



IJRASET

International Journal For Research in
Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 11 **Issue:** XII **Month of publication:** December 2023

DOI: <https://doi.org/10.22214/ijraset.2023.57543>

www.ijraset.com

Call:  08813907089

E-mail ID: ijraset@gmail.com

Using Mobile and Fixed Internet of Things Sensing Networks, Real-Time Monitoring and Prediction of Air Quality in One's Immediate Vicinity

Anusha Ayyagari¹, L Dasarada Ramaiah², V Anil Santosh³, D D D Suribabu⁴

¹M-Tech student, Department of Computer Science and Engineering, International School of Technology and Sciences (For Women), Andhra Pradesh, India

^{2,3}Associate Professor, Department of Computer Science and Engineering, International School of Technology and Sciences (For Women), Andhra Pradesh, India

⁴Associate Professor & Head of Department of Computer Science and Engineering, International School of Technology and Sciences (For Women), Andhra Pradesh, India

Abstract: A significant global problem has emerged in several locations as a result of air pollution and its adverse impact on human well-being. Recently, there has been a rise in the number of researchers that are keen on evaluating and forecasting the air quality in close proximity to individuals. The use of the Internet of Things (IoT) across several businesses has significantly enhanced people's quality of life by interconnecting multiple sensors in diverse locations. Moreover, the Internet of Things (IoT) has streamlined the task of monitoring air pollution. The conventional utilization of stationary sensors is inadequate for acquiring an accurate and all-encompassing representation of the air pollution levels in close proximity to people. This is due to the fact that the sensors in closest proximity to people may be located many kilometers apart from each other. The objective of our study is to construct a model that precisely depicts the air quality pattern within a certain geographic area. This objective will be accomplished by using a combination of stationary and portable Internet of Things sensors. The sensors are affixed to vehicles that are carrying out surveillance in the vicinity. Our methodology allows for a thorough examination of the whole spectrum of air quality fluctuations in neighboring regions. Through the use of many machine learning algorithms and real-world data, we showcase the efficacy of our methodology in accurately discerning and forecasting air quality without compromising on accuracy. The results of our research indicate that there is significant potential for efficiently monitoring and forecasting air quality in the context of a smart city.

Keywords: Internet of things, Real time monitoring, Air quality, Networks

I. INTRODUCTION

A Many nations throughout the globe are now facing a substantial issue with air pollution, which has emerged as a direct result of widespread urbanization and industrialization. In recent years, there has been an increasing acknowledgment of the significant impact of air pollution on the daily routines of individuals. In urban areas of developing nations affected by air pollution, people are sometimes required to wear a mask prior to engaging in outside activities. Two urban areas that serve as prime examples of this phenomenon are Beijing and Delhi. Furthermore, the air quality throughout the day poses an additional limitation on outdoor activities. The presence of several air contaminants may be attributed to the occurrence of air pollution. The main atmospheric pollutants include nitrogen dioxide (NO₂), carbon monoxide (CO), ozone (O₃), and sulfur dioxide (SO₂). The term "air particulate matter" (PM) refers to an additional categorization of air contaminants. The designation PM_{2.5} is often used to refer to atmospheric particulate matter with a diameter less than 2.5 μm, whereas PM₁₀ is used to signify particulate matter with a diameter smaller than 10 μm. Human beings exhibit significant anxiety over the presence of these two distinct types of particle pollution. These particles possess the capacity to elicit many respiratory and cardiovascular disorders. As a result, several local governments have implemented their own air quality monitoring stations and provide up-to-date information on the current air quality on an hourly basis.

The importance of monitoring air quality in close proximity to individuals has escalated in parallel with the rising concern around air pollution. These metrics provide people with information on the most favorable circumstances for engaging in outdoor activities and assist them in formulating more efficient plans for reaching their intended destinations. The conventional method used to monitor atmospheric parameters over a wide geographical expanse involves the establishment of monitoring stations at predetermined positions. Although the implementation of a fixed sensor-based monitoring system may seem straightforward, there are certain obstacles that need to be overcome. The initial establishment and deployment of surveillance equipment over a wide geographical region require a significant financial commitment. Additionally, the proximity of the environment has a significant impact on the system's performance, and it typically displays less accuracy when used in remote locations. When evaluating the air quality in regions next to roads, even small distances might significantly impact the data related to traffic pollution. Therefore, there is a desire for innovative approaches to getting air quality data that are both cost-effective and readily sustainable while also allowing precise predictions about air quality. One potential approach that may be used to tackle these difficulties is facilitating the movement of sensors via the deployment of Internet-of-Things (IoT) technology. For example, the viability of attaching sensors to mobile vehicles or unmanned aerial vehicles (UAVs) has been shown as a feasible approach. The scope of our research project included the conceptualization and execution of Internet of Things (IoT) sensors that were specially customized for the objective of monitoring air quality. A sensor was installed on a vehicle to gather data on air pollution levels throughout its journey through the urban environment of Incheon, a city located inside the Republic of Korea. Afterwards, the data is subjected to preparation before being sent to our server for storage. One of the key benefits attributed to the use of a mobile sensor is its capacity to provide instantaneous data on the level of air pollution inside a certain geographic location at a particular moment in time, specifically while the vehicle was moving through such an area. The use of mobile Internet of Things (IoT) sensors enables the expansion of geographical coverage and the acquisition of more accurate data at the local level. The use of a mobile sensor has difficulties in ensuring a constant dissemination of information related to a specific region, as opposed to a stationary fixed sensor, which is capable of doing this. Nevertheless, it is important to acknowledge that a mobile sensor does have the capacity to circumvent this constraint.

A. Objectives

Air pollution and the negative effects it has on human health have emerged as a significant issue in a number of locations throughout the globe. Over the last several years, there has been a growing interest among academics in the investigation and prediction of the air quality in the areas around human populations. Through the facilitation of the connectivity of multiple sensors located in various geographical locations, the Internet of Things (IoT) has made a substantial contribution to the improvement of the quality of life of residents. This is due to the widespread application of the IoT in a variety of disciplines. Consequently, it has also simplified the process of monitoring air pollution, making it more effective in comparison to the approaches that were previously used. When it comes to correctly and completely determining the levels of air pollution in the vicinity of people, the typical use of stationary sensors is not adequate. It's possible that this occurrence may be explained by the significant distance that exists between people and the sensors that are in the closest vicinity to them, which spans several kilometers. The overall objective of our research is to create a model that can be used to analyze the pattern of air quality in a certain geographical location. Through the use of a mix of stationary and mobile sensors that are linked via Internet of Things (IoT) technologies, this objective will be accomplished. The aforementioned sensors are fitted in a strategic manner on the vehicles that are deployed for the purpose of monitoring inside the defined region. Through the use of the approach that we have provided, it is now possible to conduct an exhaustive investigation of the whole spectrum of changes in air quality that exist across neighboring regions. We illustrate the usefulness of our strategy by applying a number of different machine learning algorithms and making use of data from the actual world. This allows us to accurately detect and forecast air quality while retaining a high degree of accuracy. The results of our research indicate that there is sufficient potential for the efficient monitoring and forecasting of air quality within the framework of the application of smart city technology.

II. LITERATURE REVIEW

Hazardous chemicals may be released into the environment as a result of a wide variety of activities, both natural and human, that may be responsible for their introduction. These substances have the potential to have negative impacts not just on the health of humans but also on the biodiversity of the environment. Because of human activities, there has been an increase in the burning of fossil fuels, which may be connected to the observed changes in the composition of the atmosphere that have occurred over the course of the previous century.

There are many different types of air pollutants, including carbon monoxide (CO), sulfur dioxide (SO₂), nitrogen oxides (NO_x), volatile organic compounds (VOCs), ozone (O₃), heavy metals, and respirable particulate matter (PM_{2.5} and PM₁₀), all of which are examples of air contaminants. In terms of their chemical composition, reactivity characteristics, emission patterns, disintegration rates, and potential for dispersion across a range of distances, the different pollutants are distinguished from one another. The impacts of air pollution on human health may be classified as either acute or chronic, and they have the potential to have an impact on a wide variety of organs and systems located inside the body. There is a wide range of potential health effects that are associated with this condition. These effects can range from a mild irritation of the upper respiratory tract to the development of chronic respiratory and cardiovascular diseases. Some examples of these diseases include lung cancer, acute respiratory infections in children, chronic bronchitis in adults, the worsening of pre-existing heart and lung conditions, and the occurrence of asthmatic episodes. In addition, it has been shown that exposure, whether it be for a short amount of time or for an extended length of time, is associated with an increased risk of early death as well as a decrease in the person's total life expectancy. The purpose of this research is to investigate the effects that air pollutants have on human health in a comprehensive manner, as well as to shed light on the mechanisms by which these pollutants have an impact.

Apheis: Health impact assessment of long-term exposure to PM_{2.5} in 23 European cities

Increasing the availability of information on air pollution (AP) and public health (PH) to European decision-makers, environmental health professionals, and the general public is the goal of the Advanced Public Health Information System (APHEIS). The goal of this endeavor is to provide data that is both up-to-date and easy to understand about these two topics. During the Apheis-3 phase, an evaluation was conducted to determine the effect that extended exposure to PM_{2.5} (particulate matter with a diameter less than 2.5 micrometers) had on the pH levels of 23 cities located throughout Europe. Quantification of the mortality effect linked with PM_{2.5} and investigation of the possible effect that it may have on life expectancy were both included in the research. It was determined that the Health Impact Assessment (HIA) should be carried out in accordance with the recognized techniques that were specified by the World Health Organization (WHO), as well as the specifications that Apheis had established for the gathering of data and analysis. In order to assess the potential increase in life expectancy, we used the PSAS-9 method for estimating attributable events as well as the AirQ software created by the World Health Organization (WHO). The values of PM_{2.5} for the majority of cities were calculated by using a conversion factor, which might be either local or European, depending on the PM₁₀ data that was provided. A decrease in long-term exposure to PM_{2.5} levels in each city to 15 microg/m³ might possibly result in the avoidance of 16,926 early deaths across all causes on an annual basis, according to the findings included within the Health Impact Assessment (HIA). There are a total of 11,612 deaths that have been ascribed to cardiac reasons, and there are also 1901 deaths that have been explicitly related to lung cancer. Each of these deaths has been documented. The decrease in death rates that was found across the cities that were a part of the Apheis project would result in a comparable increase in life expectancy at the age of thirty, with differences ranging from one month to almost two years. The Health Effect Assessment (HEA) that was carried out as part of our research investigated the potential influence that extended exposure to fine particles may have on the anticipated lifespan of an individual. After doing this analysis, a more in-depth comprehension of the impact that air pollution has on the health of the general population in Europe has been achieved. In addition to the number of incidents that may be attributed to the pollution, this information is also included.

III. PROPOSED METHODOLOGY

A deployment diagram is used inside the Unified Modeling Language (UML) to illustrate the depiction of the arrangement of physical artifacts on nodes. This is done in order to accomplish the aforementioned purpose. In the context of the description of a website, for instance, a deployment diagram may be used as an example of a scenario. This diagram provides a graphical representation of the hardware components, often known as "nodes," that are involved in the functioning of the website. There is a possibility that these nodes will consist of a database server, an application server, and a web server. In addition, the figure depicts the software components, which are referred to as "artifacts," that are carried out on every node. An example of these artifacts would be a web application and a database, both of which are associated with a website. In addition, the figure illustrates the connections that exist between the different components, including REST, RMI, and JDBC-based components.

The node boxes themselves are square shapes, but the rectangular forms inside them serve as representations of the allocated artifacts. There are sub-nodes that may be associated with a node. These sub-nodes are graphically represented by nested boxes that are linked to the overarching node. In a deployment diagram, it is feasible for a single node to symbolically represent numerous physical nodes. For instance, a single node might in this way represent a cluster of database servers. When it comes to deployment diagrams, the graphic that has been supplied acts as an example.

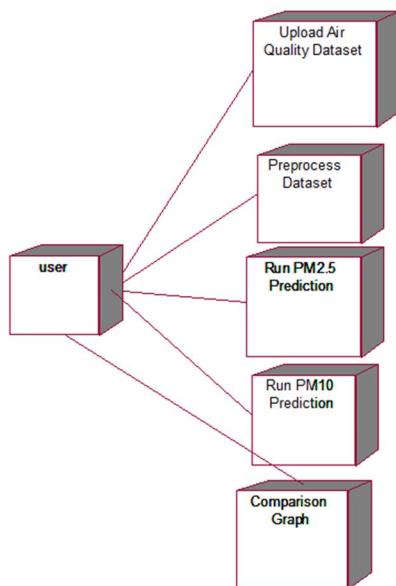


Fig 1. Deployment diagram

A. Activity Diagram

Within the Unified Modeling Language (UML), the activity diagram is an important diagram that works to describe the dynamic features of a system. Its purpose is to simplify the process of modeling. The diagram that is being shown is, in essence, a flow chart that graphically depicts the progression from one activity to subsequent actions in a sequential order. A direct correlation may be drawn between the activity being described and the functioning of the system. As a result, the control flow is moved from one operation to another in a simultaneous manner. There are three main modes that may be used to describe this flow: sequential, branching, or concurrent.

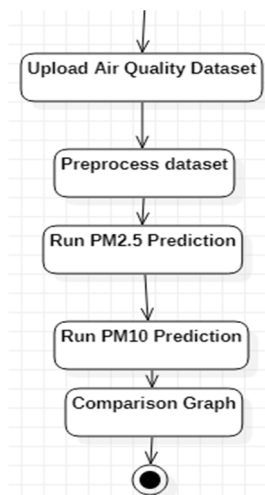


Fig 2. Activity diagram

B. Data Flow Diagram

The method by which a system processes data is shown by data flow diagrams, which also outline the inputs and outputs that the system receives. It is possible to employ data flow diagrams, often known as DFDs, which are a powerful tool that may be used to create a clear and complete depiction of a variety of business processes. The first step in the technique is to get a comprehensive view of the business, and the second is to look into the various functional areas that are important within it. The execution of the research has the capability of accurately obtaining the needed amount of information. To ensure that a comprehensive investigation is carried out, the technique makes use of the top-down growth development process.

A graphical depiction that demonstrates the flow of data inside a system or process is referred to as a data flow diagram, or DFD for short. This specific figure's name does a good job of describing the function that it is supposed to do. It is possible to generate a data flow diagram (DFD) by using simple symbols without resorting to excessive complexity. To make matters worse, the use of diagramming tools that are not only user-friendly but also freely available for download enables the rapid automation of complex processes. A data flow diagram, often known as a DFD, is a conceptual model that is used for the primary goal of developing and accessing information activities. When attempting to demonstrate the method by which information moves through a certain process, differential flow diagrams (DFD) are used as a technique for doing so. These diagrams take into consideration the many inputs and outputs that are involved. When referring to the structure of a data flow diagram (DFD), the word "process model" may also be used to refer to the structure. It is a graphical representation that illustrates a commercial or technological process, containing both incoming and outgoing data as well as the outputs that are produced as a consequence of the process. This kind of representation is known as a Data Flow Diagram (DFD).

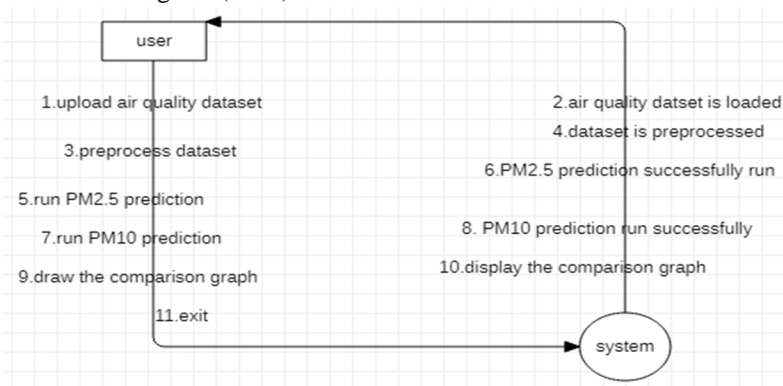


Fig 3. Data flow diagram

IV. EXPERIMENT ANALYSIS

A. This Project Implementation and Testing

Due to the fact that it includes a number of activities that are really interesting, the implementation phase of a project is of great relevance. As a consequence of this, it is essential for people to exercise caution during this time period. One of the most important stages in the process of developing a successful system is the implementation stage. This stage is essential because it is responsible for instilling faith in users about the efficiency and practicability of the new system. In addition, this stage guarantees that the system will be successful. Independent testing is performed on each individual program throughout the development process. This testing is done using sample data to check that the programs are able to interface with one another in accordance with the program specification. The fact that this observation was made implies that the definition of the program is accurate and trustworthy. The computer system and the environment in which it operates are subjected to rigorous testing in order to guarantee that the user will be satisfied with the results.

B. Taking action in Response to the Conditions that have been made Available

When compared to the phase that encompasses system design, the phase that encompasses system implementation demonstrates a degree of originality that is considerably lower. Specifically, with regard to this specific system, the user training process and the file conversion process are the most important elements to consider. The potential arises that the system could need a significant amount of user training in order to function properly. Following the execution of the programming activity, it is necessary to make adjustments to the essential settings of the system. These modifications are an absolute necessity. This article describes a technique that attempts to assist the user's comprehension and execution of different actions by offering a user-friendly approach that is defined by transparency and efficiency. The methodology is presented in this document. The user is given the option of using either the inkjet printer or the dot matrix printer, both of which are available to them. Both of these printer varieties are available to them. Because of this capability, the user is able to utilize a large variety of reports that are available to them.

The implementation of the technique that has been supplied is a rather straightforward process. In its broadest sense, the word "implementation" refers to the process of bringing a system that has been recently designed or updated into a state where it can perform its intended functions.

C. An Investigation is being Carried Out

In the process of testing, the production of test data is included. This data is then used to do independent testing on the modules, and after that, the fields are validated. Immediately after the conclusion of the testing procedure, this method is carried out. Following that, the process of system testing is carried out in order to guarantee that all of the system components carry out their functions in a cohesive manner as a unified organism. In order for the test to provide accurate findings, it is essential to choose the data that will be used in a manner that enables its applicability under any and all circumstances that might possibly arise. The testing phase of implementation is carried out before the real operation begins. The purpose of this phase is to identify whether or not the system is performing in an appropriate and efficient manner. Before beginning the actual surgical procedure, this process is carried out in order to avoid any complications. In the following paragraphs, a comprehensive description of the testing procedures that were carried out throughout the testing period is presented.

D. Evaluation of the System is Currently Ongoing at this Time

Over the course of the last several years, testing has evolved into an essential component in the process of putting many different systems and projects into development, especially in the field of information technology. For the purpose of determining whether or not a person is prepared to manage the numerous facets of their life, including their capacity to confront and triumph over challenges within certain situations, tests are used. The relevance of doing testing prior to the production of anything is of the utmost importance since it highlights the intrinsic worth and purpose of testing itself. It is essential to run tests on software once it has been developed in order to determine whether or not it is capable of accomplishing the goal that was intended for it. Before the program is made available to consumers, it is of the utmost necessity to carry out this testing. It is possible to use a wide variety of testing approaches in order to guarantee the dependability of the software. A logical examination was carried out on the program, during which the execution behavior of the program was examined several times for a certain data input. As a consequence of this, the code was subjected to careful examination in order to guarantee that it included correct information, and the outcomes that were produced were also exposed to close examination.

E. Examining the Modules in Question

When doing specialized testing on each module, the main goal is to discover and identify any possible issues that may be present. This ensures that the module is functioning properly. As a consequence of this, we are able to recognize flaws and make improvements to them without having any effect on other modules. The process of making modifications to software is needed in order to get the intended result in situations where the program is unable to perform its primary function. In light of the conditions, it is essential to expose each module to its own unique testing processes, beginning with the modules that are the least complicated and easiest to understand and gradually progressing to the next level. Extensive testing is performed on each and every component of the system in a way that is unique to that component. As an example of a module that is subjected to autonomous testing, the module that is responsible for job classification fits this description. In order to conduct a comprehensive analysis of the capabilities of this module, it is put through testing using a number of different tasks, all of which have needs for around the same amount of time. After that, a comparison is made between the outcomes of the test and the results that were obtained manually throughout the process. The comparison study reveals, on the basis of the data that was acquired, that the strategy that was recommended is more efficient in terms of its consequences when compared to the methodology that is currently being used. Extensive testing is performed on each and every component of the system in a way that is unique to that component. Within the framework of this specific system, the modules that are accountable for the classification of resources and the scheduling of work are subjected to separate testing, which results in findings that correspond to each independent test. This helps to reduce the amount of time that is spent waiting for the next phase to be finished, which in turn helps to reduce the amount of time that is spent without doing anything throughout the process.

F. The Evaluation of the process Regarding Integration

Following the conclusion of the module testing phase, the second step entails the execution of integration testing. There is a possibility that errors will occur when the process of integrating modules is carried out. The adoption of testing processes is what is used to remedy these errors. All of the parts of this system are linked to one another and are put through extensive testing for their reliability. Each and every one of the test findings is completely accurate. As a consequence of this, the system displays the capacity to properly assign workloads to resources when the resources are configured in a suitable manner.

To run project double click on 'run.bat' file to get below screen

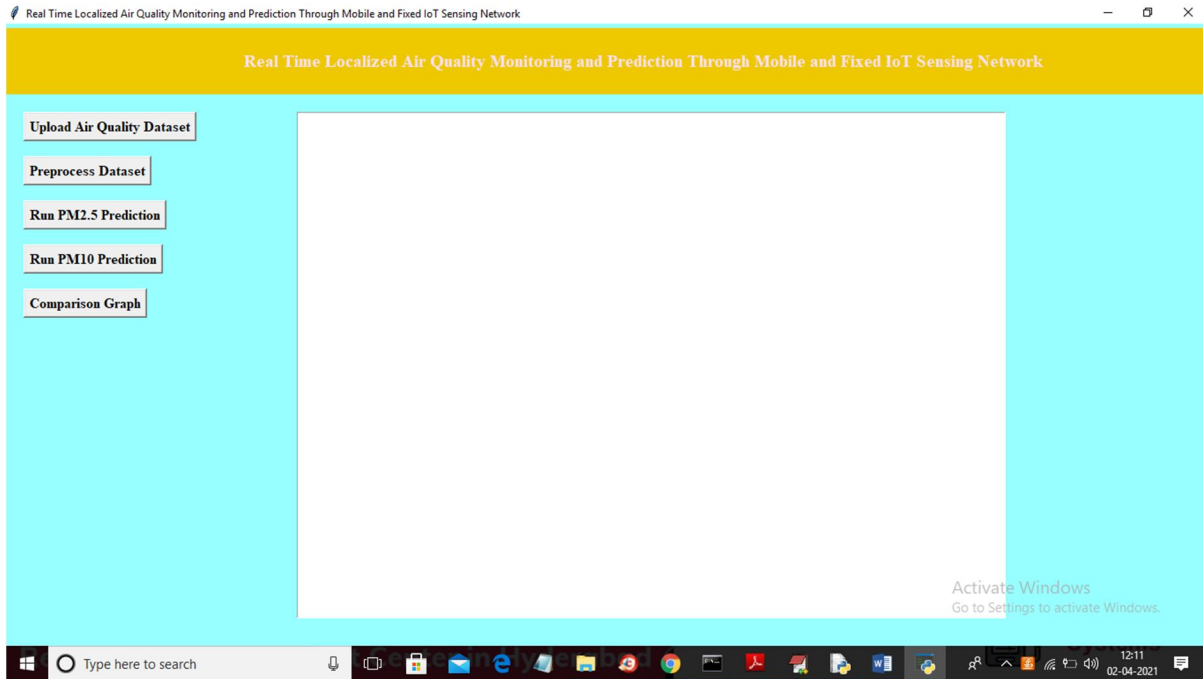


Fig 4. User interface

In above screen click on 'Upload Air Quality Dataset' button and upload dataset

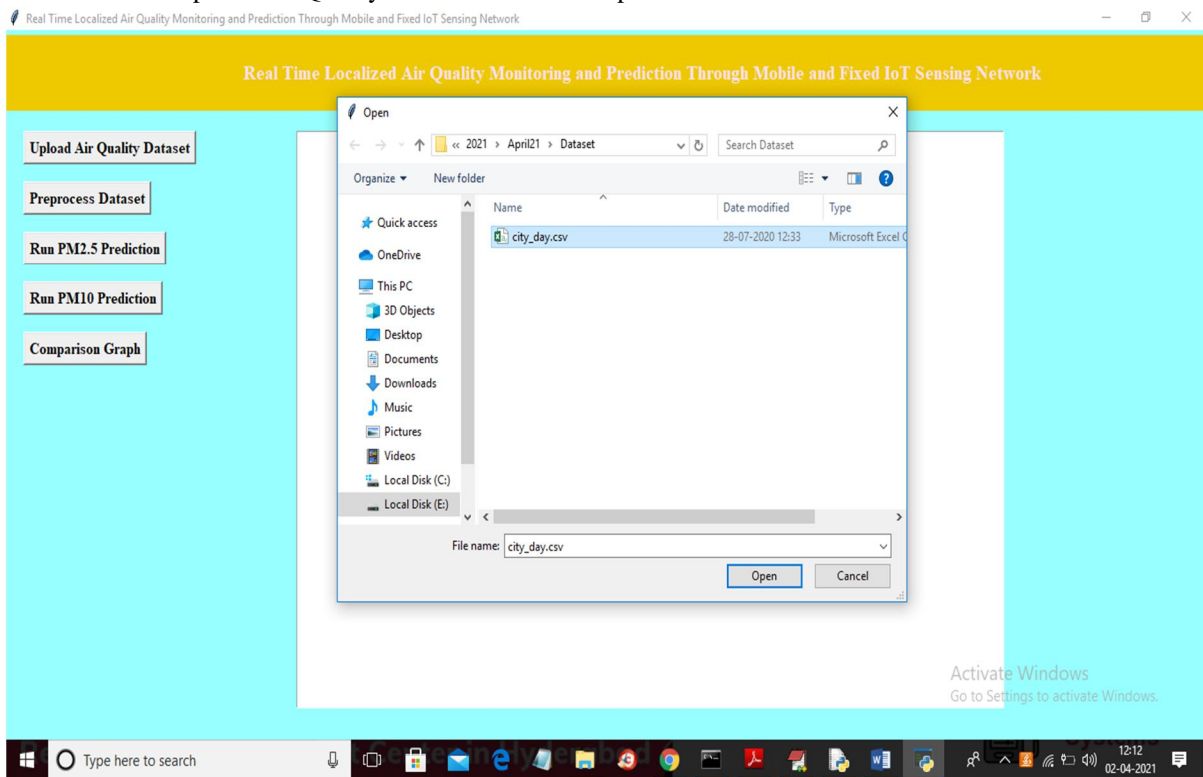


Fig 5. Upload data

The user should continue by clicking the 'Open' button in order to access the following page and load the dataset. This should be done after the user has chosen and uploaded the 'city_day.csv' file on the screen that was previously described.

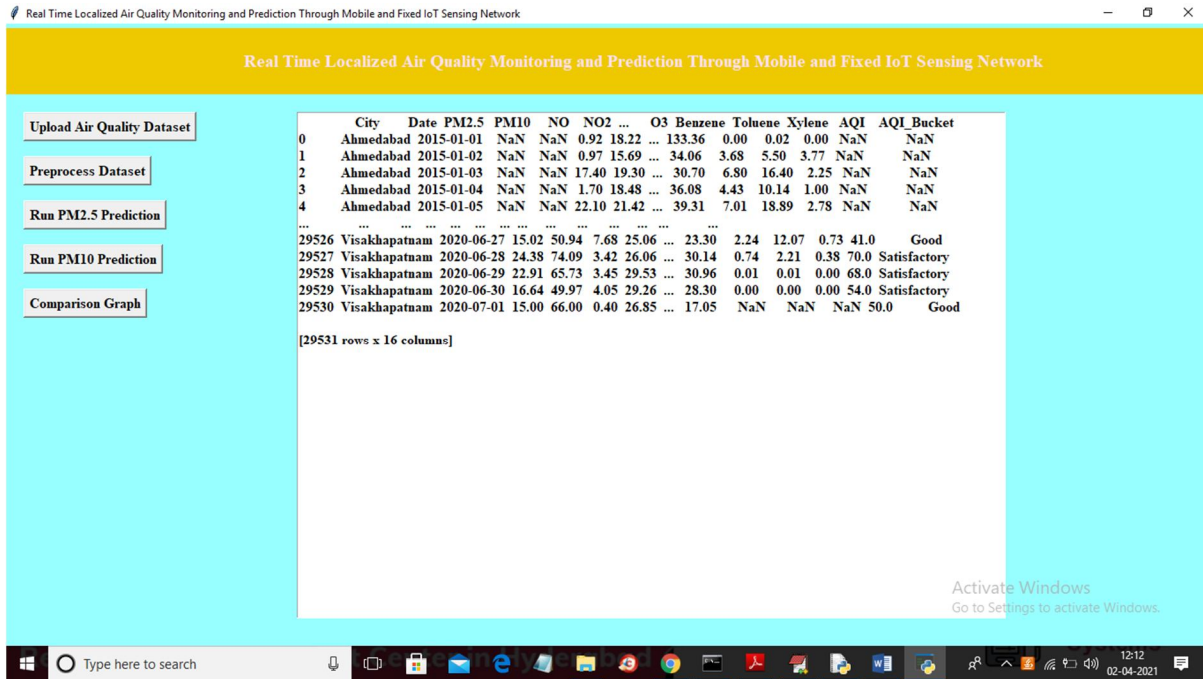


Fig 6. Analyzing data

It is clear that the data that has been presented has a number of values that are absent, which are referred to as NAN values. In light of this, it is essential to use a preprocessing method in order to replace these NAN values with numerical values. Clicking the "Preprocess Dataset" button will allow you to get the result that is shown.

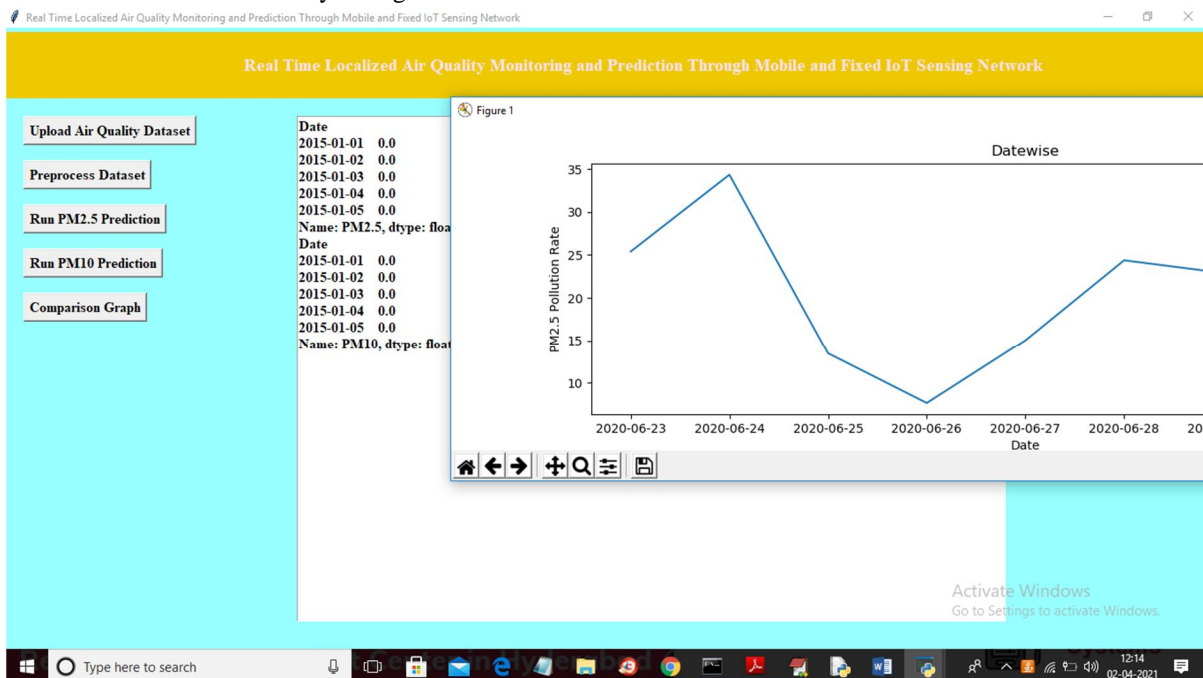


Fig 7. Developing graphs

It is clear from the information that has been presented that the values that were absent, which are represented by the symbol NAN, have been replaced with values of 0. Additionally, the graph that is seen above depicts the levels of PM2.5 that have been present throughout the course of the last ten days. The date is shown along the x-axis, and the percentage of PM2.5 that corresponds to that date is shown along the y-axis.

It would be helpful if you could dismiss the graph that is now shown and then continue by choosing the "Run PM2.5 Prediction" option in order to apply the newly provided dataset to all four strategies. After that, a prediction for the PM2.5 levels for the future thirty-day period has to be prepared, and then the computation of the root mean square error ought to be carried out. Within the ensuing console, one will have the ability to view the projected, observed, and anticipated values for each algorithm during the impending thirty-day period.

```

C:\Windows\system32\cmd.exe
Gradient Boosting PM2.5 Day=1, Predicted=17.582784, Expected=9.910000
Gradient Boosting PM2.5 Day=2, Predicted=11.812784, Expected=17.940000
Gradient Boosting PM2.5 Day=3, Predicted=18.818562, Expected=19.140000
Gradient Boosting PM2.5 Day=4, Predicted=20.118393, Expected=29.710000
Gradient Boosting PM2.5 Day=5, Predicted=29.868076, Expected=21.660000
Gradient Boosting PM2.5 Day=6, Predicted=23.562784, Expected=29.690000
Gradient Boosting PM2.5 Day=7, Predicted=30.568562, Expected=31.140000
Gradient Boosting PM2.5 Day=8, Predicted=32.118393, Expected=31.840000
Gradient Boosting PM2.5 Day=9, Predicted=33.120212, Expected=32.550000
Gradient Boosting PM2.5 Day=10, Predicted=33.830212, Expected=37.710000
Gradient Boosting PM2.5 Day=11, Predicted=38.688393, Expected=37.670000
Gradient Boosting PM2.5 Day=12, Predicted=39.232347, Expected=22.750000
Gradient Boosting PM2.5 Day=13, Predicted=24.652784, Expected=21.330000
Gradient Boosting PM2.5 Day=14, Predicted=22.892347, Expected=24.660000
Gradient Boosting PM2.5 Day=15, Predicted=25.638393, Expected=21.060000
Gradient Boosting PM2.5 Day=16, Predicted=22.622347, Expected=17.050000
Gradient Boosting PM2.5 Day=17, Predicted=18.612347, Expected=21.160000
Gradient Boosting PM2.5 Day=18, Predicted=22.138393, Expected=14.410000
Gradient Boosting PM2.5 Day=19, Predicted=16.312784, Expected=26.670000
Gradient Boosting PM2.5 Day=20, Predicted=26.828076, Expected=25.960000
Gradient Boosting PM2.5 Day=21, Predicted=27.522347, Expected=33.170000
Gradient Boosting PM2.5 Day=22, Predicted=34.148393, Expected=25.400000
Gradient Boosting PM2.5 Day=23, Predicted=27.302784, Expected=34.360000
Gradient Boosting PM2.5 Day=24, Predicted=35.238562, Expected=13.450000
Gradient Boosting PM2.5 Day=25, Predicted=15.352784, Expected=7.630000
Gradient Boosting PM2.5 Day=26, Predicted=9.532784, Expected=15.020000
Gradient Boosting PM2.5 Day=27, Predicted=15.998393, Expected=24.380000
Gradient Boosting PM2.5 Day=28, Predicted=25.258562, Expected=22.910000
Gradient Boosting PM2.5 Day=29, Predicted=24.472347, Expected=16.640000
Gradient Boosting PM2.5 Day=30, Predicted=18.542784, Expected=15.000000

SVR PM2.5 Day=1, Predicted=34.703826, Expected=9.910000
SVR PM2.5 Day=2, Predicted=28.750578, Expected=17.940000
SVR PM2.5 Day=3, Predicted=35.569267, Expected=19.140000
SVR PM2.5 Day=4, Predicted=40.335541, Expected=29.710000
SVR PM2.5 Day=5, Predicted=45.093799, Expected=21.660000
SVR PM2.5 Day=6, Predicted=38.341614, Expected=29.690000
SVR PM2.5 Day=7, Predicted=47.319267, Expected=31.140000
SVR PM2.5 Day=8, Predicted=52.296156, Expected=31.840000
SVR PM2.5 Day=9, Predicted=53.089658, Expected=32.550000
SVR PM2.5 Day=10, Predicted=53.798901, Expected=37.710000
SVR PM2.5 Day=11, Predicted=57.398040, Expected=37.670000
SVR PM2.5 Day=12, Predicted=58.938196, Expected=22.750000
  
```

Fig 8. Analyzing the data

```

C:\Windows\system32\cmd.exe
Random Forest Regressor PM2.5 Day=9, Predicted=27.143355, Expected=32.550000
Random Forest Regressor PM2.5 Day=10, Predicted=36.953072, Expected=37.710000
Random Forest Regressor PM2.5 Day=11, Predicted=30.041920, Expected=37.670000
Random Forest Regressor PM2.5 Day=12, Predicted=41.054697, Expected=22.750000
Random Forest Regressor PM2.5 Day=13, Predicted=21.764043, Expected=21.330000
Random Forest Regressor PM2.5 Day=14, Predicted=22.942720, Expected=24.660000
Random Forest Regressor PM2.5 Day=15, Predicted=27.783953, Expected=21.060000
Random Forest Regressor PM2.5 Day=16, Predicted=13.542612, Expected=17.050000
Random Forest Regressor PM2.5 Day=17, Predicted=22.685342, Expected=21.160000
Random Forest Regressor PM2.5 Day=18, Predicted=30.783409, Expected=14.410000
Random Forest Regressor PM2.5 Day=19, Predicted=22.574738, Expected=26.670000
Random Forest Regressor PM2.5 Day=20, Predicted=21.521845, Expected=25.960000
Random Forest Regressor PM2.5 Day=21, Predicted=21.750246, Expected=33.170000
Random Forest Regressor PM2.5 Day=22, Predicted=42.397767, Expected=25.400000
Random Forest Regressor PM2.5 Day=23, Predicted=25.718074, Expected=34.360000
Random Forest Regressor PM2.5 Day=24, Predicted=26.604781, Expected=13.450000
Random Forest Regressor PM2.5 Day=25, Predicted=17.062057, Expected=7.630000
Random Forest Regressor PM2.5 Day=26, Predicted=4.217415, Expected=15.020000
Random Forest Regressor PM2.5 Day=27, Predicted=16.585139, Expected=24.380000
Random Forest Regressor PM2.5 Day=28, Predicted=52.723413, Expected=22.910000
Random Forest Regressor PM2.5 Day=29, Predicted=12.399555, Expected=16.640000
Random Forest Regressor PM2.5 Day=30, Predicted=26.828327, Expected=15.000000

Model: "sequential_1"
Layer (type)                Output Shape                Param #
-----
lstm_1 (LSTM)                (1, 4)                      96
dense_1 (Dense)              (1, 1)                       5
-----
Total params: 101
Trainable params: 101
Non-trainable params: 0

None
WARNING:tensorflow:From C:\Users\Admin\AppData\Local\Programs\Python\Python37\lib\site-packages\keras\backend\tensorflow_backend.py:422: The name tf.global_variables is deprecated. Please use tf.compat.v1.global_variables instead.

Epoch 1/1
  
```

Fig 9. Checking outputs

After the random forest technique has been finished, the LSTM algorithm will begin the process of constructing a prediction model for PM2.5 levels, as can be seen on the screen that was previously indicated.

Figure 1

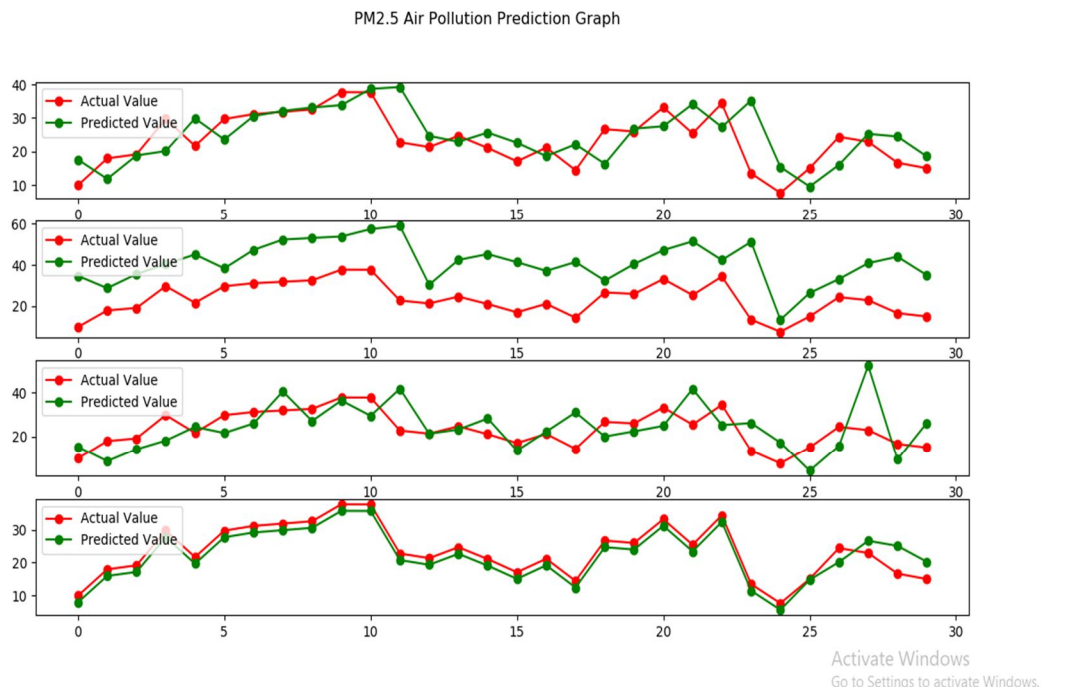


Fig 10. Air pollution prediction graph

According to the graph that is shown above, the red line represents the numbers that have been seen or are genuine, while the green line represents the values that have been projected or estimated. On top of that, it is important to point out that the horizontal axis of the graph that was just discussed relates to the duration in days, while the vertical axis represents the projected values. The initial graph that is shown in the picture that is located above is a representation of the outcomes that were achieved via the use of methods such as gradient boosting, simultaneous vector regression, random forest, and LSTM. The graph that is shown above demonstrates that the Long Short-Term Memory (LSTM) model properly predicted values that were in close alignment with the values that were actually seen via observation. In order to go to the next screen, it is strongly suggested that you disregard the graph that is now shown.

V. CONCLUSIONS & FUTURE SCOPE

A comprehensive analysis of the fundamental concepts behind the Long Short-Term Memory Network (LSTM) concept, as well as its sequential design, has been presented in this article. The development of an LSTM model that is easier to grasp and more straightforward may be accomplished by first getting a full knowledge of its operation. The broad applicability of LSTM models in the field of artificial intelligence for a variety of natural language processing applications, including language modeling and machine translation, makes it an important issue that should be discussed. LSTM has been used in a variety of fields, including voice recognition, picture captioning, handwriting identification, time series forecasting via the collection of time series data, and a great number of other applications.

During the course of this research, we investigated a novel approach to predicting the immediate air quality around people by using a combination of stationary and mobile sensors. The results of the experiments provide evidence that the hybrid distributed fixed and Internet of Things sensor system that has been built has the power to accurately anticipate air quality in areas that are in close proximity to human populations. Furthermore, our suggested system has the potential to be implemented in an effective manner via the use of public transportation systems, such as cabs and buses, which would be outfitted with Internet of Things (IoT) sensor devices in order to collect data related to a variety of geographical locations. The data that our system provides regarding the anticipated air quality has the potential to be helpful in a variety of contexts, particularly when it comes to the planning of outdoor activities.

REFERENCES

- [1] Beijing'S Air Would be Step up for Smoggy Delhi. Accessed: Jan. 26, 2014. [Online]. Available: <http://https://www.nytimes.com/2014/01/26/world/asia/beijings-air-would%20be-step-up-for-smoggy-delhi.html>
- [2] M. Kampa and E. Castanas, "Human health effects of air pollution," *Environ. Pollut.*, vol. 151, no. 2, pp. 362_367, Jan. 2008.
- [3] E. Boldo, S. Medina, A. Le Tertre, F. Hurley, H.-G. Mücke, F. Ballester, and I. Aguilera, "Aphis: Health impact assessment of long-term exposure to PM2.5 in 23 European cities," *Eur. J. Epidemiology*, vol. 21, no. 6, pp. 449_458, Jun. 2006.
- [4] J. Lin, A. Zhang, W. Chen, and M. Lin, "Estimates of daily PM2.5 exposure in beijing using spatio-temporal kriging model," *Sustainability*, vol. 10, no. 8, p. 2772, 2018.
- [5] Y. Jiang, L. Shang, K. Li, L. Tian, R. Piedrahita, X. Yun, O. Mansata, Q. Lv, R. P. Dick, and M. Hannigan, "MAQS: A personalized mobile sensing system for indoor air quality monitoring," in *Proc. 13th Int. Conf. Ubiquitous Comput. UbiComp*, 2011, pp. 271_280.
- [6] D. Zhang and S. S. Woo, "Predicting air quality using moving sensors (poster)," in *Proc. 17th Annu. Int. Conf. Mobile Syst., Appl., Services*, Jun. 2019, pp. 604_605.
- [7] Y. Zheng, X. Yi, M. Li, R. Li, Z. Shan, E. Chang, and T. Li, "Forecasting_ne-grained air quality based on big data," in *Proc. 21th ACM SIGKDDInt. Conf. Knowl. Discovery Data Mining KDD*, 2015, pp. 2267_2276.
- [8] M. Alvarado, F. Gonzalez, P. Erskine, D. Cliff, and D. Heuff, "A methodology to monitor airborne PM10 dust particles using a small unmanned aerial vehicle," *Sensors*, vol. 17, no. 2, p. 343, 2017.
- [9] Kok, M. U. Simsek, and S. Ozdemir, "A deep learning model for air quality prediction in smart cities," in *Proc. IEEE Int. Conf. Big Data (BigData)*, Dec. 2017, pp. 1983_1990.
- [10] S. Devarakonda, P. Sevusu, H. Liu, R. Liu, L. Iftode, and B. Nath, "Realtime air quality monitoring through mobile sensing in metropolitan areas," in *Proc. 2nd ACM SIGKDD Int. Workshop Urban Comput. UrbComp*, 2013, p. 15.



10.22214/IJRASET



45.98



IMPACT FACTOR:
7.129



IMPACT FACTOR:
7.429



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089  (24*7 Support on Whatsapp)