Recurrent Neural Network, specifically Long Short Term Memory, to focus on monitoring air contaminants.	The LSTM is implemented for forecasting which has a quick convergence and reduces the training cycles, and does this with a good accuracy.
2	“Air Pollution Monitoring and Prediction Using IoT and Machine Learning”	Building an IOT device to monitor and forecast the concentrations of Pollutants and using Data Science to predict the concentrations of pollutants in the future.	They put out a system that employs a number of sensors to measure the concentration of pollutants, and they created a number of ML models to project future pollution levels.	The stacking ensemble model had the highest accuracy and lowest error rate of all the implemented ML models.
3	“Air Quality Monitoring And Prediction Using IOT And Machine Learning Approaches”	Building an IOT device to monitor and forecast the concentrations of Pollutants and using Data Science to predict the concentrations of pollutants in the future.	In this study, arduino-based sensors were used to create an air quality monitoring system that could record parameters of the air quality before the sensors	The overall accuracy of the device according to testing was 91%. It demonstrates an increase in accuracy of between 6% to 11%, demonstrating the

			were calibrated to increase accuracy.	beneficial effects of the study's calibration processes.
4	"Air Quality Prediction using Machine Learning Algorithms"	Predicting Air Quality Index using numerous Machine Learning Algorithms on a monthly basis.	Dataset from Kaggle and models used are AutoRegressive model and ARIMA	Predicted the AQI of Maharashtra with an MSE of 166.358
5	"Adaptive machine learning strategies for network calibration of IoT smart air quality monitoring devices"	Overcoming the limited accuracy of long term field deployment of AQMS.	For the purpose of overcoming the impacts of drift brought on by the deployment of low cost AQMS across a number of seasons, they have achieved findings from various calibration update scenarios.	Results show that both update intervals and the number of tagged samples are important in influencing final performances. updated every day based on a single hour of very accurate data.
6	"Air Quality Prediction: Big Data and Machine Learning Approaches"	Predicting Air Quality Indexes using techniques of Big Data Analysis and some Machine Learning algorithms	The purpose of this study is to investigate several big data and machine learning approaches for air quality forecasting. It has also contrasted big data-based models with already published research articles.	The development of IoT infrastructures, Big Data Analysis, and Machine Learning Architectures has made it desirable for future smart cities to have real-time air quality monitoring and evaluation.
7	"Real Time Localized Air Quality Monitoring and Prediction Through Mobile and Fixed IoT	Using mobile and static IoT sensors as well as ML algorithms, air quality is	This study combines both stationary and mobile IoT sensors that are mounted	Gradient Boosting Regressor gave the least RMSE for P.M2.5

	Sensing Network”	monitored and predicted.	on vehicles that patrol the area to predict the pattern of air quality in that area. With this method, the complete range of how air quality varies in close-by areas may be examined.	
8	“IoT enabled Environmental Air Pollution Monitoring and Rerouting system using Machine learning algorithms”	Air pollution monitoring system used as a rerouting system to find pollution free route	The prediction study was carried out using the neural network Multilayer Perceptron and support vector machine regression learning technique using the obtained time series data.	The Google Maps API correctly updated the pollution levels, and the advised route was the one with the lowest levels of pollution based on those levels.
9	“Air pollution prediction through internet of things technology and big data analytics”	Using Big Data Analytics and the Internet of Things, anticipate air pollution	In the context of forecasting air pollution that results from the introduction of dangerous compounds, such as NO2, SO2, CO, and O3, into the Earth's atmosphere, we explore the potential for a fusion between the two new ideas big data and internet	With the use of the US pollution dataset and the Databricks platform's Spark technology, it was able to create a precise model that is capable of producing reliable forecasts for air quality.

			of things in this research.	
10	"Developing a risk-based air quality health index"	Finding out the different pollutants in the air and analysing how the affect the Air Quality and consequently the health of people	The relative risks of hospital admissions for respiratory and cardiovascular disorders linked to four air pollutants—sulfur dioxide, nitrogen dioxide, ozone, and particulate matter with an aerodynamic diameter smaller than 10 mm—were determined using time series analyses (PM10).	For all age groups combined, the RRs of emergency hospital admissions for respiratory and cardiovascular disorders were considerably increased for all four air pollutants (NO2, O3, PM10, and SO2).
11	"Air pollution monitoring system using IoT and Artificial Intelligence".	Air Pollution monitoring system to detect the concentrations of pollutants, calculate and predict AQI, built using Big Data Analytics and Data Science	In this work, they suggest a system that tracks CO2, dust, and methane levels. To measure the concentration of these gases in the atmosphere, they employed an IoT device.	They successfully implemented the system and were able to detect the concentration of the various gases in the atmosphere. If a particular gas concentration was greater than the set threshold, an alarm was triggered.
12	"Modelling air quality prediction using a deep learning approach: Method optimization and evaluation"	Modelling Air Quality prediction using Deep Learning and allied fields is the problem chosen by these authors.	A deep learning framework to forecast air quality for the next 24 hours is suggested in this study. a neural network	Comparing the suggested model to multiple linear regression, support vector, the conventional LSTM, and the

			extended model with temporal sliding long short-term memory (TS-LSTM)	LSTME, the new model performed better in terms of stability and performance with a high correlation coefficient R Squared.
13	“Impact of lockdown on air quality in India during COVID-19 pandemic”	Analysing the Air Quality during lockdown to find out the main contributors to pollution and what caused its reduction	The HYPSPLIT model was used to do the trajectory analysis. an examination of the air quality across India utilising ground-based and satellite-based measurements (particulate matter-PM2.5, Air Quality Index, and tropospheric NO2).	The key takeaways from the paper were the correlation between PM2.5 and AQI and how Vehicular Motion is one of the major contributors of Air Quality being poor.
14	“A Review on Indoor Air Quality Monitoring using IoT at Campus Environment”	Analysis on various indoor air quality measuring systems	The best approach to create a low-cost IAQMS is using an Arduino board. Temperature, humidity, and CO2 are the pollutants that are utilised in observations the most frequently. The IAQ monitoring system's next task is to increase the sensor's accuracy under the	The goal of this study was to make it easier for researchers to use the Internet of Things to build a monitoring system for indoor air quality.

			circumstance of "1 ppm".	
15	"Air Quality Monitoring System Based on IoT using Raspberry Pi"	IoT-Based Air Quality Monitoring System Using Raspberry Pi	An independent real-time air quality monitoring system is presented in this paper. A cutting-edge method for better managing data from various sensors may be found in the Internet of Things and cloud computing. This data is collected and transferred by the low-cost, ARM-based Raspberry Pi minicomputer.	A cutting-edge method for improved data management is provided by the Internet of Things and cloud computing. Data from various sensors is collected and transferred by the low-cost, ARM-based Raspberry Pi minicomputer.
16	"Air Quality Index Forecasting using Auto-regression Models"	To forecast the AQI while taking into consideration external factors such as Wind speed, Atmospheric Pressure, Mean temperature, etc.	The ARIMA model is employed. Detrending, differencing, and other techniques can be used to change a non-stationary series into a stationary series.	The MSE was 27.00 when auto-regression was used on a time series dataset to forecast the Air Quality Index value seven days in advance.
17	"Forecasting of Daily Air Quality Index in Delhi"	To forecast the Daily Air Quality index in Delhi to improve the decision making process of the Government and	3 different models were built; ARIMA, PCR (a combination of PCA and MLR), and a combination	The 3rd model; a combination of PCR and ARIMA, seemed to understand the underlying correlations and

		help people take necessary measures.	of the aforesaid 2 models.	patterns between the data points to give the best results.
18	“Daily Air Quality Index forecasting with hybrid models: A Case in China”	To build better and more accurate models than the currently existing models such as ARIMA, multiple linear models and SVR.	Two hybrid models(EMD-SVR-Hybrid and EMD-IMFs-Hybrid) are proposed.	The results of these hybrid models are better than ARIMA, SVR and other models such as EMD-GRNN, GRNN, etc.

CHAPTER-5

SYSTEM REQUIREMENT SPECIFICATION

5.1 Functional Requirements

1. **Connectivity**: Ability of the system to connect and communicate with various sensors, devices and networks
2. **Data collection and storage**: Ability of the system to collect data from various sensors, store it and upload it the cloud storage
3. **Integration**: Ability to integrate the system with other systems like database and cloud offerings
4. **Alerts and Notifications**: Ability to issue the appropriate notifications and alerts for specific conditions.
5. Ability to analyse the collected data and perform accurate prediction in real time.

5.2 Non-functional Requirements

1. **Accuracy**: Accuracy is predictions and issuing alerts and notifications should be greater than 97%
2. **Reliability**: Ability to ensure consistent and accurate data collection, data analysis and predictions
3. **Usability**: Ability to provide an extremely interactive and user friendly interface so that the users can have a good experience using our system
4. **Cost Effectiveness**: We want to provide a cost effective solution which is economic and affordable.
5. **Robustness**: Ability to work with noisy and corrupted data and filter them out.

6. **Scalability**: Ability to handle the enormous amount of data and processing, and also to allow the implementation of more number of devices

7. **Supportability**: Ability to perform updates and modifications with ease

5.3 Hardware Requirements

Sl. No.	**Component**
Basic Components	
1	Arduino Uno
2	Breadboard
3	WiFi module
4	Jumper Wires
7	Resistor set
8	Capacitors
9	Computer/laptop with internet connection
Sensors	
1	MQ131 Ozone sensor
2	MQ135 Air Quality sensor
3	PM 2.5 Particle sensor
4	MQ7 CO sensor
5	DHT11 sensor

Table 5.3.1 Hardware requirements

5.4 Software Requirements

- IDEs such as Visual Studio Code, Google Colab and Arduino IDE
- Python and libraries: NumPy, Pandas, Tensorflow, Seaborn, Statsmodels, Seaborn
- MySQL database
- Web hosting service
- TeraTerm software

CHAPTER-6

SYSTEM DESIGN

6.1 System Design

6.1.1 System Architecture

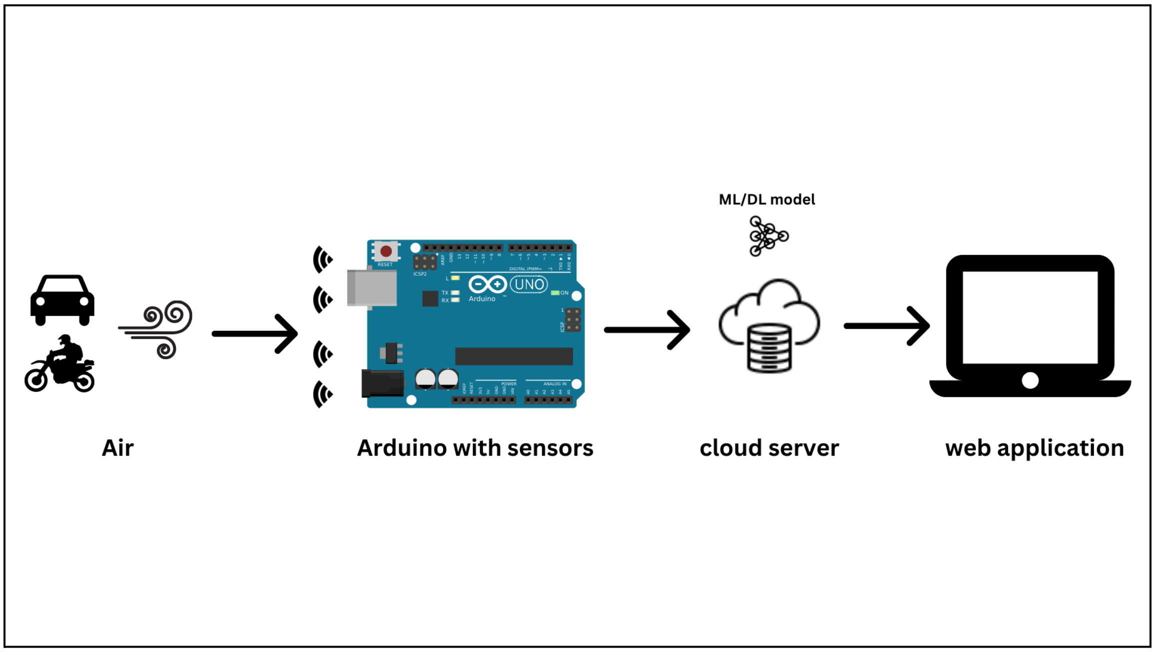

Fig. 6.1.1 System Architecture

1. **Arduino and sensor**s: The sensors collect the concentration of various air pollutants in that locality.

2. **Gateway**: Acts like a bridge between the Arduino and our cloud database. Wifi module or Node MCU will be used to transfer and store data in a cloud database like Firebase

3. **Cloud Platform**: Used to store the data and also run our ML and DL models

4. **Web application**: Displays the data to the users along with suggestions if air pollution level in too high

6.1.2 Module Design

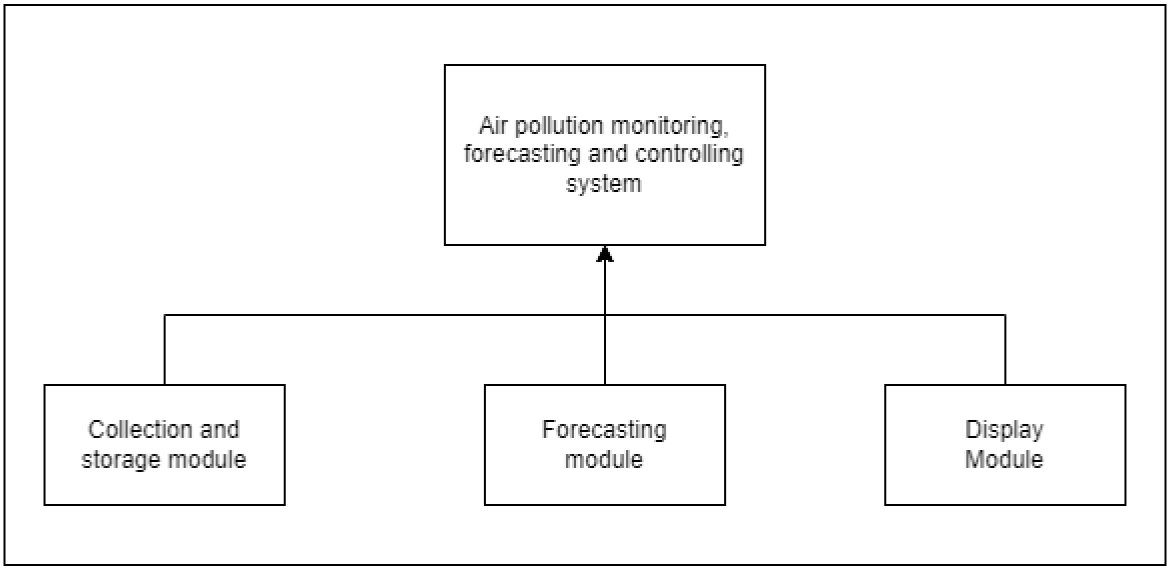

Fig 6.1.2 Module Design

6.1.2.1 Collection and storage of data

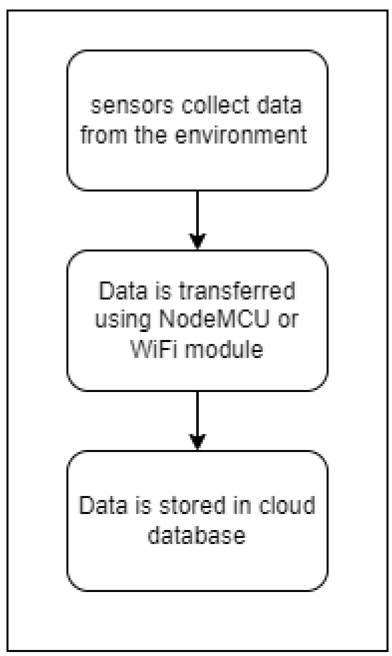

Fig 6.1.2.1 Collection and storage of data module design

6.1.2.2 Forecasting future predictions

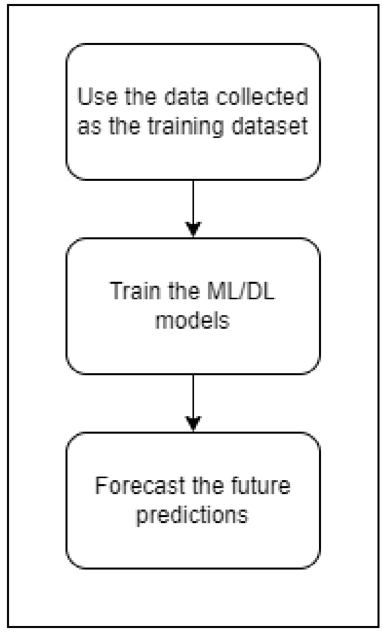

Fig 6.1.2.2 Forecasting future predictions module design

6.1.2.3 Displaying the results

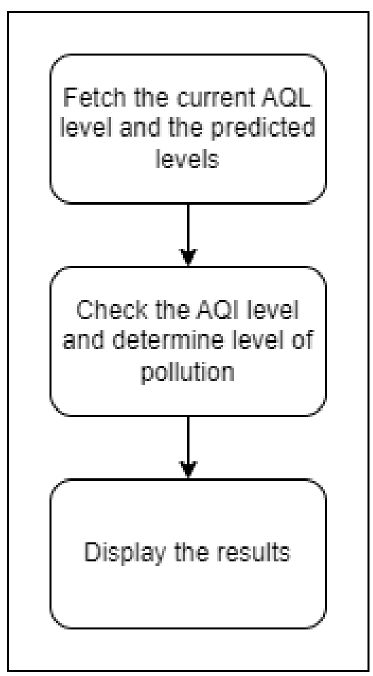

Fig 6.1.2.3 Displaying results module design

6.2 Detailed Design

6.2.1 Class Diagram

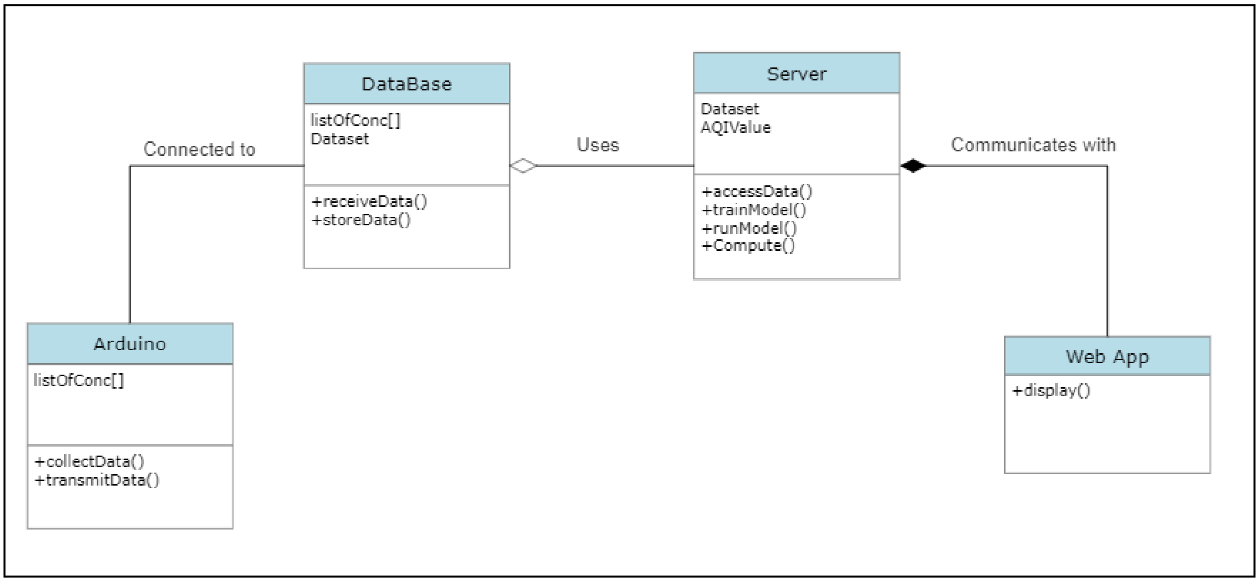

Fig 6.2.1 Class Diagram

6.2.2 Activity Diagram

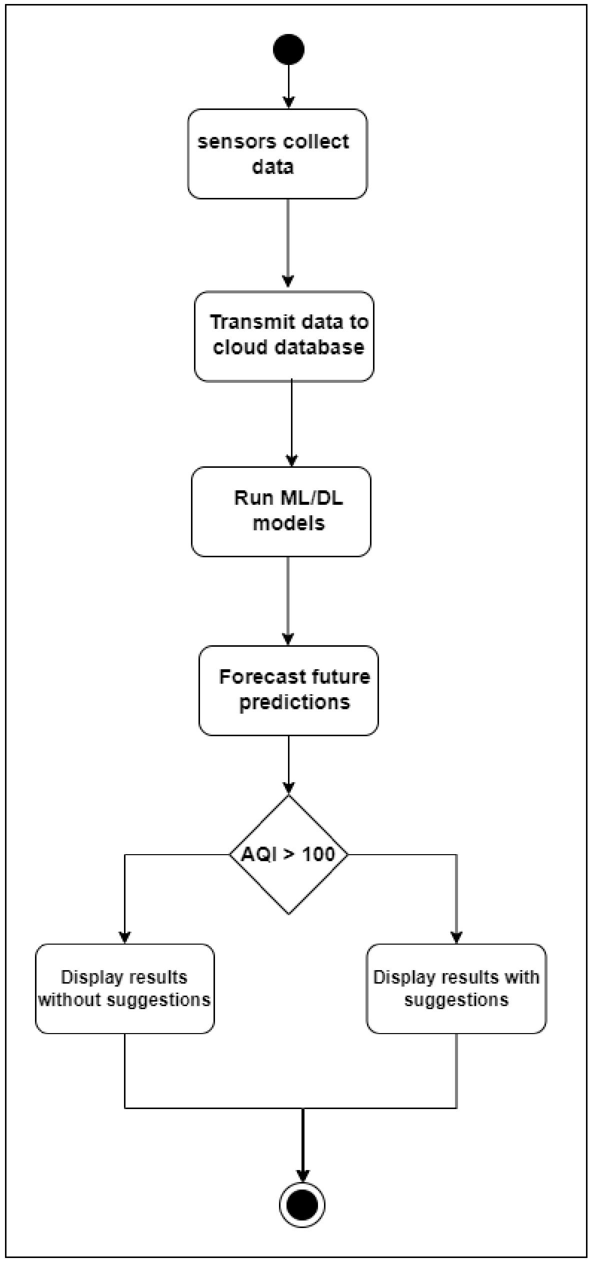

Fig 6.2.2 Activity flow diagram

6.2.3 Use Case Diagram

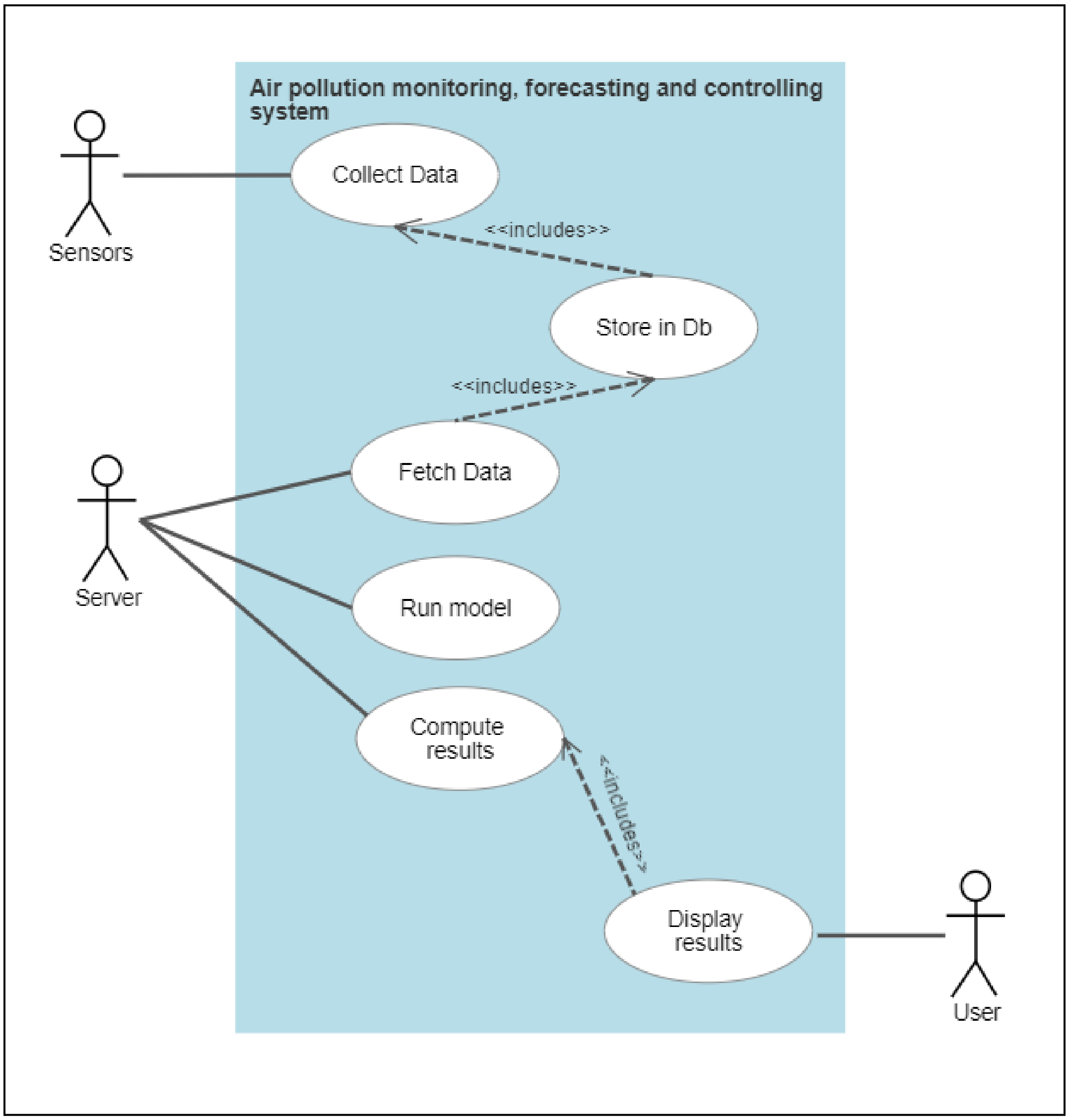

Fig 6.2.3 Use Case Diagram

6.2.4 Scenarios

There are three possible scenarios for our project:

- Air Quality Index has a value below 50: The air pollution level is minimum and the air quality is at its best. It is the most ideal situation to live in.
- Air Quality Index has a value between 50 and 100: Humans and animals can still live in this condition without severe impact on health. Air pollution level is low.

Precautionary actions must be taken to reduce air pollution or at least maintain the same levels.

- Air Quality Index has a value greater than 100: The quality of air is poor. Humans and animals are at high risk of getting impacted from the negative effects of high levels of air pollution. Immediate actions must be taken to reduce air pollution and safeguard human and animal health.

CHAPTER-7

IMPLEMENTATION

7.1 Hardware Implementation

Arduino UNO:

Arduino Uno is an open-source microcontroller board based on the ATmega328P microcontroller. It has 14 digital input/output pins, 6 analog inputs, a 16 MHz quartz crystal, a USB connection, a power jack, an ICSP header, and a reset button.

The Uno can be programmed using the Arduino IDE (Integrated Development Environment) and can be used to control a wide range of electronic projects.

Applications: Home Automation, IoT, Robotics etc.

MQ135 Sensor:

The MQ135 sensor is a gas sensor that is capable of detecting various gases in the atmosphere and provides the AQI value of the air. The sensor is based on the electrochemical principle and can detect a wide range of gases, including carbon monoxide, methane, ammonia, and other organic compounds. It has a small size and low power consumption, making it suitable for use in portable electronic devices. The MQ135 sensor operates by heating a sensing element and measuring the changes in its resistance when exposed to different gases. The sensor is calibrated to provide an analog voltage output that corresponds to the concentration of the detected gas.

Applications: air quality monitoring systems, etc. indoor air quality monitoring

MQ131 Sensor:

MQ131 is a gas sensor that is designed to detect the presence of ozone gas in the air. It is based on the electrochemical principle and can detect a wide range of ozone concentrations, from 10ppb to 2ppm. The sensor has a small size and low power consumption, making it suitable for use in portable electronic devices.The MQ131 sensor operates by measuring the changes in its resistance when exposed to ozone gas. The sensor is calibrated to provide an analog voltage output that corresponds to the concentration of ozone in the air.

Applications: The MQ131 sensor is commonly used in air quality monitoring systems, especially in industries where ozone is used or produced, such as water treatment plants, semiconductor manufacturing, and air purifiers. It is also used in environmental monitoring and research applications, such as studying the impact of ozone on plant growth and human health.

DHT11 Sensor:

DHT11 is a low-cost digital temperature and humidity sensor that can be used to measure temperature and humidity in various electronic projects. It has a small size and low power consumption, making it suitable for use in portable electronic devices.

The DHT11 sensor operates by using a thermistor to measure the temperature and a capacitive humidity sensor to measure the relative humidity. The sensor is calibrated to provide a digital output. The DHT11 sensor is easy to use and can provide accurate temperature and humidity readings with a relatively low cost.

Applications: Weather stations, home automation systems, HVAC systems etc.

MQ7 Sensor:

MQ7 is a gas sensor that is designed to detect the presence of carbon monoxide (CO) gas in the air. It is based on the electrochemical principle and can detect a wide range of CO concentrations, from 20 to 2000 ppm. The sensor has a small size and low power consumption, making it suitable for use in portable electronic devices. The MQ7 sensor operates by measuring the changes in its resistance when exposed to CO gas. The sensor is calibrated to provide an analog voltage output that corresponds to the concentration of CO in the air.

Applications: Indoor air quality monitoring, gas leak detection systems, automotive exhaust monitoring etc. It is also used in industries where CO is produced or used, such as coal mines and metallurgical plants.

<u>PM2.5 Dust Smoke Particle Sensor:</u>

This is a dust and smoke particle sensor that is designed to detect and measure the concentration of fine particles, such as dust and smoke, in the air. The sensor uses an optical sensing method based on light scattering to measure the concentration of particulate matter. The sensor operates by using a laser light scattering technique to measure the concentration of fine particles in the air. The sensor measures the amount of light scattered by the particles, which is proportional to their concentration in the air. It has a compact size and low power consumption, making it suitable for use in portable electronic devices.

<u>Applications</u>: Used in air quality monitoring systems to measure the concentration of fine particles in the air. Indoor air quality monitoring, outdoor air quality monitoring, and industrial hygiene monitoring.

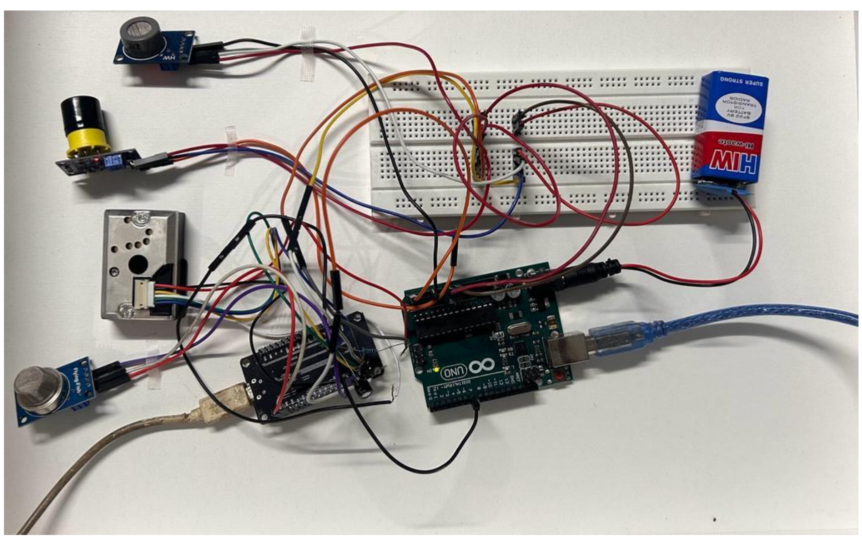

Fig 7.1.1 IoT Implementation

7.2 Data Science Implementation

The data used to train and test the Machine Learning model was a combination of the data collected from the Central Pollution Control Board (CPCB - https://cpcb.nic.in/) and our IoT sensor explained in the previous section. The data collected was dated from 25th June 2018 to 1st May 2023. The Air Pollution monitoring station is situated in Jayanagar, 5th Block, Bangalore. From the open source data made available by CPCB, the following data points were collected (Note, all data points are daily averages) :- Concentrations of Particulate Matter 2.5, Particulate Matter 10, Oxides of Nitrogen, Carbon Monoxide, Ozone as we had found these to be the major contributing pollutants during the calculations of AQI. The other data points collected were Relative Humidity, Average Temperature, Wind Direction and Barometric Pressure. A separate data point called “Diwali” was added, as it was noted that the AQI was higher on the days of Diwali.

The basic Machine Learning algorithm used for training the model was AutoRegression. Variants of this algorithm such as ARMA, ARMAX, ARIMA, ARIMAX, SARIMA were explored. Breaking down the acronyms further;

AutoRegression (AR): An AutoRegressive model is a statistical model which forecasts future values based on past values. Just like how linear models make predictions based upon a combination of independent variables, AR models make predictions based on a combination of past values. This is known as autocorrelation.

$$\hat{y}_t = c + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \ldots + \phi_p y_{t-p} + \epsilon_t$$

Fig 7.2.1 AutoRegression Formula

Integrated Differencing (I): When the data is non-stationary, this method is performed to make the data stationary. That is, prediction is not made upon the data, rather on the differences between every 2 points of data. This may be done multiple times to remove non-stationarity. However, this is just one of the many methods of making the model deal with non-stationarity. The other methods will be discussed later.

$By_t = y_{t-1}$ where B is called a backshift operator

Thus, a first order difference is written as

$$y'_t = y_t - y_{t-1} = (1 - B)y_t$$

In general, a d th-order difference can be written as

$$y'_t = (1 - B)^d y_t$$

Fig 7.2.2 Integrated Differencing Formula

Moving Average(MA): Moving average model denotes that the forecast or result of the model depends linearly on the historical data. Additionally, it implies that predicting mistakes are linear functions of prior errors. Keep in mind that statistical moving averages are not the same as moving average models.

$$\hat{y}_t = c + \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2} + \ldots + \theta_q \epsilon_{t-q}$$

Fig 7.2.3 Moving Average Formula

eXogenous Variables(X): These are external variables that may affect the time series forecast. For example, in the case of our modeling, Average Temperature, Diwali, etc. are eXogenous Variables. These alone are insufficient to forecast the values of AQI, however, they contribute indirectly.

Seasonality(S): This lets the model know if there is any seasonality in the data. Seasonality refers to known spikes or troughs in the time series data at a particular time every period. A period can be a day, week, month, quarter, year, etc.

In our case, we had daily data for about 5 years and a yearly seasonality where the AQI increased every winter. This made it hard for the model to catch the seasonality. Hence, in our case, instead of using a Seasonality model, it was better to model in Average Temperature and Relative Humidity as eXogenous Variables.

The parameters of the ARIMA model:
ARIMA(endog, exog=None, order=(p,d,q), seasonal_order=(P,D,Q,S), trend=None, enforce_stationarity=True, enforce_invertibility=True, concentrate_scale=False, trend_offset=1, dates=None, freq=None, missing='none', validate_specification=True)

By varying the parameters of this model, one can build AR, ARMA, ARMAX, ARMAX, ARIMAX or SARIMAX models. For instance, if the values of P,D,Q,S are all 0s and p=1, d=0, and q=1, and exog= Null, it becomes an ARMA model. If we pass exogenous variables, it becomes an ARMAX model.

Although ARIMA models and its variants are known to be very advantageous in modeling time series, finding the right values for all parameters requires a lot of time and good modeling expertise. There is also a chance of overfitting the training data. A quick side note on SARIMA models; while they are state of art when it comes to forecasting time series data with seasonality, they work best when the period of seasonality is low. For instance, our data consists of daily AQI values and a yearly trend or in other words, daily data and a seasonality period of 365 days. This becomes very cumbersome for the model and hence, we opted to model the seasonality by using Average Temperature as an exogenous variable.

7.3 Web Application

A web application has been developed using HTML, CSS, Javascript and PHP. It has a very user-friendly interface and an aesthetically pleasing design. It displays the current AQI, the AQI forecast for the next 5 days and the controlling measures.

CHAPTER-8

RESULTS

8.1 Model Building

MODEL (Parameters)	MSE
Predicting based on previous day	577.05
AutoReg(lags=17, trend='t')	917.961
ARMA: ARIMA(p=4, d=0, q=1, trend='t', exog=Null)	449.86
ARIMA(p=3, d=1, q=1, trend='t', exog=Null)	450.586
ARMAX(p=4,d=0,q=1, trend='t', exog=['Average Temperature'])	421.21
ARMAX(p=4,d=0,q=1, trend='t', exog=['Average Temperature','Diwali'])	320.125
ARMAX(p=4,d=0,q=1, trend='t', exog=['Relative Humidity','Wind Direction'])	376.463
ARMAX(p=4,d=0,q=1, trend='t', exog=['Average Temperature','Diwali','Relative Humidity','Wind Direction','Barometric Pressure'])	295.075
ARMAX(p=4,d=0,q=1, trend='t', exog=['Average Temperature','Diwali','Relative Humidity'])	**279.228**

Table 8.1.1 Results

8.3 Web Application

Fig 8.3.1 Homepage

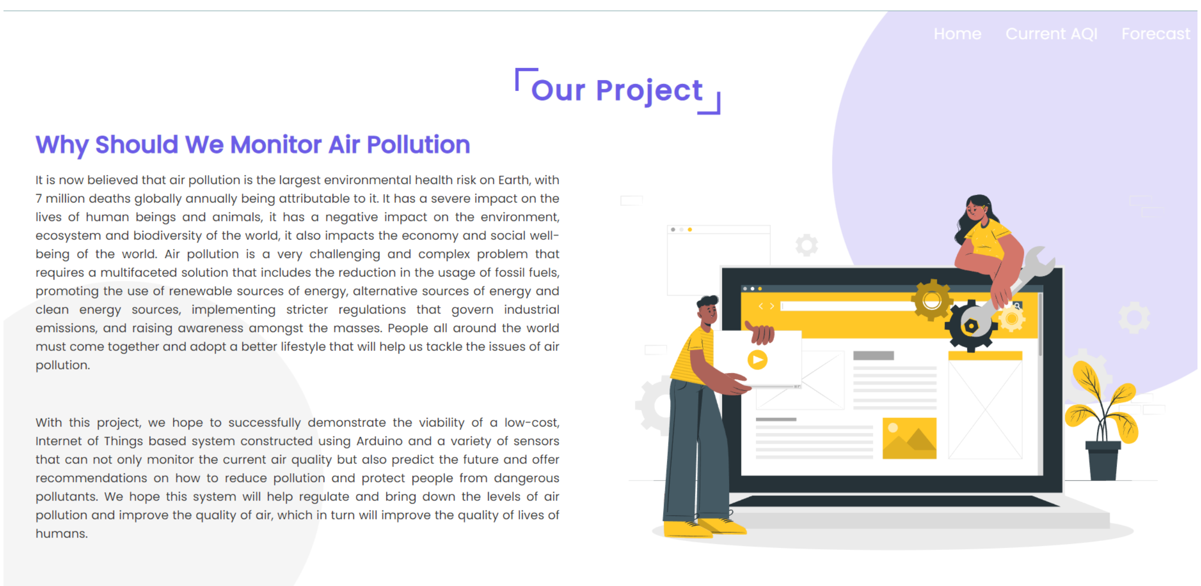

Fig 8.3.2 About our project

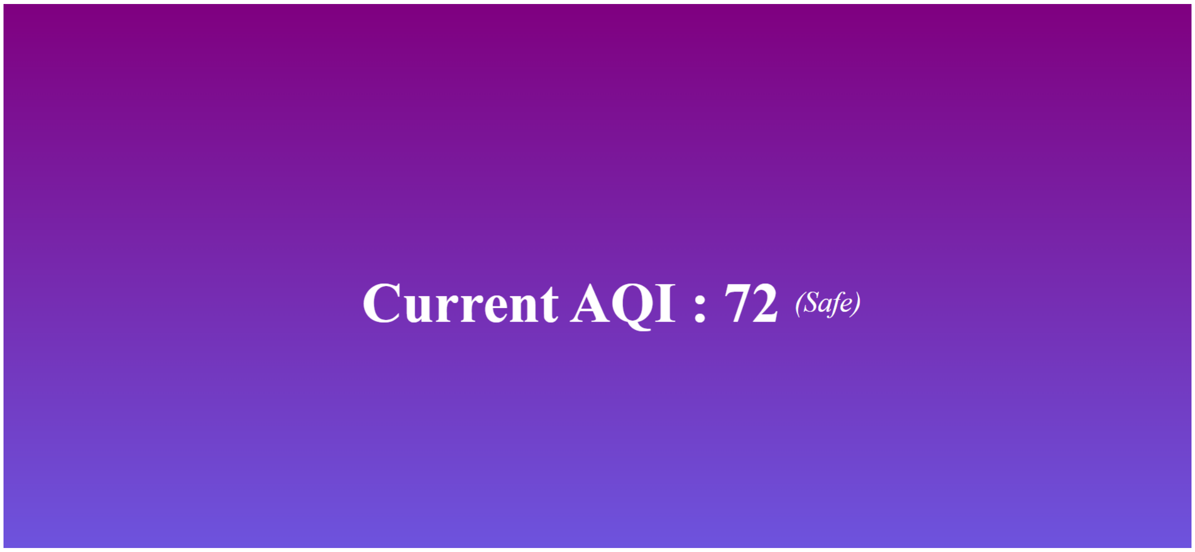

Fig 8.3.3 Current Measured AQI

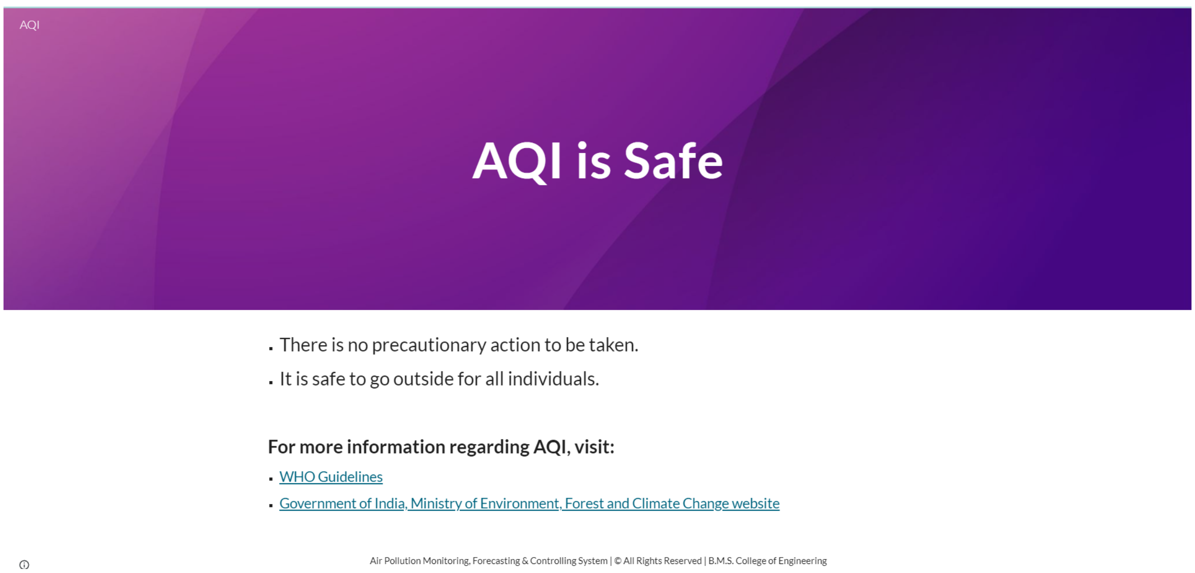

Fig 8.3.4 Guidelines for current AQI

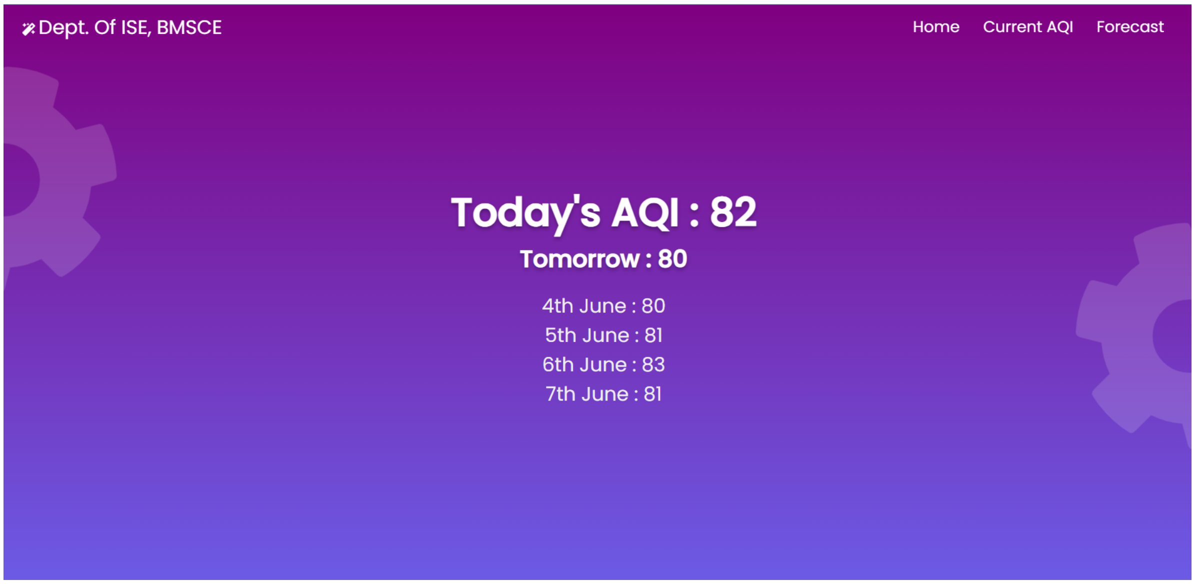

Fig 8.3.5 AQI Forecast

CHAPTER-9

TESTING

TEST CASE ID	TEST SCENARIO	TEST CASE DESCRIPTION	PRE CONDITION	TEST STEPS	TEST INPUT	EXPECTED BEHAVIOUR	ACTUAL RESULT	TEST CASE STATUS (PASS/FAIL
TC1.1	Verify the AQI against Quality of Air	To verify if the AQI range displays the correct description of Air Quality	The website must display an AQI value	1. Visit the website 2. Read the AQI 3. Check in which range AQI is present	If the AQI value is between 0-50	Description of Air Quality - Good	Description of Air Quality - Good	PASS
TC1.2	Verify the AQI against Quality of Air	To verify if the AQI range displays the correct description of Air Quality	The website must display an AQI value	1. Visit the website 2. Read the AQI 3. Check in which range AQI is present	If the AQI value is between 51-100	Description of Air Quality - Moderate	Description of Air Quality - Moderate	PASS
TC1.3	Verify the AQI against Quality of Air	To verify if the AQI range displays the correct description of Air Quality	The website must display an AQI value	1. Visit the website 2. Read the AQI 3. Check in which range AQI is present	If the AQI value is between 101-150	Description of Air Quality - Unsafe for sensitive groups	Description of Air Quality - Unsafe for sensitive groups	PASS
TC1.4	Verify the AQI against Quality of Air	To verify if the AQI range displays the correct description of Air Quality	The website must display an AQI value	1. Visit the website 2. Read the AQI 3. Check in which range AQI is present	If the AQI value is between 300 or higher	Description of Air Quality - Hazardous	Description of Air Quality - Hazardous	PASS

TEST CASE ID	TEST SCENARIO	TEST CASE DESCRIPTION	PRE CONDITION	TEST STEPS	TEST INPUT	EXPECTED BEHAVIOUR	ACTUAL RESULT	TEST CASE STATUS (PASS/FAIL
TC2.1	Verify the AQI measured and forecasted by our system with the AQI levels from the Central pollution board	To verify if the AQI measured and forecasted by the system with the AQI levels from the Central pollution Board (CPCB)	The model should be able to measure AQI through sensors data	1. Read the AQI display by the model 2. Read the AQI displayed on the CPCB website 3. Calculate the range of accepted AQI (Current AQI + or - 5)	Measured AQI = [Current AQI-5,Current AQI+5]	Measured AQI is the in range of accepted value	Measured AQI is the in range of accepted value	PASS
TC3.1	Verify that the arduino reads the data correctly	To verify if the arduino is reading the data within the expected range	The arduino must be properly connected with each sensors	1. Connect the Arduino properly with the sensors 2. Read the data displayed by the arduino	Arduino is connected to all the sensors	The data produced is within the expected range	The data produced is within the expected range	PASS
TC3.2	Verify the connection of the arduino to the database	To verify if the arduino is connected properly to the database	The arduino must be functional	1. Arduino is connected to the database 2. Database reads the data stored in the Arduino	Arduino sends the data to the database	The database properly stores the collected sensor data	The database properly stores the collected sensor data	PASS
TC3.3	Verify if the arduino is collecting the data accurately	To verify if the data stored in the database is matching the data displayed by sensors	The arduino must be properly connected with the database	1. Arduino sends the data from the sensor 2. Database is collecting the data from Arduino	Database is collecting the data from Arduino	The data stored in the database is matching the calue displayed by sensor	The data stored in the database is matching the calue displayed by sensor	PASS

These test cases include tests for the hardware system that gathers data from many sensors and stores it in a database, an ML model that gathers sensor data and forecasts the Air Quality Index (AQI), and a web application that shows the current AQI level and the predicted AQI levels.

1) Testing for the gathering of accurate sensor data demonstrates that the model accurately interprets accurate sensor data inputs, such as temperature, particulate matter (PM), humidity, and other pertinent characteristics.
2) Data storage that assesses the Arduino's capacity to connect to the database and accurately save the sensor data.
3) Test for calculation accuracy that compares the model's air quality index to the Central Pollution Board's air quality index levels.
4) The AQI range's accuracy test confirms that it gives the right representation of air quality.

These test cases check that the ML model successfully gathers sensor data, determines the current AQI, and produces accurate results under a variety of conditions.

CONCLUSION AND FUTURE ENHANCEMENTS

Air pollution is currently considered the biggest environmental health threat on the planet. It causes millions of deaths worldwide annually. Apart from causing severe issues to human health, it also poses a significant threat to animal life, the environment, ecology and biodiversity of the planet. It also impacts the economy and social well-being of the world. Air pollution is a very challenging and complex problem that requires a multifaceted solution.

To tackle this complex issue, we are proposing the implementation of a holistic approach to tackle air pollution. With this project we not only aim to reduce the levels of air pollution, but also positively impact the lives of people by helping them change their lifestyle in such a way that it will encourage them to adopt some habits that will help reduce air pollution, which will help them, their pets and their environment to be healthier.

We started this project with the aim of building a holistic solution to monitor, forecast and control the levels of AQI in Jayanagar region in Bangalore. We intended the project to be a Proof Of Concept to model the AQI in other regions of Bangalore and India similarly. We were able to build an IoT device that met our expectations, as well as model the AQI for Jayanagar to forecast the AQI for the coming few days. We have also mentioned a few controlling and precautionary measures for every bracket of the AQI.

We would like to conclude by whole heartedly thanking BMS College of Engineering and the Department of Information Science and Engineering for providing us an opportunity to work on this project. We would also like to express our heartfelt gratitude to our guide, Prof. Pallavi B for encouraging us, guiding us and being there in every step of the way. We hope to do our bit in making our planet a better place to live in by successfully implementing this project.

A few possible future enhancements are:

- Including Traffic Density and Industrial Presence as Exogenous variables as these impact the AQI the most, and will give more accurate predictions.

- Building multiple IoT devices to cater to various areas in Bengaluru and building predictive models for the same.
- Integrating the AQI forecast into Google Maps. Just like how Google gives push notifications of the weather rain forecast, it should give AQI forecasts as well.
- Automating the complete line of work, the device and model should work without any human interference, except for maintenance.

REFERENCES

[1] Ayele, Temesegan Walelign, and Rutvik Mehta. "Air pollution monitoring and prediction using IoT." 2018 second international conference on inventive communication and computational technologies (ICICCT). IEEE, 2018.

[2] Balasubramanian, Ssneha, Sneha, Talapala, B, Vinushiya and S, Saraswathi. "Air Pollution Monitoring and Prediction using IOT and Machine Learning." (IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 12 (2), 2021.

[3] Bhalgat, Pooja, Sachin Bhoite, and Sejal Pitare. "Air quality prediction using machine learning algorithms." International Journal of Computer Applications Technology and Research 8.9 (2019): 367-370.

[4] Tomar, Nimisha, Durga Patel, and Akshat Jain. "Air Quality Index Forecasting using Auto-regression Models." 2020 IEEE International Students' Conference on Electrical, Electronics and Computer Science (SCEECS). IEEE, 2020.

[5] De Vito, Saverio, et al. "Adaptive machine learning strategies for network calibration of IoT smart air quality monitoring devices." Pattern Recognition Letters 136 (2020): 264-271.

[6] Meivel S, Mahesh M, Mohnish S, Mukesh Pandian S, “Air Pollution Monitoring System Using IOT And Artificial intelligence”, IJMA, vol. 10, no. 2, pp. 3085 - 3089, May 2021.

[7] Kumar, Anikender, and P. Goyal. "Forecasting of daily air quality index in Delhi." Science of the Total Environment 409.24 (2011): 5517-5523.

[8] Zhu, Suling, et al. "Daily air quality index forecasting with hybrid models: A case in China." Environmental pollution 231 (2017): 1232-1244.

[9] Zhang, Dan, and Simon S. Woo. "Real time localized air quality monitoring and prediction through mobile and fixed IoT sensing network." IEEE Access 8 (2020): 89584-89594.

[10] Moses, Leeban. "IoT enabled Environmental Air Pollution Monitoring and Rerouting system using Machine learning algorithms." IOP Conference Series: Materials Science and Engineering. Vol. 955. No. 1. IOP Publishing, 2020.

[11] Kumar, Somansh, and Ashish Jasuja. "Air quality monitoring system based on IoT using Raspberry Pi." 2017 International conference on computing, communication and automation (ICCCA). IEEE, 2017.

[12] Mao, Wenjing, et al. "Modeling air quality prediction using a deep learning approach: Method optimization and evaluation." Sustainable Cities and Society 65 (2021): 102567.