DEVELOPMENT OF A PREDICTIVE IOT-BASED SYSTEM FOR REAL-TIME AIR QUALITY MONITORING AND ROUTE OPTIMIZATION

Final Presentation

2024 - 078

Meet Our Team !

Supervisor : Ms. Chathurangika Kahandawarachi

Co-supervisor : Ms. Pipuni Wijeseri

Inthikhaff M. I. M

IT21058028

Rimas M. J. M

IT21003714

KUMARI J.M.D

IT21152832

Pasan M.G.R

IT21057588

SNAPS FROM THE FIELD VISITS

Air Quality Data

NE NBRO ESSD <nbro.essd@gmail.com>
4/1/2024 4:53 PM

To: mhdinthikaff@gmail.com

Save all attachments

- AAQ daily Data Colombo_MET... 71.84 KB
- AAQ daily data_Mobile... 108.07 KB
- AAQ hourly data_Mobil 254.81 KB
- Air Quality hourly data_... 216.08 KB

Dear Student,

Please find related docs.

Thanks & Best Regards

Environmental Studies and Services Division
National Building Research Organisation
No. 99/1, Jawatte Road,
Colombo 05
Tel: 0112588946, 0112503826 Fax: 0112588365
Ext : 617 - Head of Division
589 - Air Quality, Noise & Vibration Studies Unit
621 - Land & Water Quality Studies Unit

RESEARCH PROBLEM

- How can air quality monitoring and analysis be integrated into a predictive system to forecast and mitigate environmental risks?
- How we define Air pollution with the creation of IoT device.
- How we can develop user-friendly and easily understandable Pollution map?
- How can we encourage people to travel with minimal levels of pollution ?
- How can we enhance transparency and raise awareness among the public regarding air pollution and pollution levels?
- How can we find a solution for school students who are being affected by air pollution?

RESEARCH OBJECTIVES

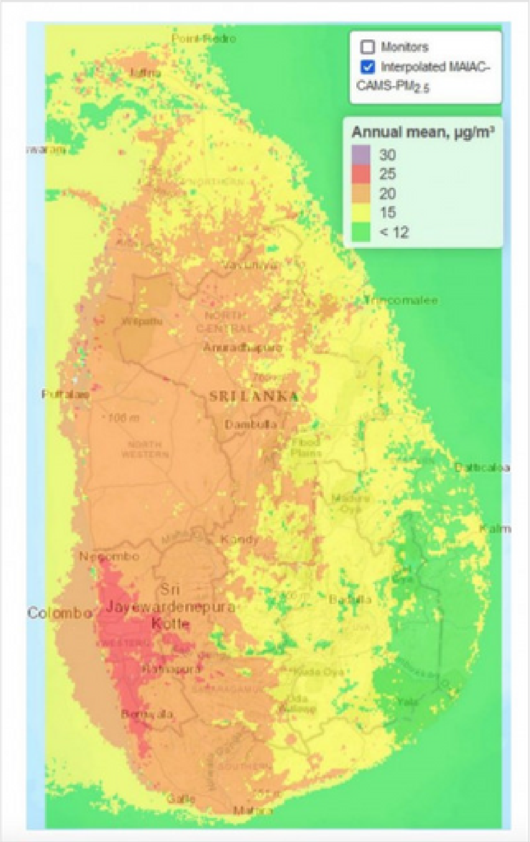
Main Objective

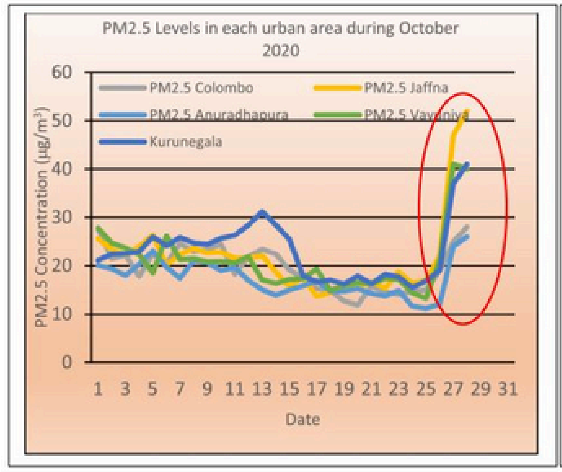
Minimize the adverse effects of air pollution for the general public population while travelling.

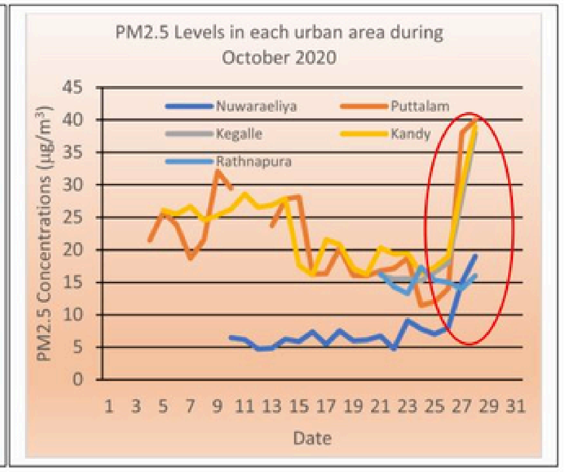
Sub Objectives

1. **Develop sensor-based devices with a focus on accuracy and value creation on Air pollution.**
2. **Implement future pollution level prediction and forecasting mechanisms**
3. **Create a user-friendly real-time pollution heatmap accessible to all.**
4. **Develop real-time route generation based on pollution levels.**

Significant of this study







Kriged MAIAC-CAMS-PM2S concentrations at 1 km intervals over Sri Lanka.

- **This map illustrates the air quality across Sri Lanka.**
- **Several areas are experiencing unexpected increases in air pollution levels.**

These table Illustrate the information about the selected disease categories and the discharge mode.

- More than one fourth Of the deaths are reported from the ischemic heart diseases.
- ***Almost one fourth Of the patients are discharged live those Who had the 'Other acute lower respiratory infections".***

Disease category		Live discharge	Transfers	Deaths	LAMA	Missing	Total
Acute upper respiratory infections	n	51023	2015	6	2	3	53049
	%	10.2%	5.1%	0.0%	2.7%	11.1%	9.5%
Influenza and pneumonia	n	12075	855	2328	4	0	15262
	%	2.4%	2.2%	15.8%	5.3%	0.0%	2.7%
Other acute lower respiratory infections	n	109910	5773	307	1	2	115993
	%	21.9%	14.7%	2.1%	1.3%	7.4%	20.8%
Other diseases of upper respiratory tract	n	10622	230	7	0	0	10859
	%	2.1%	0.6%	0.0%	0.0%	0.0%	2.0%
Chronic lower respiratory diseases	n	97055	4694	1182	9	4	102944
	%	19.3%	11.9%	8.0%	12.0%	14.8%	18.5%
Lung diseases due to external agents	n	646	43	805	0	0	1494
	%	0.1%	0.1%	5.5%	0.0%	0.0%	0.3%
Other respiratory diseases principally affecting the interstitium	n	631	24	120	1	0	776
	%	0.1%	0.1%	0.8%	1.3%	0.0%	0.1%
Suppurative and necrotic conditions of lower respiratory tract	n	590	28	28	0	0	646
	%	0.1%	0.1%	0.2%	0.0%	0.0%	0.1%
Other diseases of pleura	n	1265	115	30	1	0	1411
	%	0.3%	0.3%	0.2%	1.3%	0.0%	0.3%
Other diseases of the respiratory system	n	9554	432	140	2	0	10128
	%	1.9%	1.1%	1.0%	2.7%	0.0%	1.8%
Hypertensive diseases	n	39803	3350	393	9	5	43560
	%	7.9%	8.5%	2.7%	12.0%	18.5%	7.8%
Ischaemic heart diseases	n	45621	11424	3871	9	4	60929
	%	9.1%	29.1%	26.3%	12.0%	14.8%	11.0%
Cerebrovascular diseases	n	21388	6005	2270	12	1	29676
	%	4.3%	15.3%	15.4%	16.0%	3.7%	5.3%
Diabetes mellitus	n	37932	2345	626	7	6	40916
	%	7.6%	6.0%	4.3%	9.3%	22.2%	7.4%
Disorders of lens	n	41046	55	0	4	0	41105
	%	8.2%	0.1%	0.0%	5.3%	0.0%	7.4%
Malignant neoplasms of respiratory and intrathoracic organs	n	7619	140	536	10	2	8307
	%	1.5%	0.4%	3.6%	13.3%	7.4%	1.5%
Pulmonary heart disease and diseases of pulmonary circulation	n	2524	78	126	0	0	2728
	%	0.5%	0.2%	0.9%	0.0%	0.0%	0.5%
Other forms of heart disease	n	12977	1708	1929	4	0	16618
	%	2.6%	4.3%	13.1%	5.3%	0.0%	3.0%
Total	n	502281	39314	14704	75	27	556401
	%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%

Cross-tabulation of selected diseases and the discharge mode

Comparison of deaths due to the air pollution related diseases and the N02 and S02 concentrations

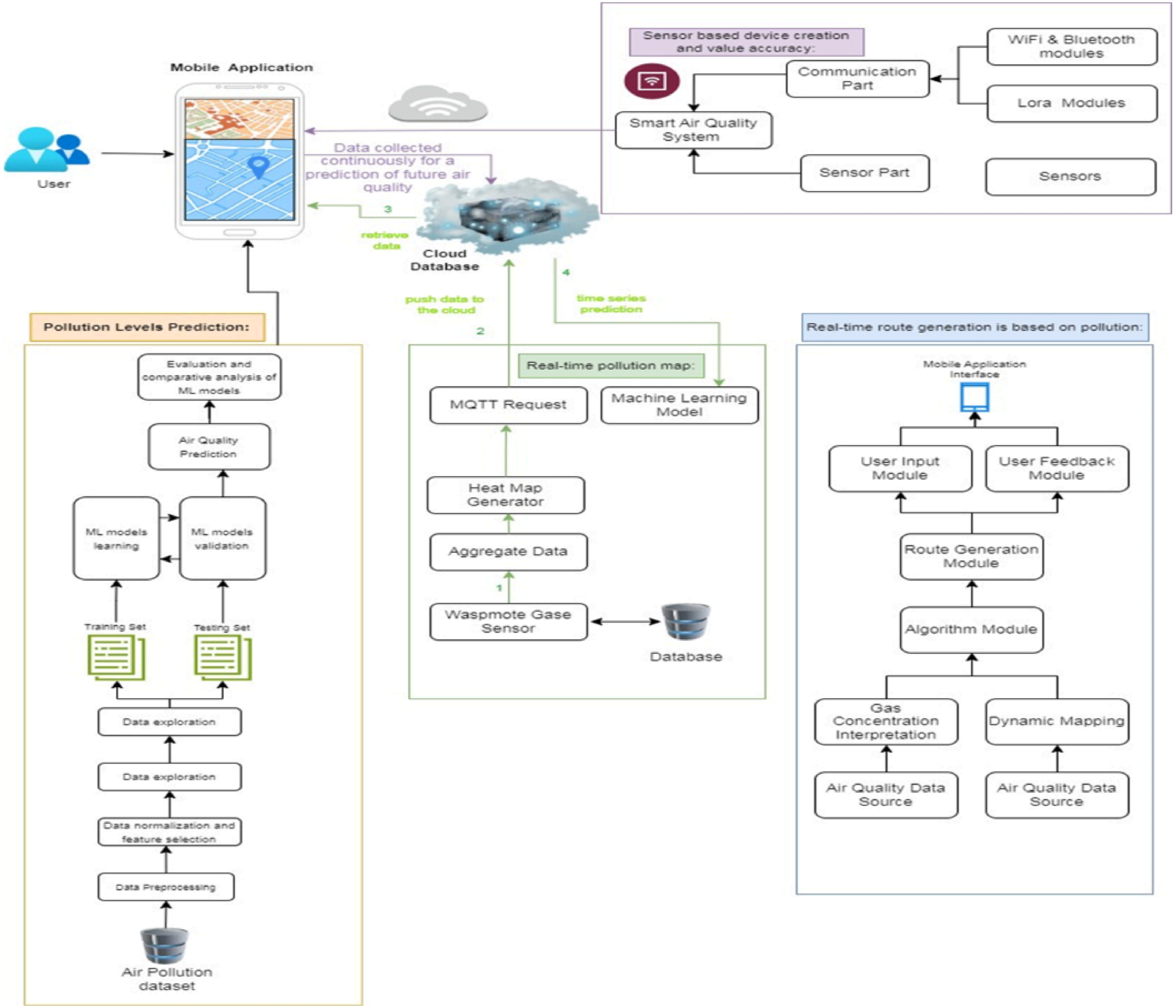
	Deaths	Population	Adjusted population*	Total estimated deaths**	Crude Death rate***	NO_2	SO_2
Anuradhapura	1261	918000	918000	1261	1.374	23	39
Colombo	2717	2419000	2177100	5846	2.685	39	51
Galle	1765	1113000	1113000	1765	1.586	27	48
Gampaha	2297	2391000	2151900	2297	1.067	32	49
Kalutara	1260	1271000	1143900	1260	1.101	32	38.5
Kandy	2002	1452000	1452000	2002	1.379	39	44
Kurunegala	2112	1694000	1694000	2112	1.247	41	49
Ratnapura	1290	1151000	1151000	1290	1.121	30	43

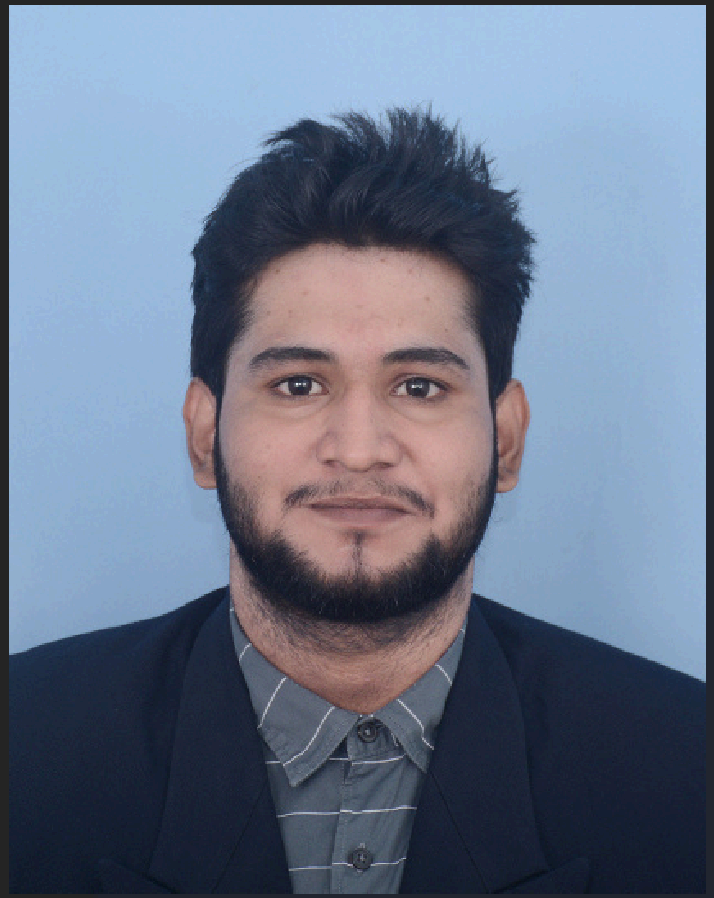
* Population consider the proportions obtain government health sector care
**Total deaths based on the estimation for NHSL
***Per 1000 population

Comparison of death rate and air quality levels in selected districts

Addressing air pollution exposure among school children and the elders is a timely research topic.
By doing so, we can effectively mitigate the adverse effects of air pollution.

System Diagram





IT21003714 | RIMAS M. J. M

Specialization : Information Technoogy

Component 2 : Sensor based device creation and value accuracy

INTRODUCTION

Background

- Urban areas in Sri Lanka face severe air pollution, mainly due to vehicle emissions, leading to serious health problems such as respiratory diseases and increased mortality. Existing air quality monitoring systems are often too expensive, offer limited coverage, or lack real-time data capabilities. This situation underscores a critical need for a more affordable, accurate, and real-time air quality monitoring solution. Such a system would help better understand and address the impact of vehicle emissions on both public health and the environment.

- The primary focus of this research is to develop and implement a cost-effective IoT-based air pollution monitoring system specifically designed to measure vehicle emissions in urban areas of Sri Lanka

Introduction

Research Question

- What is the performance of the IoT-based air pollution monitoring system in terms of data accuracy and transmission reliability ?

- How can MQ4, MQ7, and MQ135 sensors be calibrated and integrated with an Arduino ESP32 to ensure accurate detection and measurement of pollutants?

Specific and Sub Objective

Sensor based device creation and value accuracy

Integrate MQ4 and MQ7 sensors with an Arduino board and get the reading values

Set up MongoDB database to store sensor data securely

Implement Real-Time Data Collection and Transmission(MQTT)

Sensor Integration and Calibration

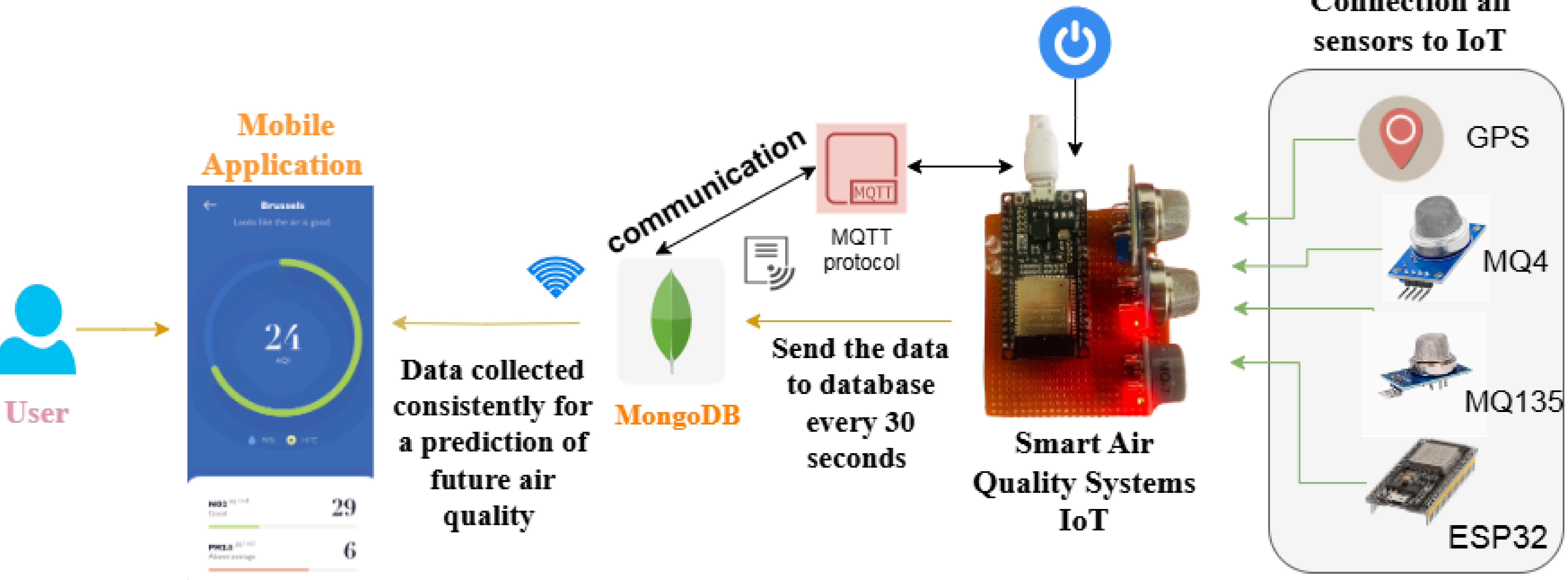
Methodology

Existing Studies on sensor based device creation and value accuracyon

Author	Application	Techniques	Remark
Rauniyar, A., and colleagues	IoT-based car emissions monitoring system for smart cities	Reacts, a PostgreSQL database	85.2% of the vehicles were classified as normal emitters, 7.1% were medium emitters, while 7.7% of the vehicles were classified as high emitter vehicles
Kaivonen, S. and Ngai, E.C.H	real-time air pollution monitoring using wireless sensors on public transport vehicles	Python scripts, MongoDB, JSON, HTTP API, PHP	
Zhang, D. and Woo, S.S.	air quality pattern in the area using both moving and fixed IoT sensors mounted in the vehicles	mobile and stationary IoT sensors integrated into the vehicles	machine learning algorithms and real-time
Kumar, A	IoT-based system for monitoring vehicle pollution	proposed an Internet of Things (IoT) solution aimed at monitoring vehicle emissions	
Asha, P., et	IoT enabled environmental toxicology for air pollution monitoring using AI techniques	Artificial Algae Algorithm	model, which involves the monitoring of eight pollutants (NH_3, CO, NO_2, CH_4, CO_2, PM2.5, temperature, and humidity) through an IoT-based sensor
JunHo Jo	IoT-driven Smart-Air system for effective air quality monitoring and real-time	AWS server, PHP, MySQL, IoT-based, JavaScript, web server developed for Android OS	

existing studies

System diagram



here we are using mosquito mqtt sever there something call message brokers there are publisher and subscriber when publiser send the some data if there is a subscriber here receive it

IoT data which we are having topic called IOT/data which we created (customizable) and what happening here is that tops are in the MQTT server (broker).When an IoT device is acting as a publishes to that topic (IOT/data) from this channel our toT device Send/ the (Posts) data

Used Techniques and Technologies

Techniques

- Sensor Calibration and Data Conversion
- Use of MongoDB database for storing sensor data securely.
- Develop an IoT-Based Air Pollution Monitoring System
- Implementation of Arduino code on the Arduino Uno board to handle

Technologies

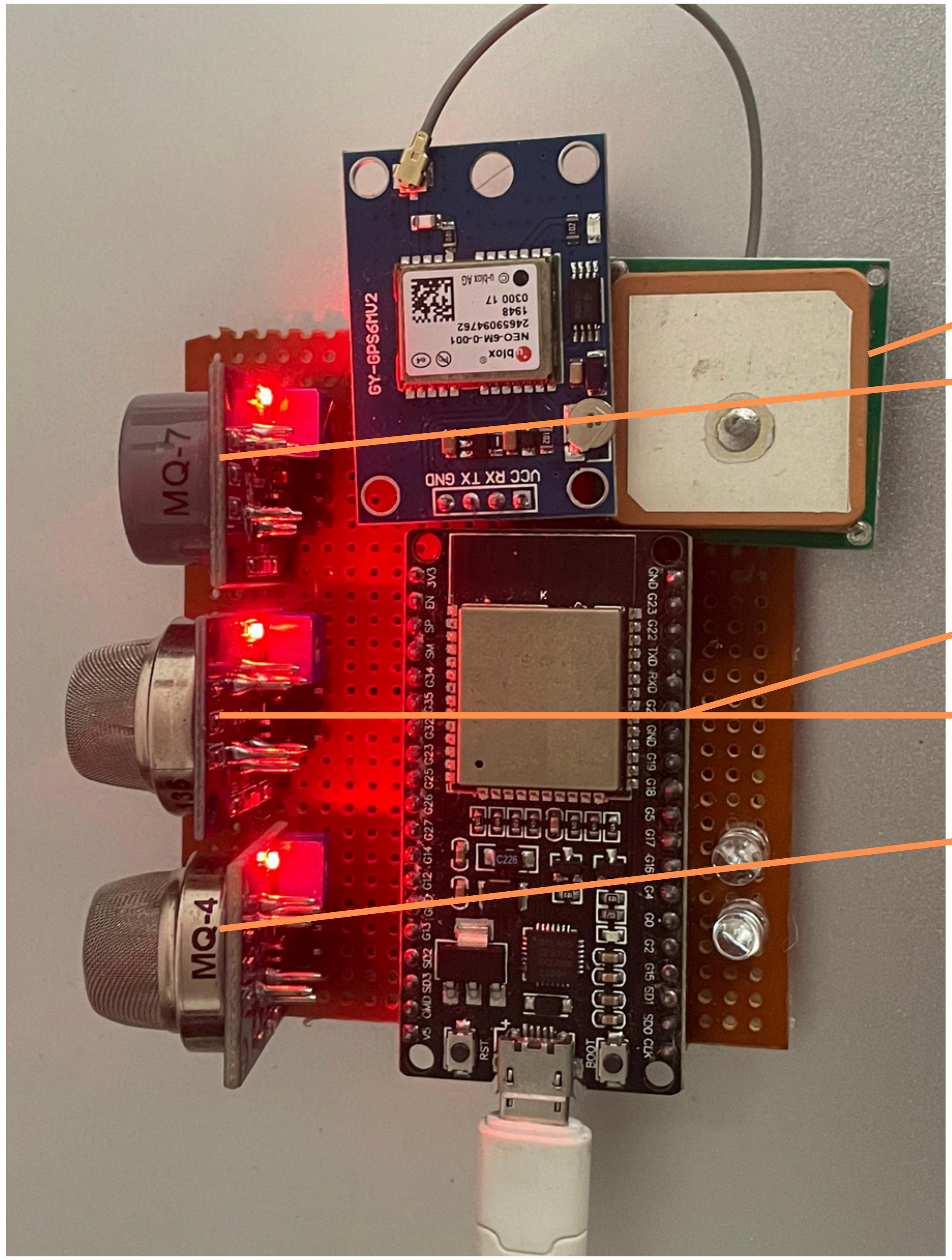
- React Native
- MongoDB
- Node Js
- Fast API
- Arduino IDE
- Visual Studio Code

Evidence of Completion

- create the IoT device
- Connect with Backend server from IoT devices and send the sensor data in database(DB)
- get the real-time Sensor Data from IoT devices
- Host the finalized model in the fast server
- calculate the sensor data values (PPM) concentration
- Displayed the results in mobile application
- Using the MQTT Server to send the data to back end

IoT Devices

Evidence of Completion



GPS

MQ7

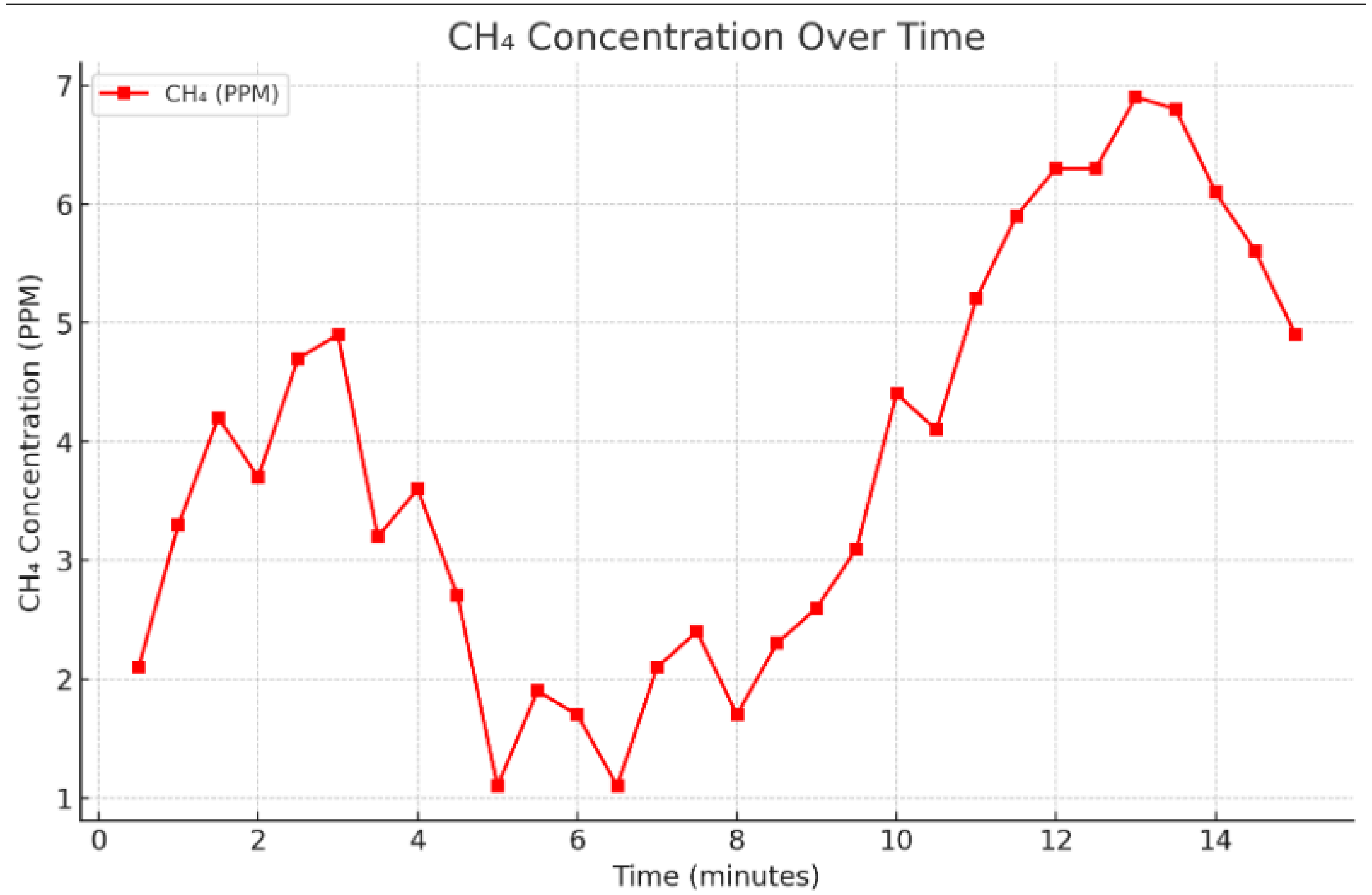
Arduino ESP32

MQ135

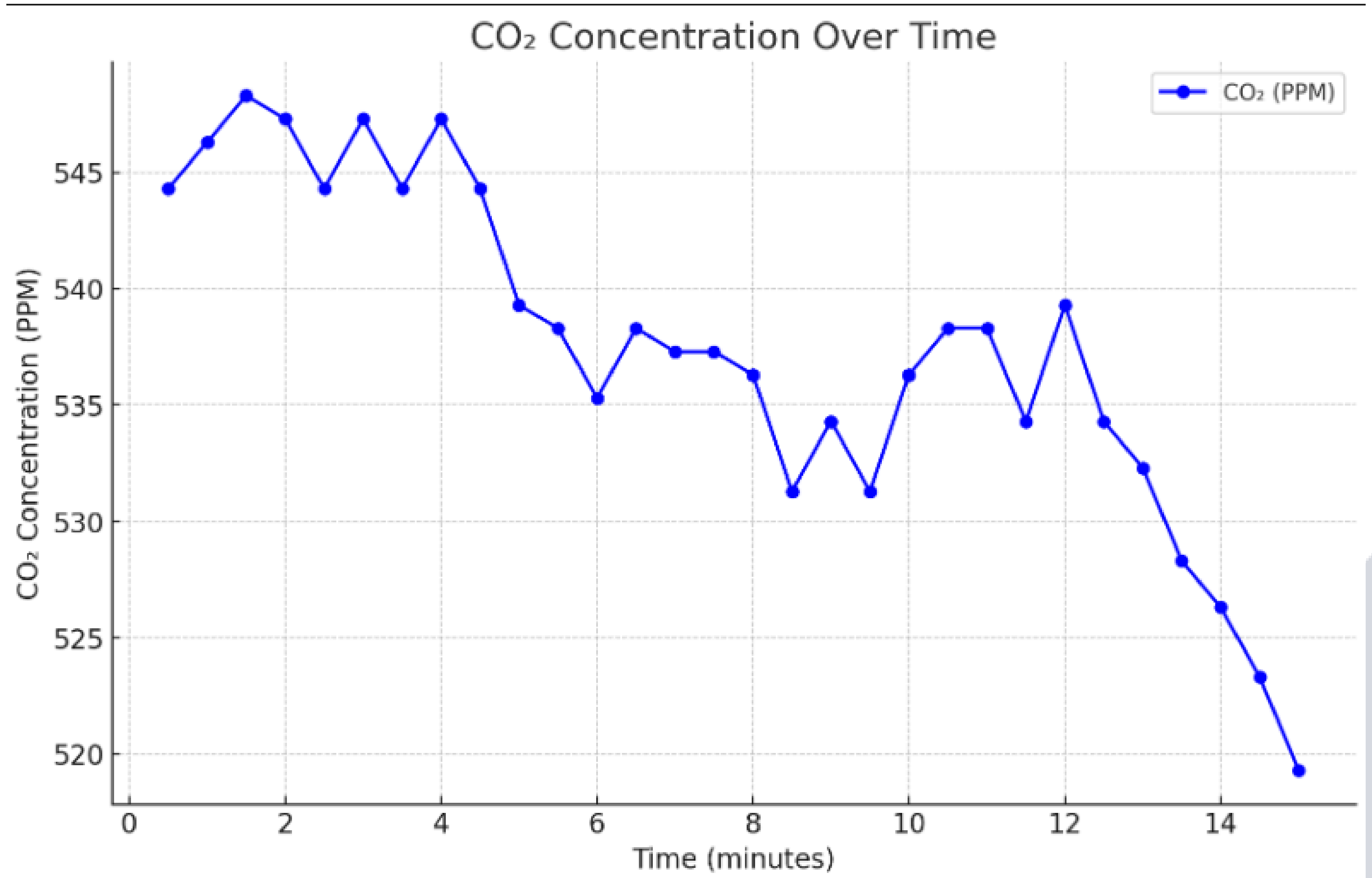
MQ4

REAL TIME SENSOR DATA COLLECTION(FROM IOT)

Evidence of Completion



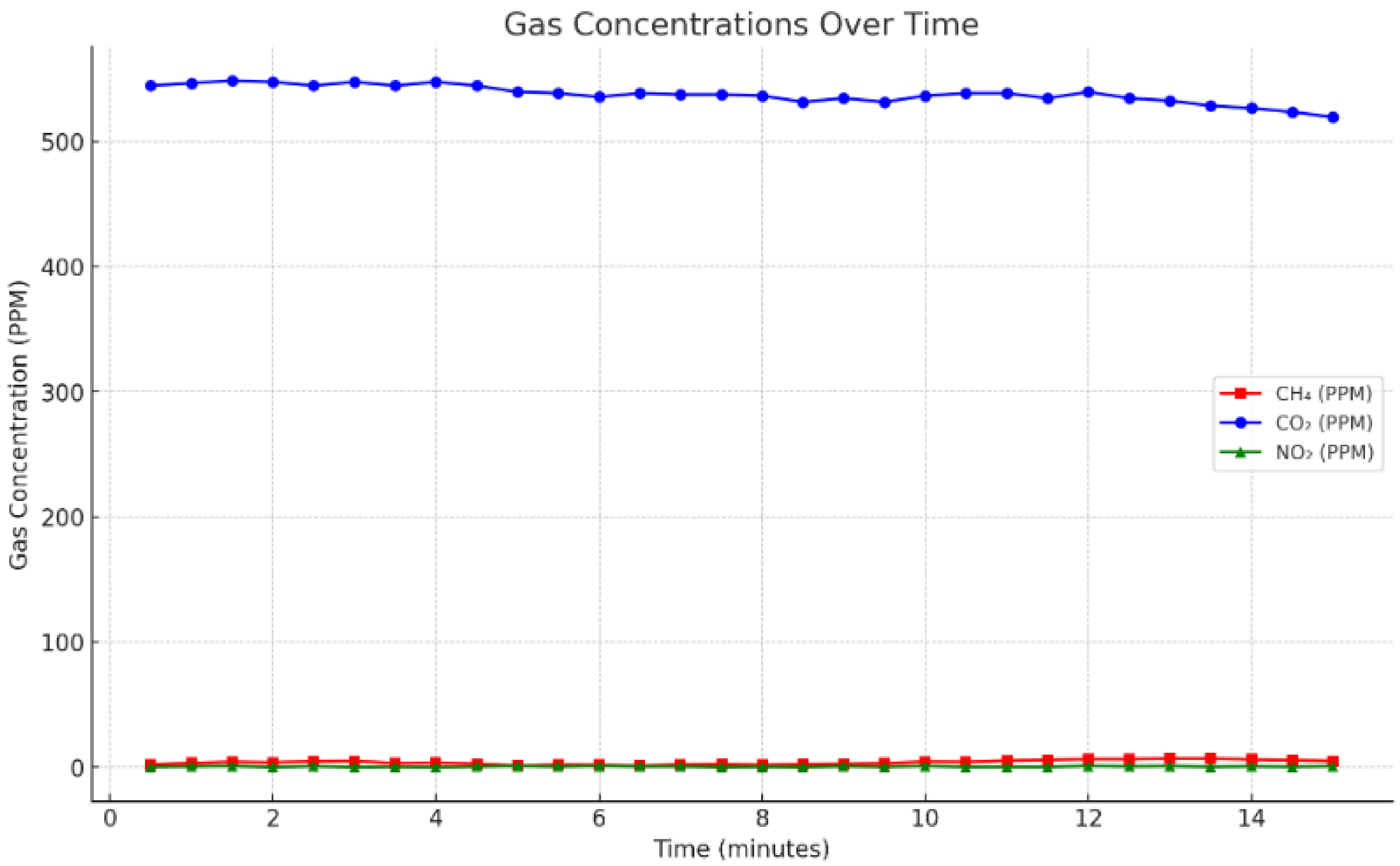
ch4 level time



co2

REAL TIME SENSOR DATA COLLECTION(FROM IOT)

Evidence of Completion



overall behavior

REAL TIME DATA COLLECTION FROM MOBILE APPLICATION

Evidence of Completion

```
{
  iotDeviceId: 'IOT1',
  gpsLatitude: 6.915652275,
  gpsLongitude: 79.9731369,
  co2Level: 334.7085571,
  no2Level: 0.166248277,
  ch4Level: 0.332214385,
  _id: new ObjectId('66de0d7b7eeb59d59ae2d694'),
  timestamp: 2024-09-08T20:47:55.728Z
}
Data saved: {
  iotDeviceId: 'IOT1',
  gpsLatitude: 6.915652275,
```

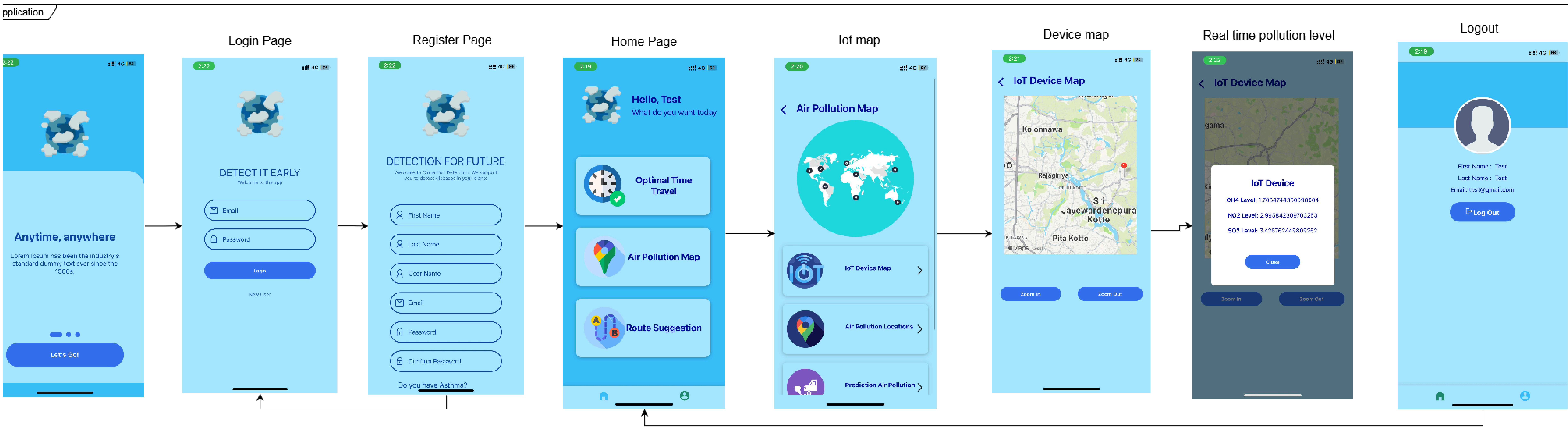
This is the real-time IOT devices data

```
LOG  Connected!
LOG  mqtt
LOG  {"ch4Level": 2.638250188668174, "co2Level": 465.93713642545174, "gpsLatitude": 6.946132, "gpsLongitude": 79.939988, "iotDeviceId": "sim1", "no2Level": 0.8617419092438405, "so2Level": 1.094279035584621}
LOG  mqtt
LOG  {"ch4Level": 0.0001, "co2Level": 0.0001, "gpsLatitude": 6.9089, "gpsLongitude": 79.9314, "iotDeviceId": "sim2", "no2Level": 0.0001, "so2Level": 0.0001}
```

MQTT server transfer the data

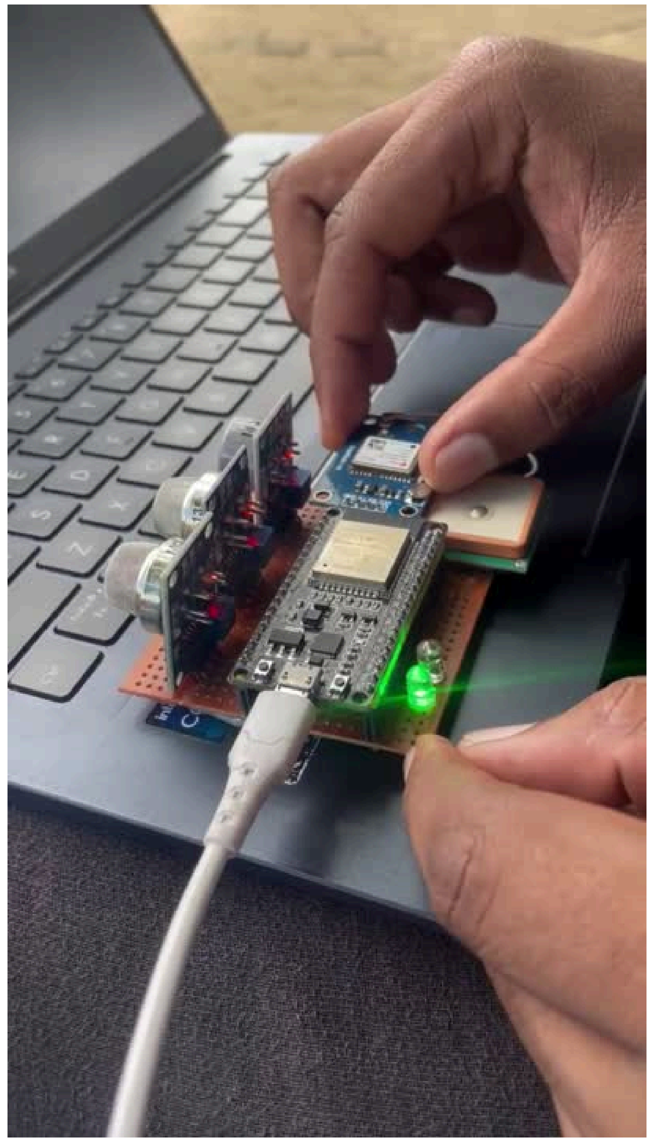
REAL TIME DATA COLLECTION FRONT-END UI PART

Evidence of Completion

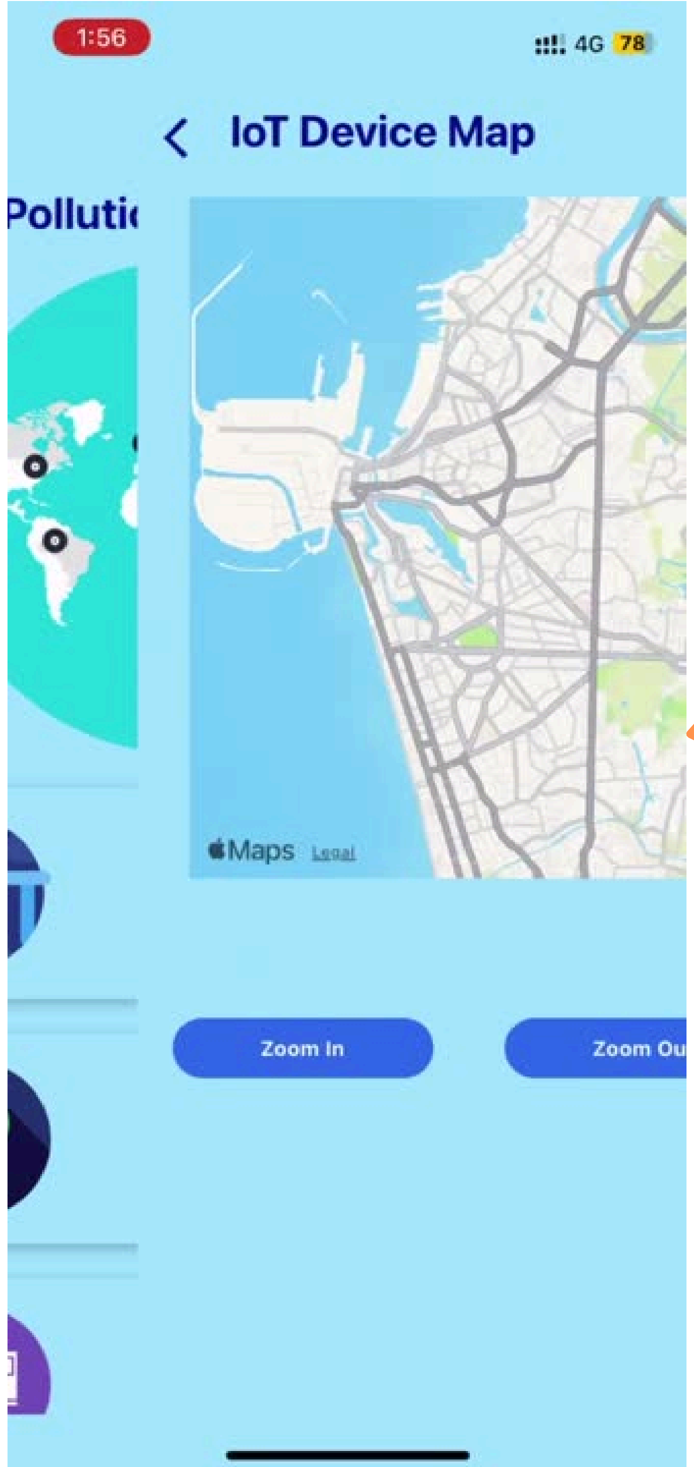


REAL TIME DATA COLLECTION FRONT-END UI PART

Evidence of Completion



Real time environmental tasting

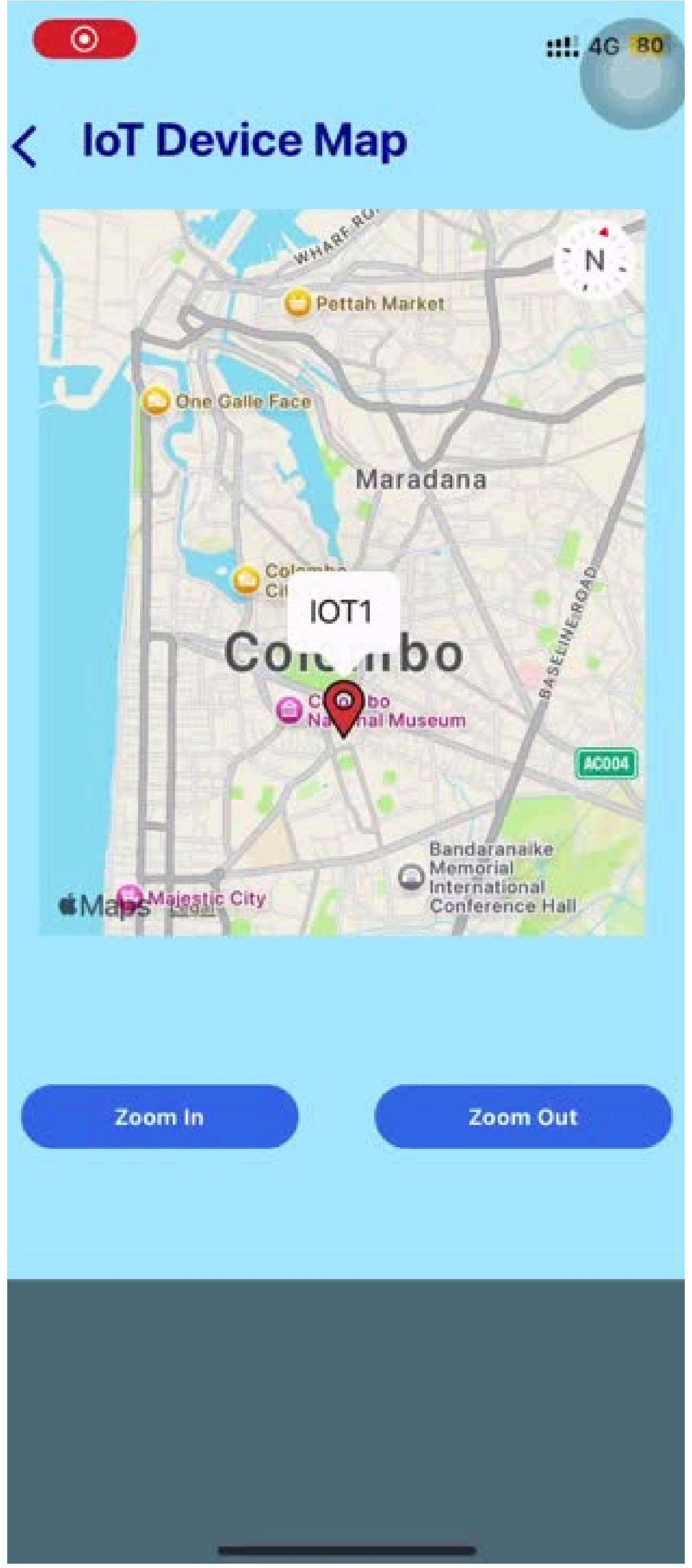


Your paragraph text

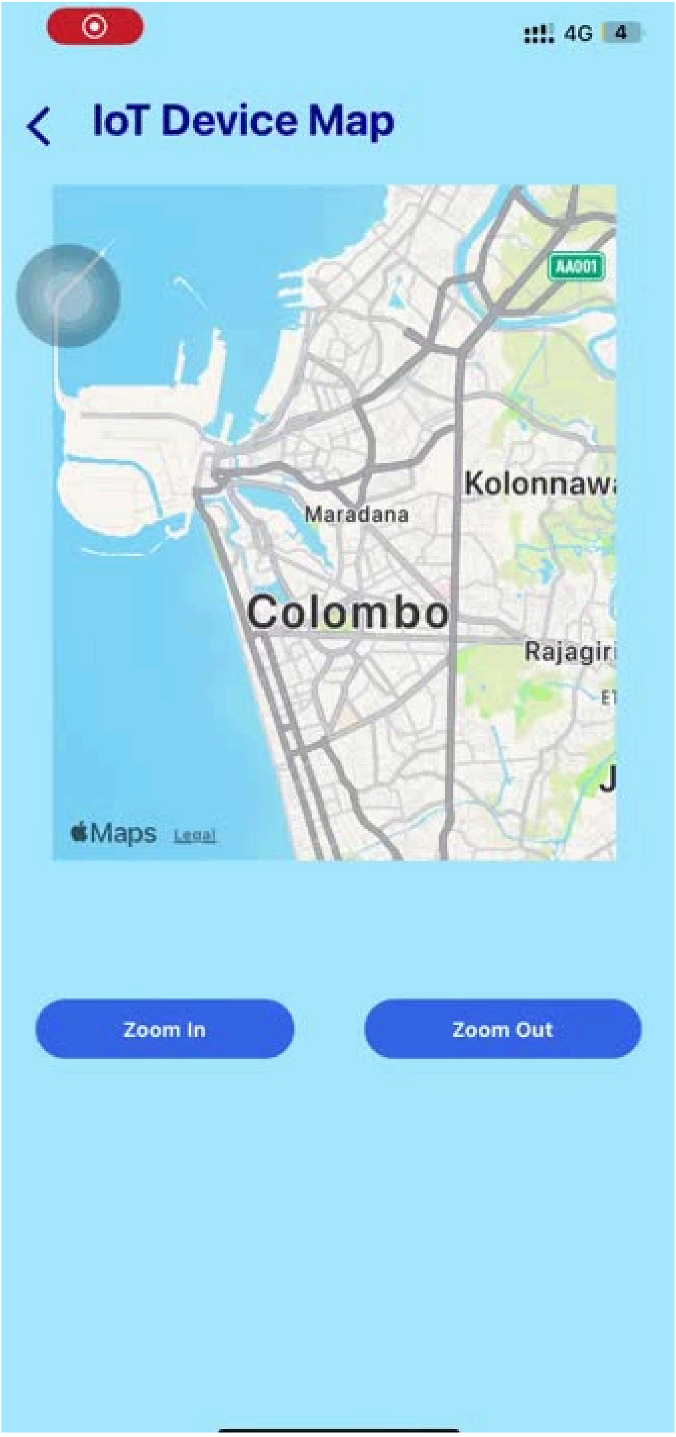
Data transferring method i am using **Playload method** in here i am using rest API call every 30 second make rest call in the playload method to send the data packet

REAL TIME DATA COLLECTION FRONT-END UI PART

Evidence of Completion



Real time environmental tasting



turn of the Iot device can't show the location and device

Functional, Non-Functional and Personnel Requirements

Functional Requirements

- Real-Time Data Transmission
- Sensor Data Collection:
- Set up MongoDB database to store real-time air quality data.
- Calculating the PPM value
- Implement data transmission at regular intervals (e.g., every 30 seconds) to maintain up-to-date information

Non-Functional Requirements

- User-friendly interfaces.
- Provide continuous operation and data transmission
- Ensure compatibility with different mobile platforms (iOS, Android).

Personnel Requirements

- Identify the sensors and board details to creat an IoT device.

Completion and Future works

100% Completion of the components

- Build an IoT Device
- Reading the data from MQ4 AND MQ7 AND MQ135 sensors
- Calculating the PPM value from meter Readings
- **Connect with the Backend server from IoT devices and send the sensor data to the database**
- **Displayed the real-time pollution level in the mobile application**

UI Enhancement | **Increasing The Performance** | **Bug Fixing**

References

[1] Naik, U. U., Salgaokar, S. R., & Jambhale, S. (2023). IOT based air pollution monitoring system. Int. J. Sci. Res. Eng. Trends, 9, 835-838.

[2] Pandithurai, O., Jawahar, M., Arockiaraj, S., & Bhavani, R. (2023). IOT Technology Based Vehicle Pollution Monitoring and Control (Doctoral dissertation, Department of Electrical and Electronics Engineering, Mepco Schlenk Engineering College, Sivakasi).

[3] Dhanalakshmi, M. (2021). A survey paper on vehicles emitting air quality and prevention of air pollution by using IoT along with machine learning approaches. Turkish Journal of Computer and Mathematics Education (TURCOMAT), 12(11), 5950-5962.

[4] Potbhare, Piyush Devidas, et al. "IoT based Smart Air Pollution Monitoring System." 2022 International Conference on Applied Artificial Intelligence and Computing (ICAAIC). IEEE, 2022..

[5] [1]Mihăiță, A. S., Dupont, L., Chery, O., Camargo, M., & Cai, C. (2019). Evaluating air quality by combining stationery, smart mobile pollution monitoring and data-driven modelling. Journal of cleaner production, 221, 398-418..

[6] [1] .N.S. Attanayake and R.A.B. Abeygunawardana, A Comprehensive Comparison of Air Pollution in Main Cities in Sri Lanka Department of Statistics, Faculty of Science, University of Colombo, Colombo 03, Sri Lanka

IT21058028 | Inthikhaff M. I. M

Specialization : Information Technoogy

Component 1 : Air Pollution Level Prediction

Introduction

Background

- **This component main objective is to predict the Air Pollution Level** using regression models such as Random Forest, Linear Regression, Xgbhoost, Support Vactor Regression and Prophet Method.
- Our primary focus is on enhancing existing methods to accurately predict and understand air quality.
- Colombo consistently experiences an average unhealthy AQI, making reliable predictions crucial for guiding public precautions regarding indoor and outdoor activities.
- The correlation between AQI and threats to human health, including short-term and long-term effects, emphasizes the importance of accurate predictions.
- Through out this various Model training, we are finding the best model which gives high accuracy in air pollution level prediction.

Introduction
Research Question

How to predict air pollution levels of a given time in future and forecasted?

How to filter which inputs(atributes) perform more side effect on air pollution

Specific and Sub Objective

Implement future pollution level prediction and forecasting mechanisms

- **Identification of data set, cleaning and preprocessing.**
- **Training data set using different model and Finding best model which gives high level accuracy.**
- **Graphical and comparative visual representation.**
- **Notification to user when predicted polution level is high**

Methodology

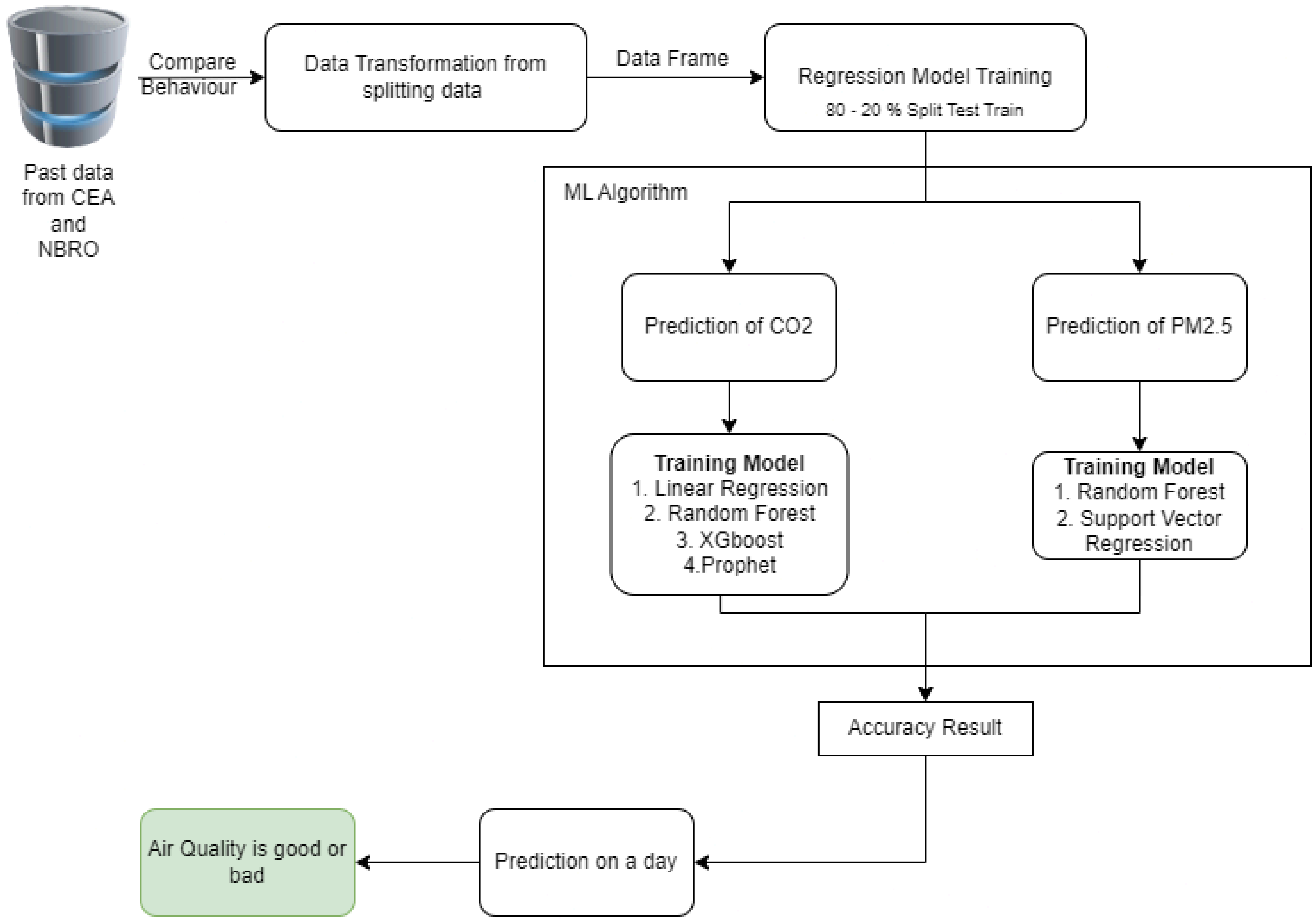
Existing Studies on Pollution Level Prediction

Author	Application	Techniques	Remark
Min Lee and others	Air pollution prediction	• Deep Learning	• Predict against PM 2.5, PM 10 particulars. • Accuracy based on PM 10 is very low. • Accuracy based on PM2.5 is very high.
Usha Mahalingam and others	Air quality prediction	• Neural Networks • Support Vector Machine	• Accuracy of 91.62% for neural network. • Accuracy of 97.3% for support vector machine
S. Silva and others	Air quality prediction for smart cities	• Support Vector Regression	• Predict PM 2.5 levels variability. • Model is suitable for predict hourly air pollution. • Obtain an accuracy of 94.1%
Timothy M. A. and others	Development of Air Quality Monitoring model	• Naive Bayesian • KNN • Support Vector Machines • Neural Networks • Random Forest	• Highest accuracy was obtained through Neural Networks. • Sometimes Neural Network leads slower response.
C. Zhao and others	Air Quality Index Prediction	• Linear regression	• AQI Prediction based on a year data of PM2.5, PM10 etc. • There is a deviation between predicted results and actual date.
Ismail Ahmadi	Air pollution prediction	• Data mining • Decision Tree	• Used clementine software for data clustering. • Data sample include climate data
Colin Belinger and others	A systematic review based on Machine Learning and data mining for Air Pollution.	• Machine Learning Algorithms • Data Mining • Big Data	• Refer 400 research papers & Reduce to 47 after the inclusion/exclusion criteria's. • Divided research papers into three categories • End of the Literature survey that highest accuracy levels always obtain in Machine Learning Algorithms based approaches.

	Time	CO2	PM2.5	PM10	Unnamed: 4	Unnamed: 5	Unnamed: 6	Unnamed: 7	Unnamed: 8
0	2022-12-21 00:00:00	562	15	16.0	NaN	NaN	NaN	NaN	NaN
1	2022-12-21 01:00:00	511	35	40.0	NaN	NaN	NaN	NA- Not Available	NaN
2	2022-12-21 02:00:00	511	36	42.0	NaN	NaN	NaN	MF- Malfunctioned	NaN
3	2022-12-21 03:00:00	507	38	44.0	NaN	NaN	NaN	PM10	Units - µg/m3
4	2022-12-21 04:00:00	498	39	46.0	NaN	NaN	NaN	PM2.5	Units - µg/m3
...	...	...	...	...	...	...	...	...	...
5938	2024-03-21 09:00:00	MF	42	50.0	NaN	NaN	NaN	NaN	NaN
5939	2024-03-21 10:00:00	MF	38	45.0	NaN	NaN	NaN	NaN	NaN
5940	2024-03-21 11:00:00	MF	34	38.0	NaN	NaN	NaN	NaN	NaN
5941	2024-03-21 12:00:00	MF	33	37.0	NaN	NaN	NaN	NaN	NaN
5942	2024-03-21 13:00:00	MF	26	27.0	NaN	NaN	NaN	NaN	NaN

From these I have used Time, CO2, PM 2.5 and PM 10 for prediction

Methodology
System Diagram



Used Techniques and Technologies

Techniques

- DataCollection
- Data Preprocessing
- Normalization
- Regularization
- Algorithm tuning

Algorithm

- Linear Regression
- Random Forest
- XGBoost (Extreme Gradient Boosting)
- Support Vector Regressor
- Prophet (Facebook's Forecasting Model)

Technologies

- React Native
- Python
- Fast API
- Node Server
- Jupyter Notebook
- Google Colab
- VS code

python

React Native

jupyter

Methodology
Evidence of Completion

- Data collection
- Data pre-processing
- Model Training
- Model Evaluation
- Model Testing
- Data Visualization
- Host the finalized model in the fast server
- Displayed the results in mobile application

Data Collection and Pre-processing

- **Data pre-processing implemetation**

```
# Differentiate with time vs predicted variable(creating seperate)

PM_25_dataframe = dataframe_cleaned[['Time',"PM2.5"]]
PM_10_dataframe = dataframe_cleaned[["Time","PM10"]]
CO2_dataframe = dataframe_cleaned[['Time','CO2']]
```

```
#This code filters the DataFrame CO2_dataframe to remove rows where the value in the "CO2" column is equal to "MF".
condition = CO2_dataframe["CO2"] == "MF"
CO2_dataframe = CO2_dataframe[~condition]
```

```
condition = PM_25_dataframe["PM2.5"] == "MF"
PM_25_dataframe = PM_25_dataframe[~condition]
```

```
PM_25_dataframe.dropna(inplace=True)
```

```
<ipython-input-116-56ec3935287e>:1: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy
  PM_25_dataframe.dropna(inplace=True)
```

```
PM_25_dataframe
```

	Time	PM2.5
0	2022-12-21 00:00:00	15
1	2022-12-21 01:00:00	35
2	2022-12-21 02:00:00	36
3	2022-12-21 03:00:00	38
4	2022-12-21 04:00:00	39
...	...	...
5938	2024-03-21 09:00:00	42
5939	2024-03-21 10:00:00	38
5940	2024-03-21 11:00:00	34
5941	2024-03-21 12:00:00	33
5942	2024-03-21 13:00:00	26

Classes	Train, Validation and Test data
CO2 (5702)	Train - 4562 Test - 1140
PM 2.5 (3057)	Train - 2446 Test - 611
PM 10 (5703)	Train - 4563 Test - 1140

- **Data collection implemetation**

```
from google.colab import drive
drive.mount('/content/drive')
```

```
Mounted at /content/drive
```

```
Start coding or generate with AI.
```

oad the hourly data set and clean the data

```
import pandas as pd

dataframe = pd.read_excel(r"/content/drive/MyDrive/air_pollution_project/Air Quality hourly data_ Colombo_MET Dept.xlsx")
```

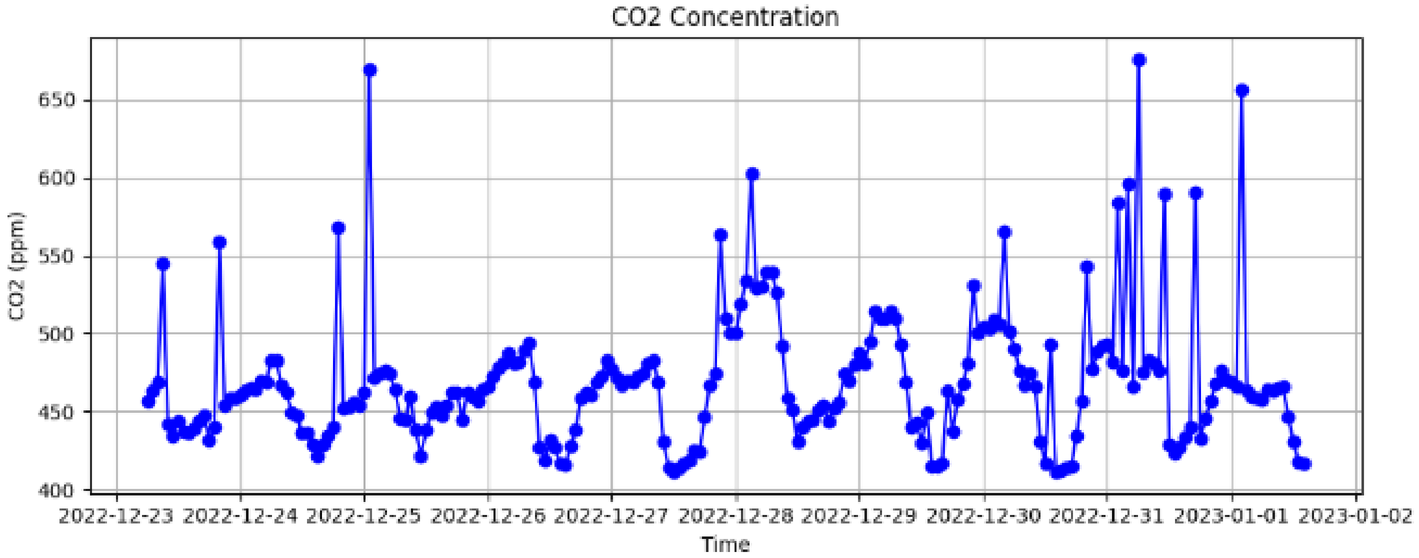
```
dataframe
```

	Time	CO2	PM2.5	PM10	Unnamed: 4	Unnamed: 5	Unnamed: 6	Unnamed: 7	Unnamed: 8
0	2022-12-21 00:00:00	562	15	16.0	NaN	NaN	NaN	NaN	NaN
1	2022-12-21 01:00:00	511	35	40.0	NaN	NaN	NaN	NA- Not Available	NaN
2	2022-12-21 02:00:00	511	36	42.0	NaN	NaN	NaN	MF- Malfunctioned	NaN
3	2022-12-21 03:00:00	507	38	44.0	NaN	NaN	NaN	PM10	Units - µg/m3
4	2022-12-21 04:00:00	498	39	46.0	NaN	NaN	NaN	PM2.5	Units - µg/m3
...	...	...	...	...	...	...	...	...	...
5938	2024-03-21 09:00:00	MF	42	50.0	NaN	NaN	NaN	NaN	NaN
5939	2024-03-21 10:00:00	MF	38	45.0	NaN	NaN	NaN	NaN	NaN
5940	2024-03-21 11:00:00	MF	34	38.0	NaN	NaN	NaN	NaN	NaN
5941	2024-03-21 12:00:00	MF	33	37.0	NaN	NaN	NaN	NaN	NaN
5942	2024-03-21 13:00:00	MF	26	27.0	NaN	NaN	NaN	NaN	NaN

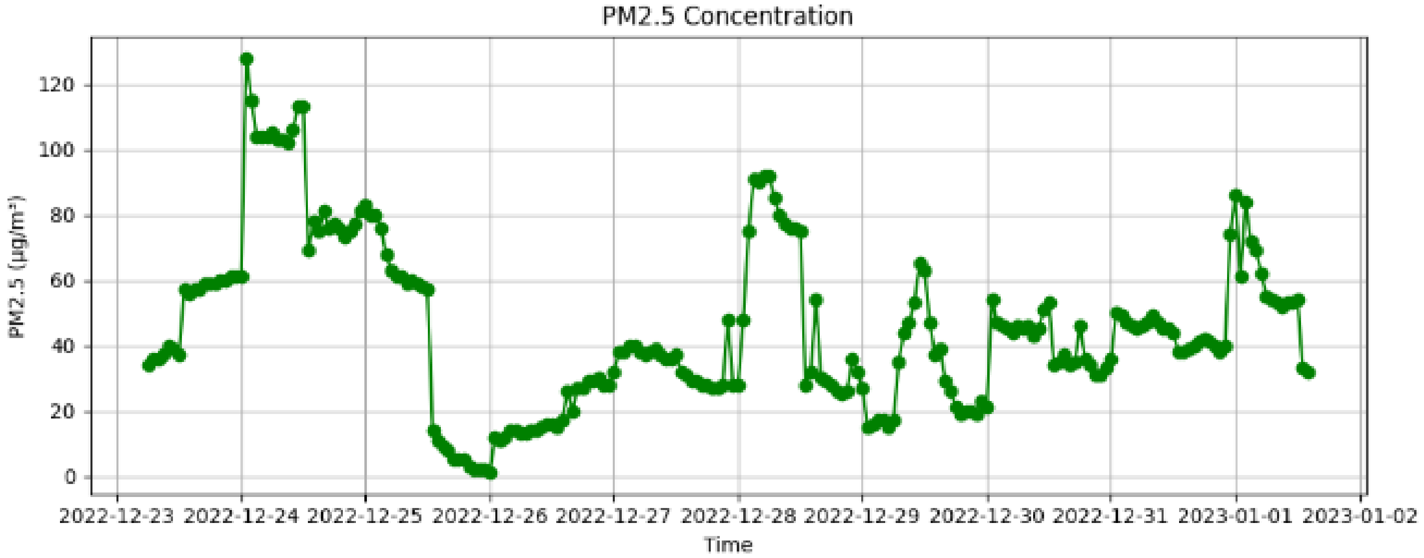
5943 rows × 9 columns

Total Data : 5943 | Preprocessed Data : 5703 | 80 -20 % Train Test Split

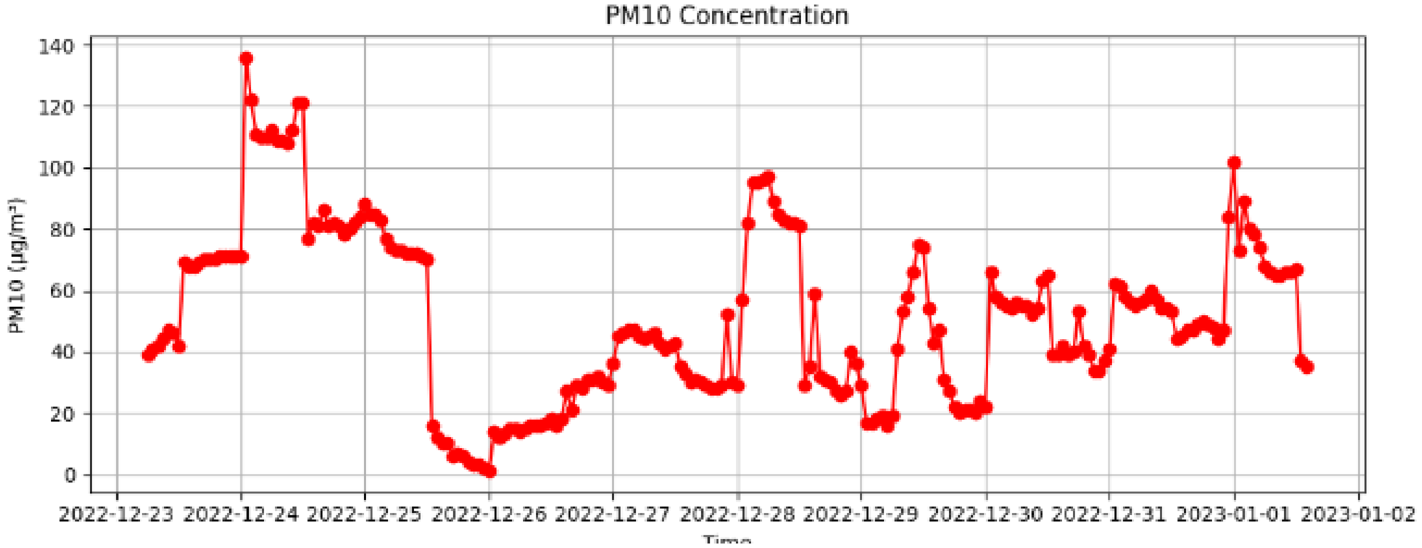
Actual Representation of Data
Evidence of Completion



CO2



PM 2.5



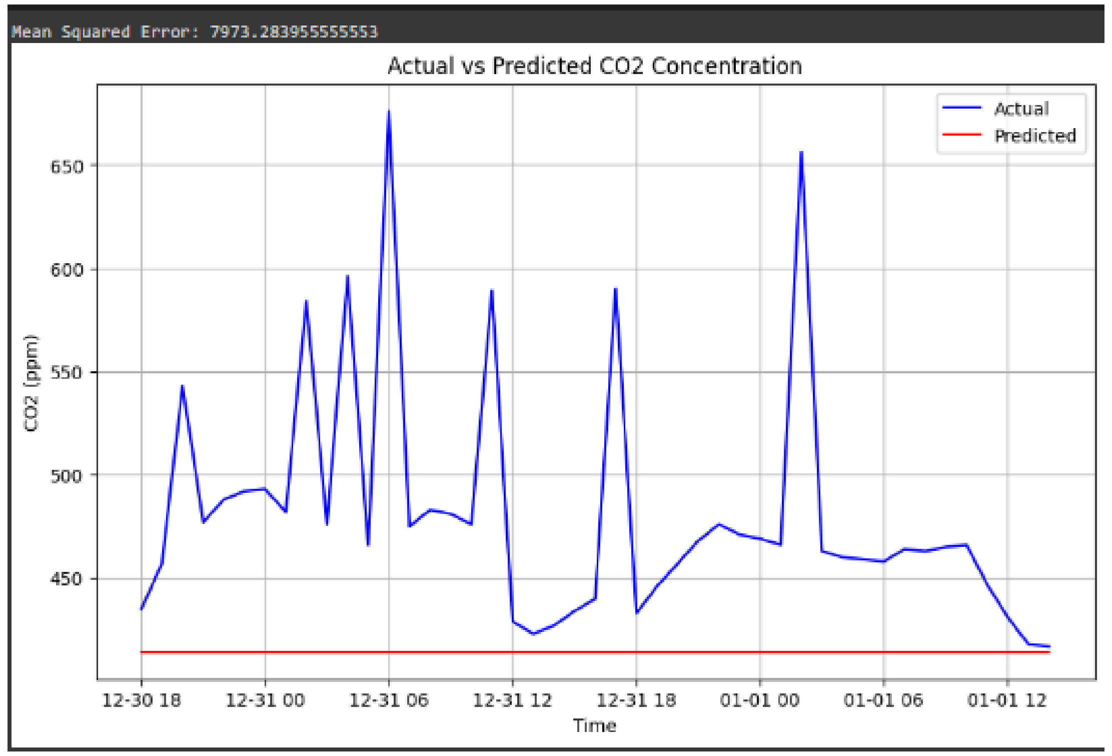
PM 10

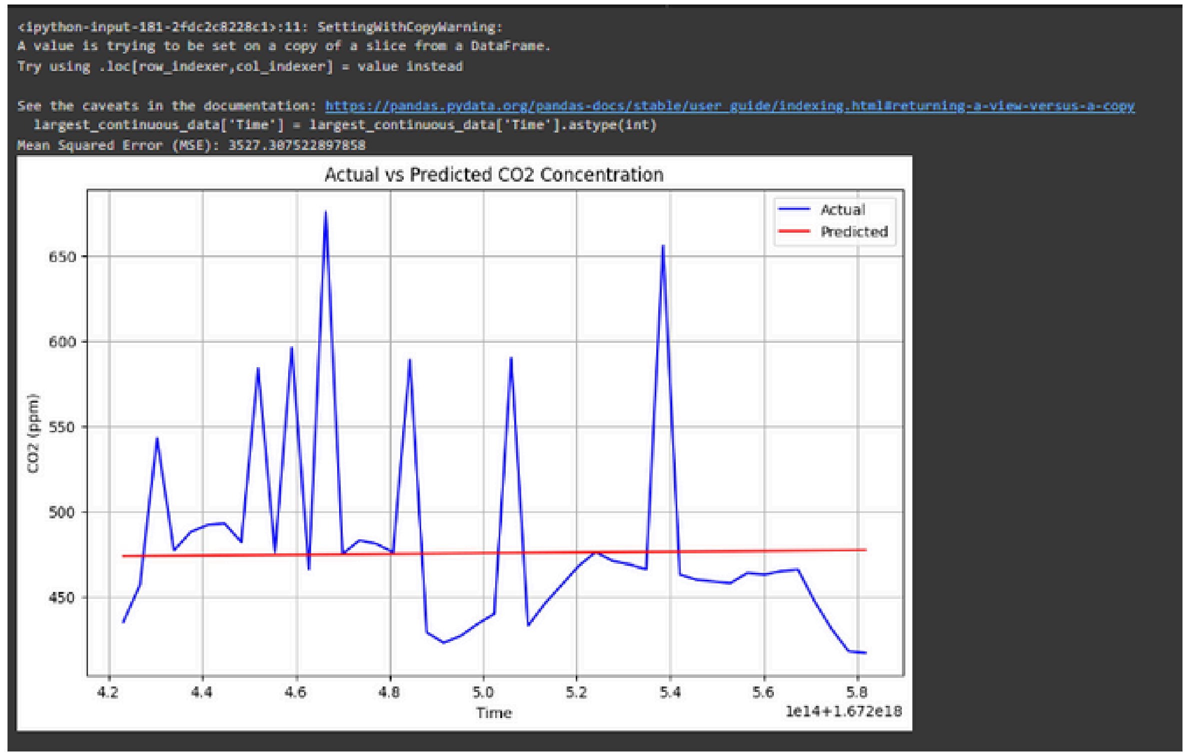
Train Model - CO2
Evidence of Completion

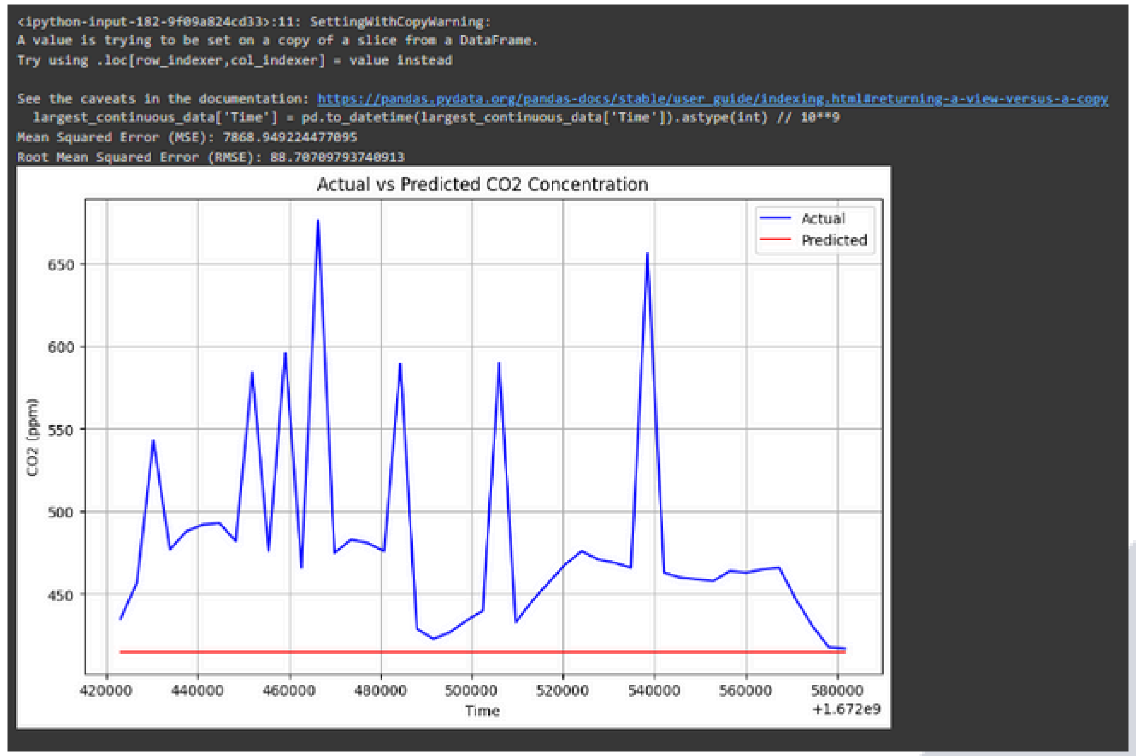
Random Forest

Linear Regression

XGbhoost








```
Mean Squared Error: 7973.283955555553
```

```
<ipython-input-181-2fdc2c8228c1>:11: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy
  largest_continuous_data['Time'] = largest_continuous_data['Time'].astype(int)
Mean Squared Error (MSE): 3527.307522897858
```

```
<ipython-input-182-9f09a824cd33>:11: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy
  largest_continuous_data['Time'] = pd.to_datetime(largest_continuous_data['Time']).astype(int) // 10**9
Mean Squared Error (MSE): 7868.949224477095
Root Mean Squared Error (RMSE): 88.70709793740913
```

MSE - 7973.28

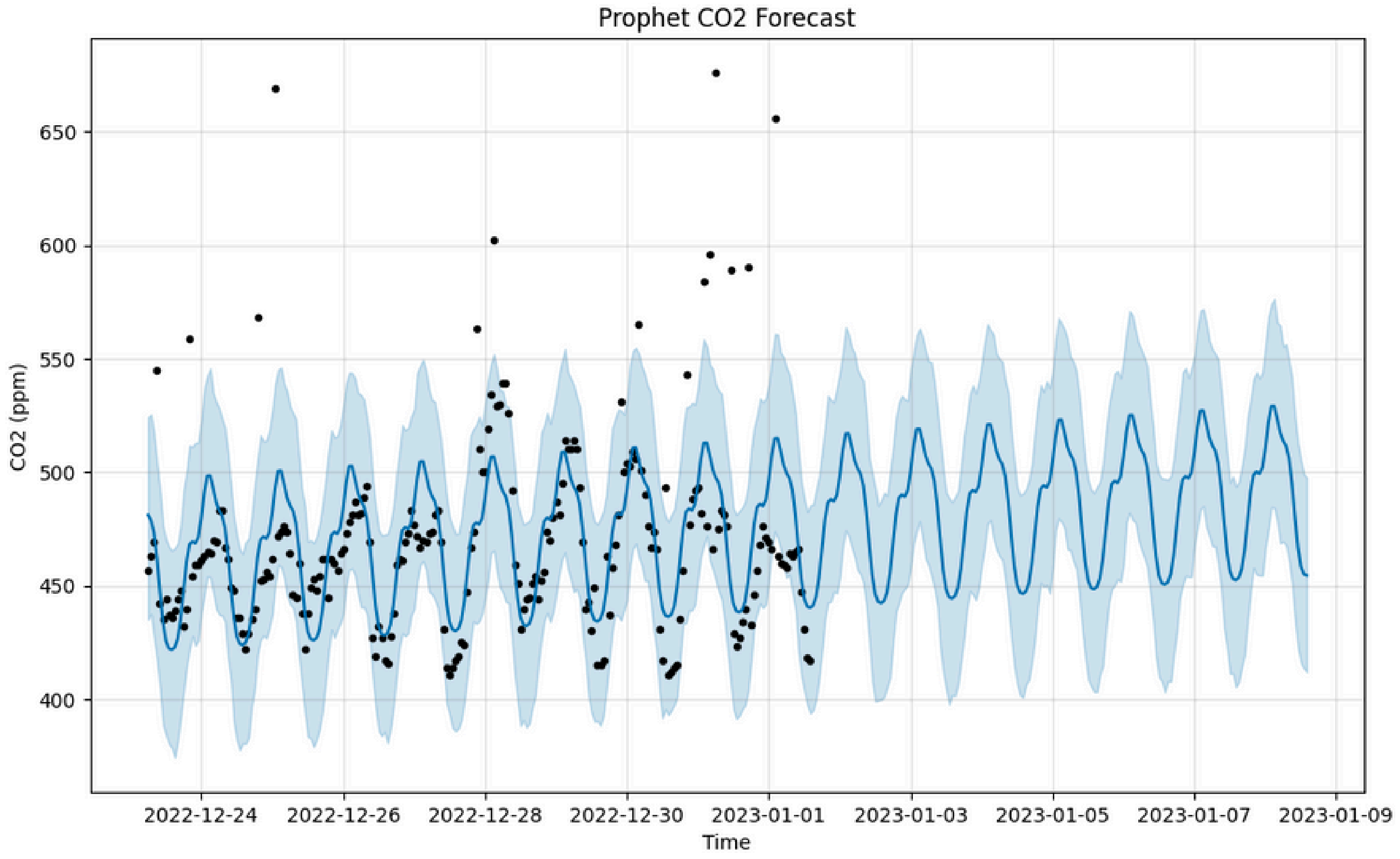
MSE - 3527.31

MSE - 88.71

Lover value indicates a better fit in MSE

Train Model - CO2
Evidence of Completion

Prophet




```
Root Mean Squared Error (RMSE): 54.452076807335985
```

MSE - 54.45

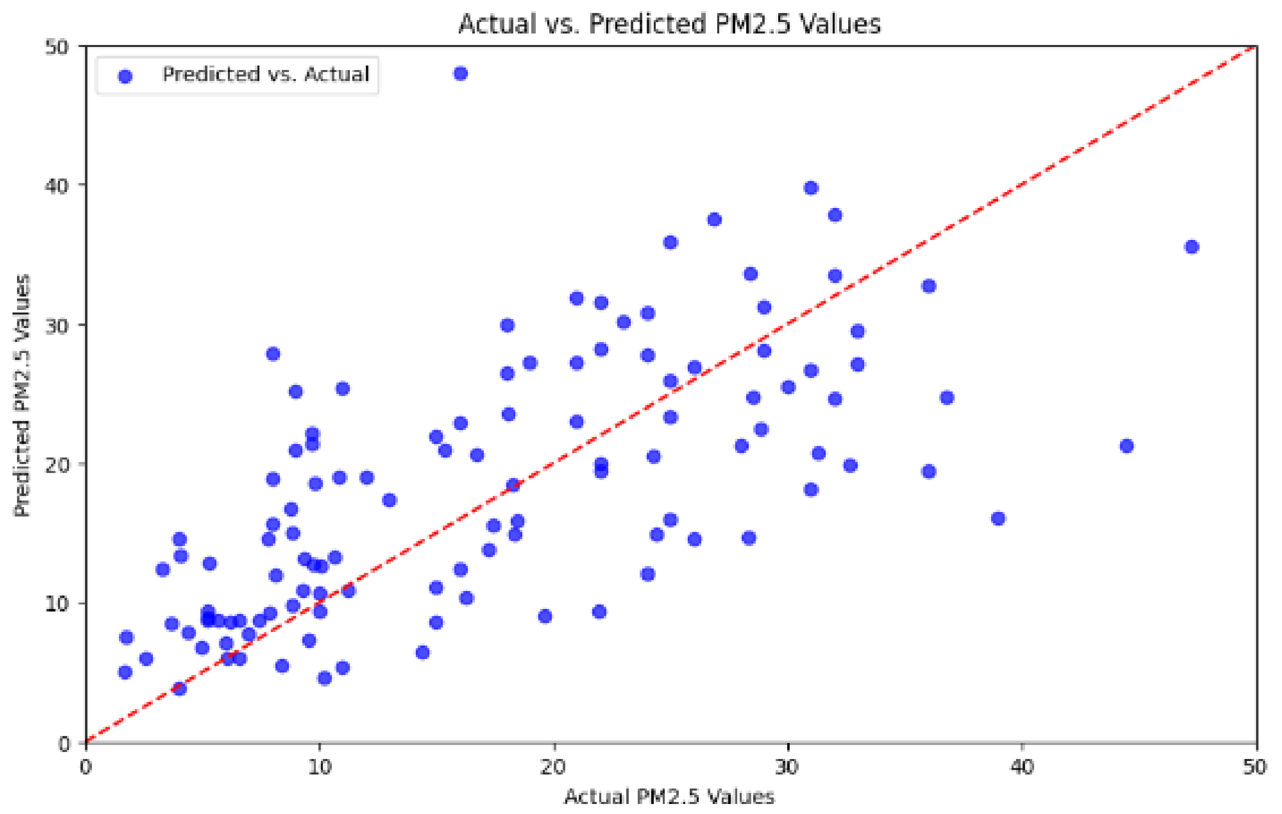
According to MSE values

Random Forest < Linear Regression < XGbhoost < Prophet

Lover value indicates a better fit in MSE

Training Model - PM 2.5
Evidence of Completion

Random Forest




```
R-squared score: 0.362
Mean Squared Error: 123.266
```

MSE - 123.26

Suppor Vector Regressor

```
from sklearn.svm import SVR
from sklearn.metrics import mean_squared_error

# Create and train the SVR model
svr_model = SVR(kernel='rbf')
svr_model.fit(X_train, y_train)

# Make predictions on the test set
y_pred_svr = svr_model.predict(X_test)

# Calculate the Mean Squared Error
r2 = r2_score(y_test, y_pred_svr)
print(f"R-squared score: {r2:.3f}")

mse_svr = mean_squared_error(y_test, y_pred_svr)
print(f"Mean Squared Error for SVR: {mse_svr:.3f}")
```

```
R-squared score: 0.316
Mean Squared Error for SVR: 132.065
```

RMSE - 132.065

Lover value indicates a better fit in MSE

Functional, Non-Functional and Personnel Requirements

Functional Requirements

- Polluted Air Data Collection and Preprocessing(clean and transform)
- Feature Selection and Analysis
- Model Development and Training
- Explore and select the most suitable predictive model based on accuracy metrics.
- Dynamic Predictions for Real-time Adjustments.
- Real-time Pollution Data and Comparison and Visualization.

Non-Functional Requirements

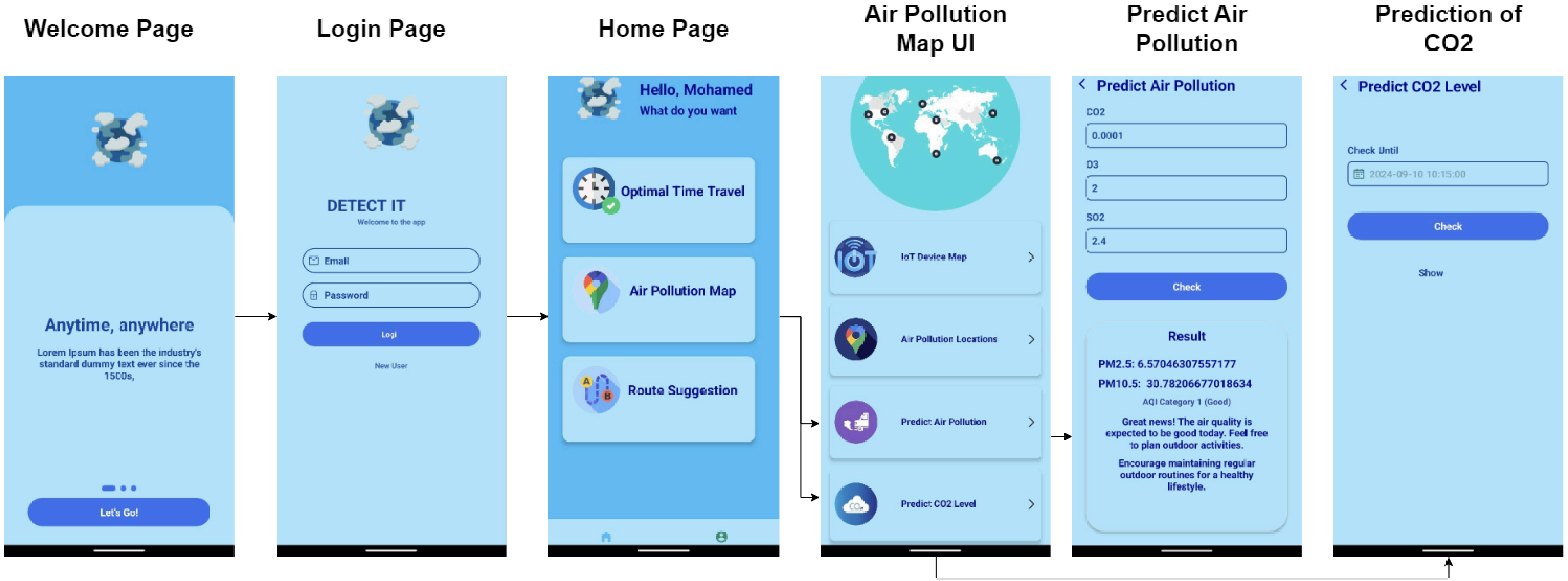
- User-friendly Interfaces.
- Should properly work in cross platforms (for android and IOS devices)
- The application should be reliable.
- Higher accuracy of results.
- Results should be more efficient.

Personnel Requirements

Resources and Dataset Air pollution level in Colombo

- Central Environment Authority (CEA)
- The National Building Research Organization (NBRO).

Evidence of Completion



PM 2.5 threshold values to say air quality is good or bad?

0 ≤ PM2.5 ≤30: AQI Category 1(Good)

31≤ PM2.5 ≤ 60: AQI Category 2(Satisfactory)

61 ≤ PM2.5 ≤ 90: AQI Category 3(Moderate)

91 ≤ PM2.5 ≤120: AQI Category 4(Poor)

121 ≤ PM2.5 ≤250: AQI Category 5(Very Poor)

251 ≤ PM2.5: AQI Category 6(Severe)

Completion and Future works

90%

Completion of the components

- Polluted Air Data Collection and Preprocessing, Cleaning and Transforming.
- Identification of the best Model for predicting CO2 level in the environment - (Sub-Objective 02
- Identification of the best model for predicting PM 2.5 - (Sub-Objective 03)
- Host the finalized model in the fast server
- Displayed the results in mobile application

10%

Future Implementations

- UI Enhancement
- Increasing The Performance
- Bug Fixing

References

[1] S. Mahanta, T. Ramakrishnudu, R. R. Jha, and N.Tailor, "Urban Air Quality Prediction Using RegressionAnalysis," in TENCON 2019 - 2019 IEEERegion 10 Conference (TENCON), Kochi, India,Oct. 2019, pp. 1118–1123. doi: 10.1109/TENCON.2019.8929517.

[2] M. Castelli, F. M. Clemente, A. Popoviˇc, S. Silva,and L. Vanneschi, "A Machine Learning Approachto Predict Air Quality in California," Complexity,vol. 2020, pp. 1–23, Aug. 2020, doi:10.1155/2020/8049504.

[3] S. Zhong, Z. Yu, and W. Zhu, "Study of the Effectsof Air Pollutants on Human Health Based onBaidu Indices of Disease Symptoms and Air QualityMonitoring Data in Beijing, China," IJERPH,vol. 16, no. 6, p. 1014, Mar. 2019, doi:10.3390/ijerph16061014.

[4] Y. L. S. Nandasena, A. R. Wickremasinghe, andN. Sathiakumar, "RAesierarpchoarltilculetion andhealth in Sri Lanka: a review of epidemiologic studies,"p. 14, 2010.

[5] T. Xayasouk and H. Lee, "AIR POLLUTION PREDICTIONSYSTEM USING DEEP LEARNING,"Naples, Italy, Jun. 2018, pp. 71–79. doi:10.2495/AIR180071.

[6] D. Iskandaryan, F. Ramos, and S. Trilles, "AirQuality Prediction in Smart Cities Using MachineLearning Technologies based on Sensor Data: AReview," Applied Sciences, vol. 10, no. 7, p. 2401,Apr. 2020, doi: 10.3390/app10072401.

IT21152832 | Kumari J.M.D

Specialization : Information Technology

Component 3 : Realtime Pollution Heatmap and Risk Exposure Analysis

BACKGROUND

Discover the key aspects of my component

- Collection of air pollutants (CH_4, CO_2, NO_2) using IoT sensors, How real-time visualizations of pollution levels influence user behavior to minimize exposure.
- Analyse how health conditions (such as asthma, COPD) impact personalized risk exposure recommendations.
- Explore how personalized notifications can help reduce exposure for users with specific health conditions.
- Determine the optimal travel times based on current and forecasted air pollution levels

SPECIFIC AND SUB OBJECTIVES

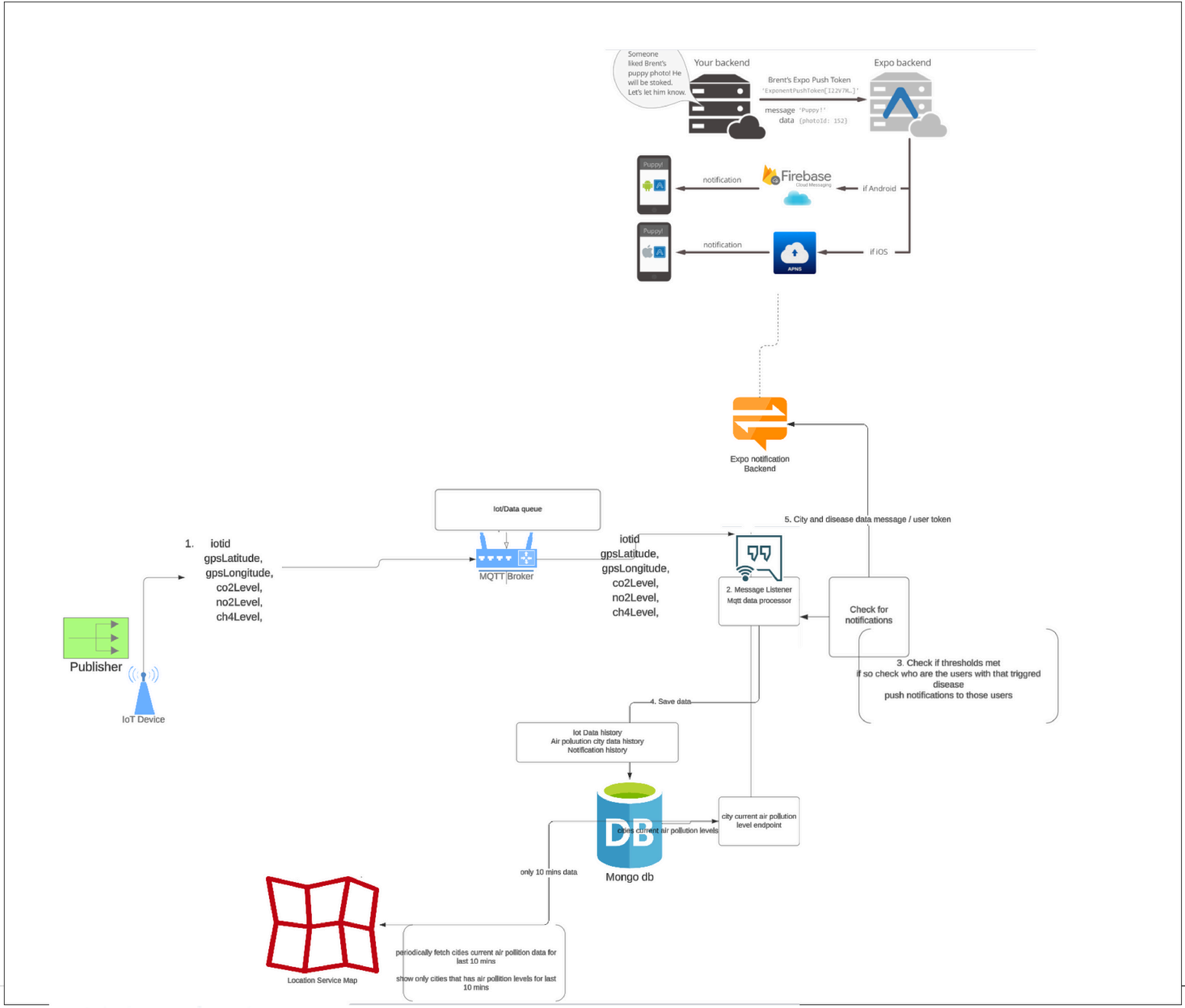
SPECIFIC OBJECTIVE

Develop a real-time air quality heatmap that provides accurate and timely data, allowing users, especially those with health conditions, to receive personalized notifications and take action to reduce their exposure to harmful pollutants.

SUB OBJECTIVE

- **Develop a Real-Time Air Quality Heatmap that visualizes real-time air pollution data, enabling users to see current air quality levels across different cities.**
- **Implement a Real-Time Risk Exposure Analysis and automatically send notifications to users whose health may be impacted by pollution, based on their location and registered conditions.**
- **Develop an Optimal Travel Time Recommendation (the optimal time for users to travel based on current and forecasted pollution levels)**

SYSTEM DIAGRAM



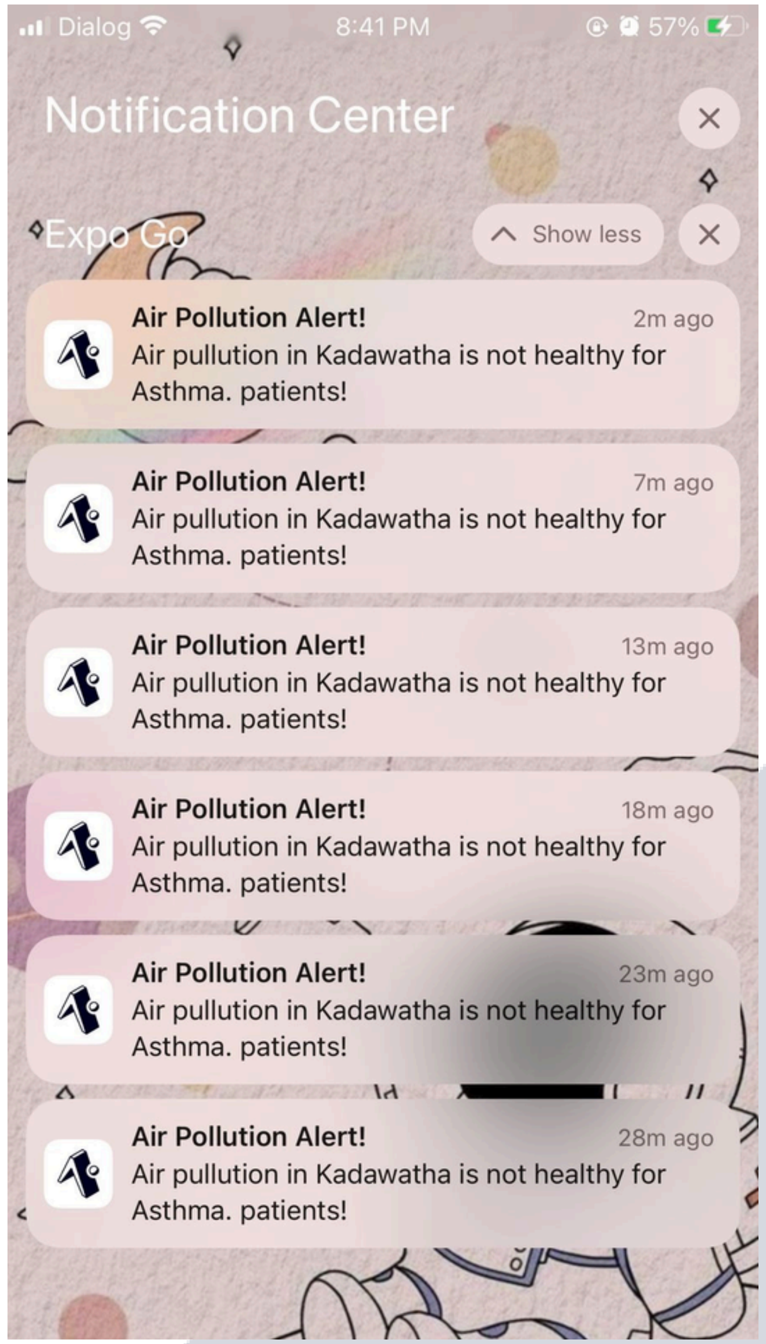
EVIDENCE OF COMPLETION


```
LOG  items:  {"__v": 0, "_id": "6719f1a55aa3f5030e19b7ac", "asthmaStat": false, "averageCH4Level": 0.25284326192307693, "averageCO2Lev
el": 102.09843914538462, "averageNO2Level": 0.06286970615384616, "bronchitisStat": false, "city": "Isurupura", "copdStat": false, "gpsL
atitude": 6.916452408, "gpsLongitude": 79.97393799, "lungCancerStat": false, "timestamp": "2024-10-24T07:05:09.600Z"}
LOG  items:  {"__v": 0, "_id": "6719f14a5aa3f5030e19b791", "asthmaStat": false, "averageCH4Level": 0.2654181601428572, "averageCO2Leve
l": 103.84848513142856, "averageNO2Level": 0.06374706278571429, "bronchitisStat": false, "city": "Isurupura", "copdStat": false, "gpsLa
titude": 6.915932178, "gpsLongitude": 79.97341919, "lungCancerStat": false, "timestamp": "2024-10-24T07:03:38.227Z"}
LOG  items:  {"__v": 0, "_id": "6719f12b5aa3f5030e19b785", "asthmaStat": false, "averageCH4Level": 0.2620375262142857, "averageCO2Leve
l": 104.1542914242857, "averageNO2Level": 0.06389985499999999, "bronchitisStat": false, "city": "Isurupura", "copdStat": false, "gpsLat
itude": 6.916482449, "gpsLongitude": 79.97396851, "lungCancerStat": false, "timestamp": "2024-10-24T07:03:07.768Z"}
LOG  items:  {"__v": 0, "_id": "6719f0ef5aa3f5030e19b76a", "asthmaStat": false, "averageCH4Level": 0.26083653366666665, "averageCO2Lev
el": 104.54221700266666, "averageNO2Level": 0.0640928438, "bronchitisStat": false, "city": "Isurupura", "copdStat": false, "gpsLatitude
": 6.915942192, "gpsLongitude": 79.97342682, "lungCancerStat": false, "timestamp": "2024-10-24T07:02:07.011Z"}
LOG  items:  {"__v": 0, "_id": "6719f0d05aa3f5030e19b75e", "asthmaStat": false, "averageCH4Level": 0.26145549013333336, "averageCO2Lev
el": 104.54221700266666, "averageNO2Level": 0.06409284380000001, "bronchitisStat": false, "city": "Isurupura", "copdStat": false, "gpsL
atitude": 6.916452408, "gpsLongitude": 79.97393799, "lungCancerStat": false, "timestamp": "2024-10-24T07:01:36.563Z"}
```


Capture real-time average pollution levels (CH4, CO2, NO2) from particular location, including geolocation, timestamps, and health risk assessments for conditions like asthma and COPD, with all risks currently marked

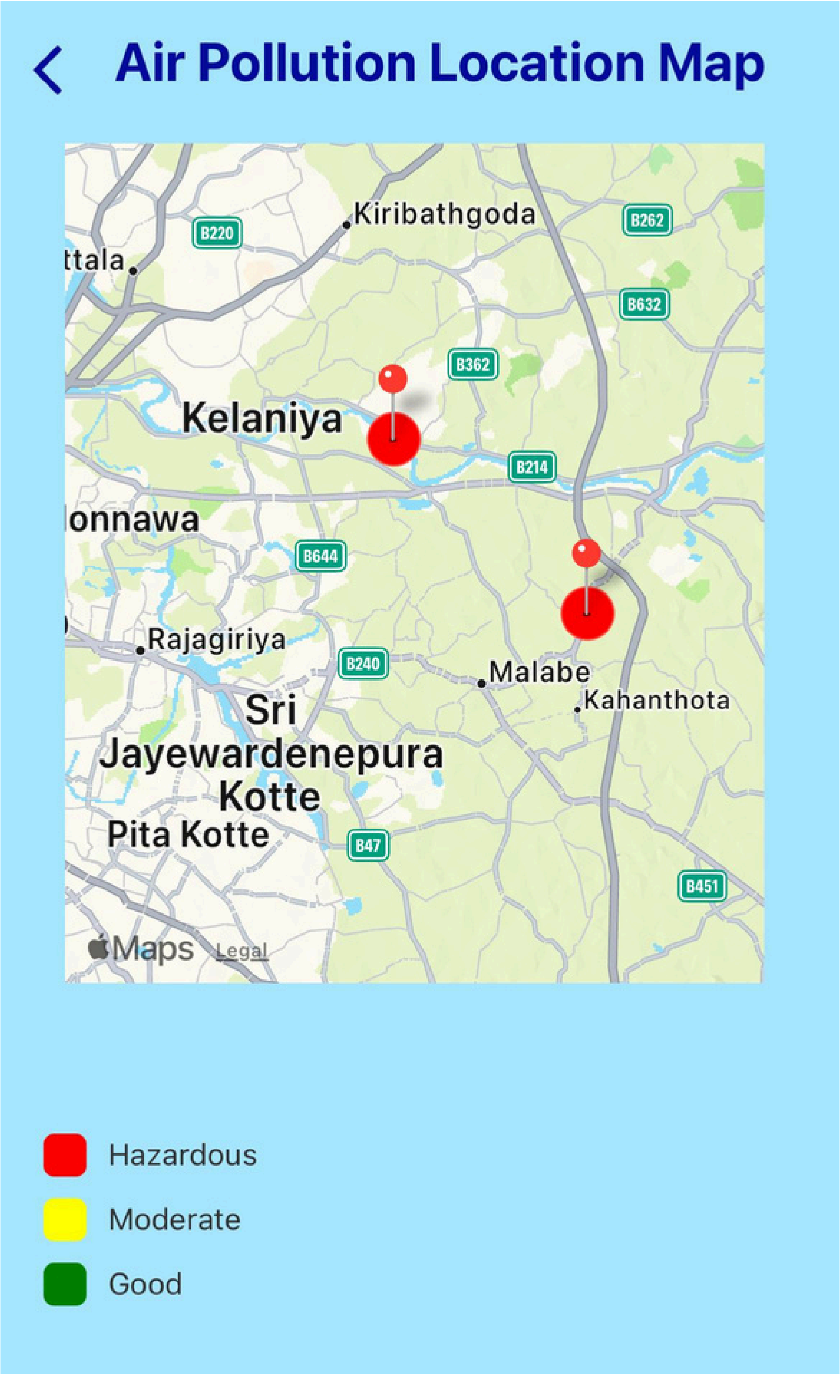
Pollutant	Low Risk (ppm)	High Risk (ppm)	Related Diseases
SO_2	$\leq$ 0.015 ppm	$\geq$ 0.048 ppm	Asthma, COPD, Bronchitis
NO_2	$\leq$ 0.013 ppm	$\geq$ 0.087 ppm	Asthma, COPD, Lung Cancer, Heart Disease, Bronchitis
CO	$\leq$ 3.50 ppm	$\geq$ 25.29 ppm	Asthma, Lung Cancer, Heart Disease
CO_2	$\leq$ 1000 ppm	$\geq$ 5000 ppm	Asthma, COPD, Lung Cancer, Heart Disease, Bronchitis

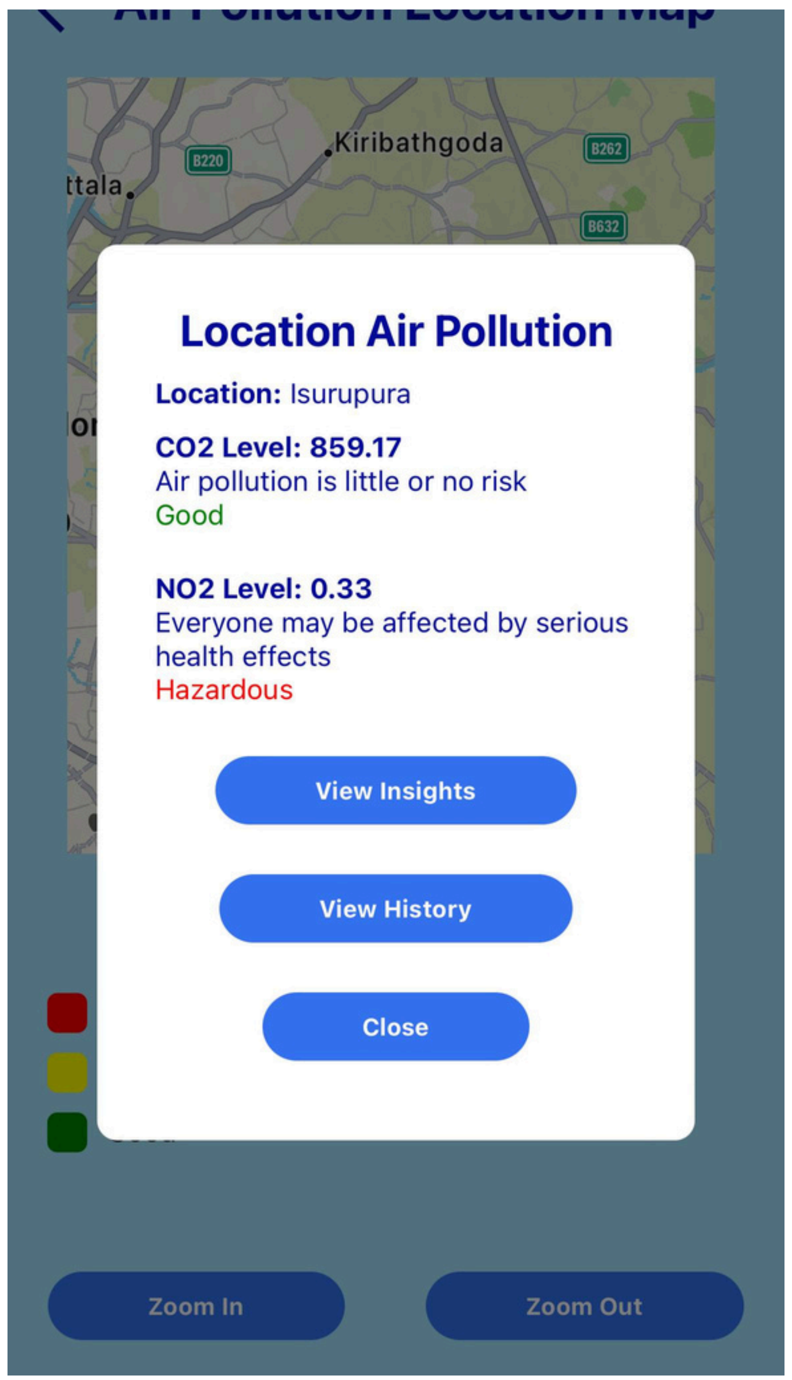
GENARALIZED HEALTH DISEASES AND RISK CLASSIFICATION BASED ON DIFFERENT POLLUTION LEVELS

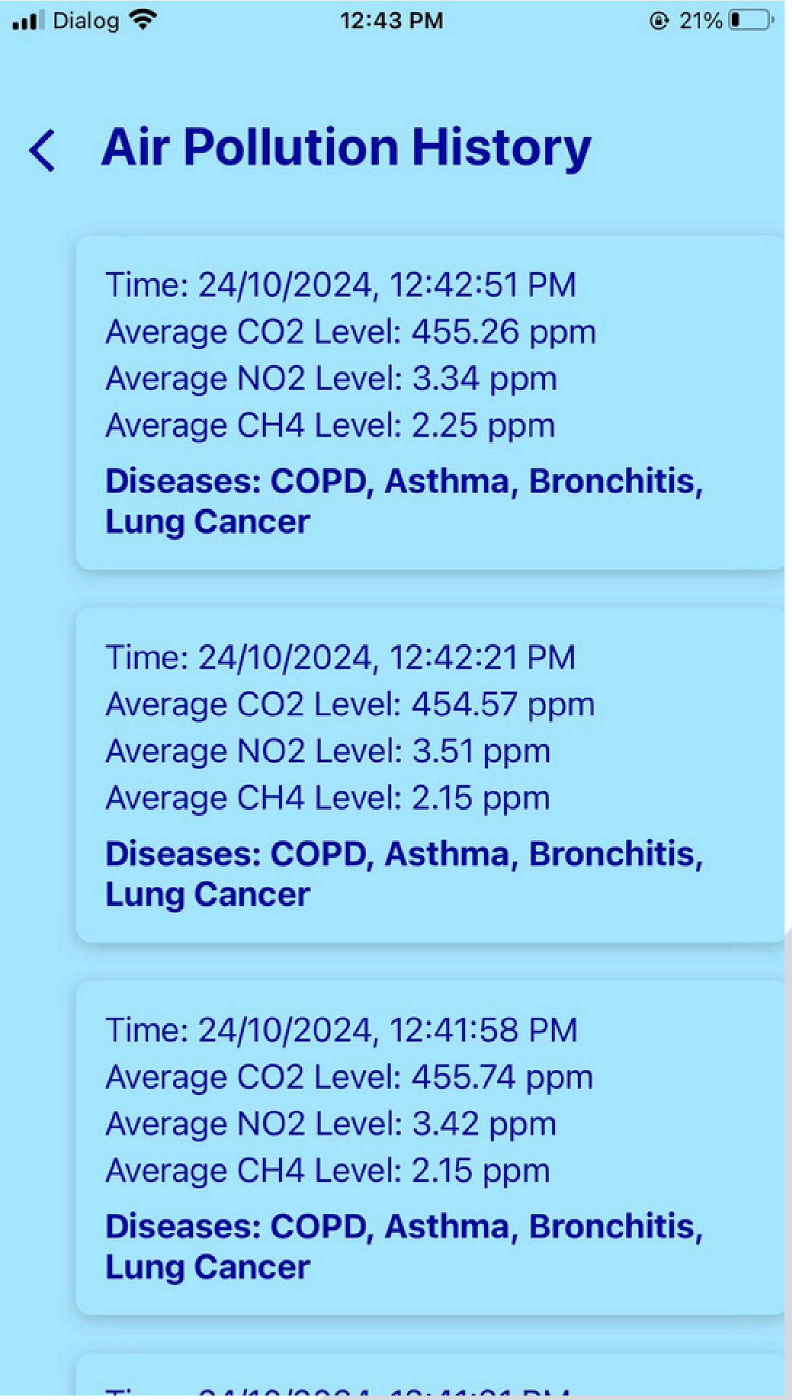


REALTIME PUSH NOTIFICATIONS

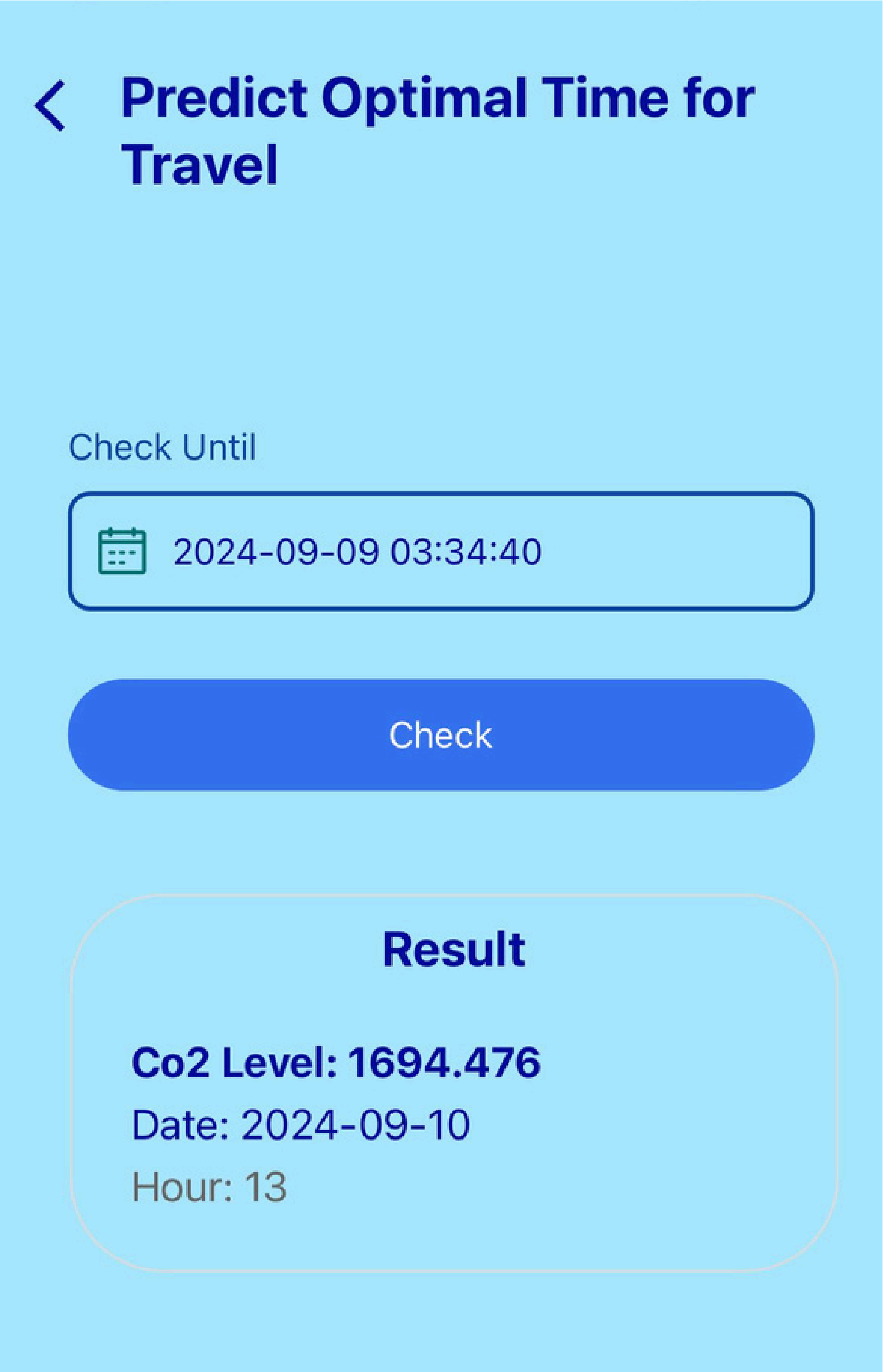
FRONTEND IMPLEMENTATION

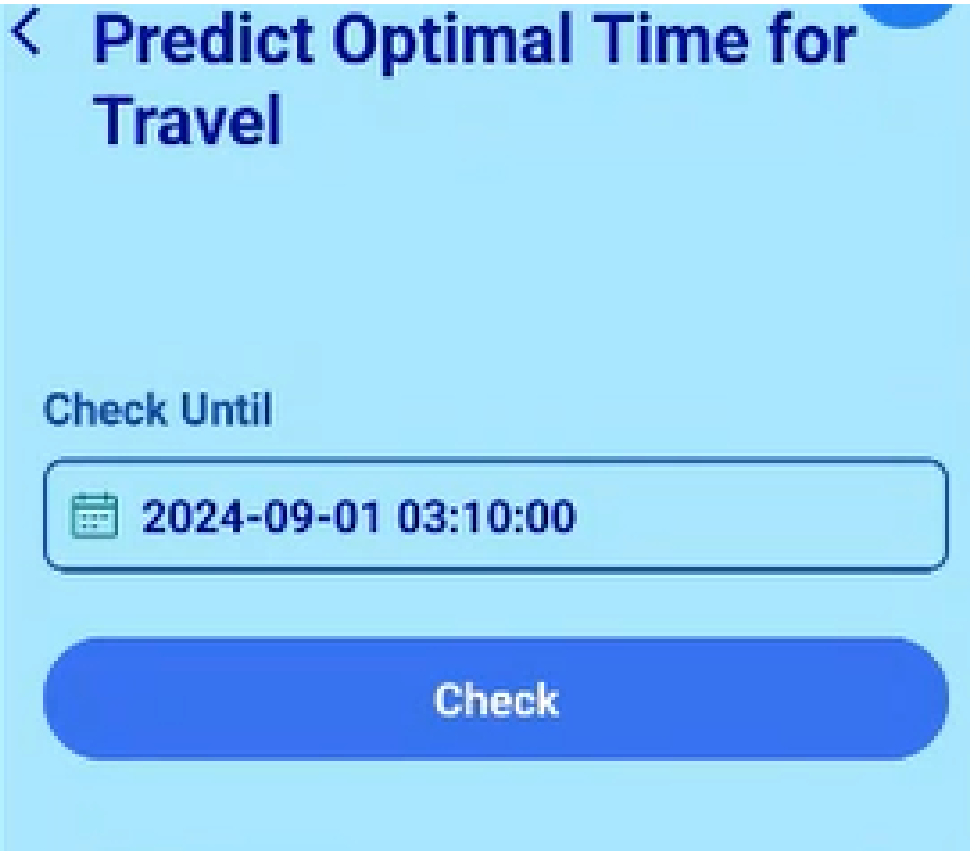


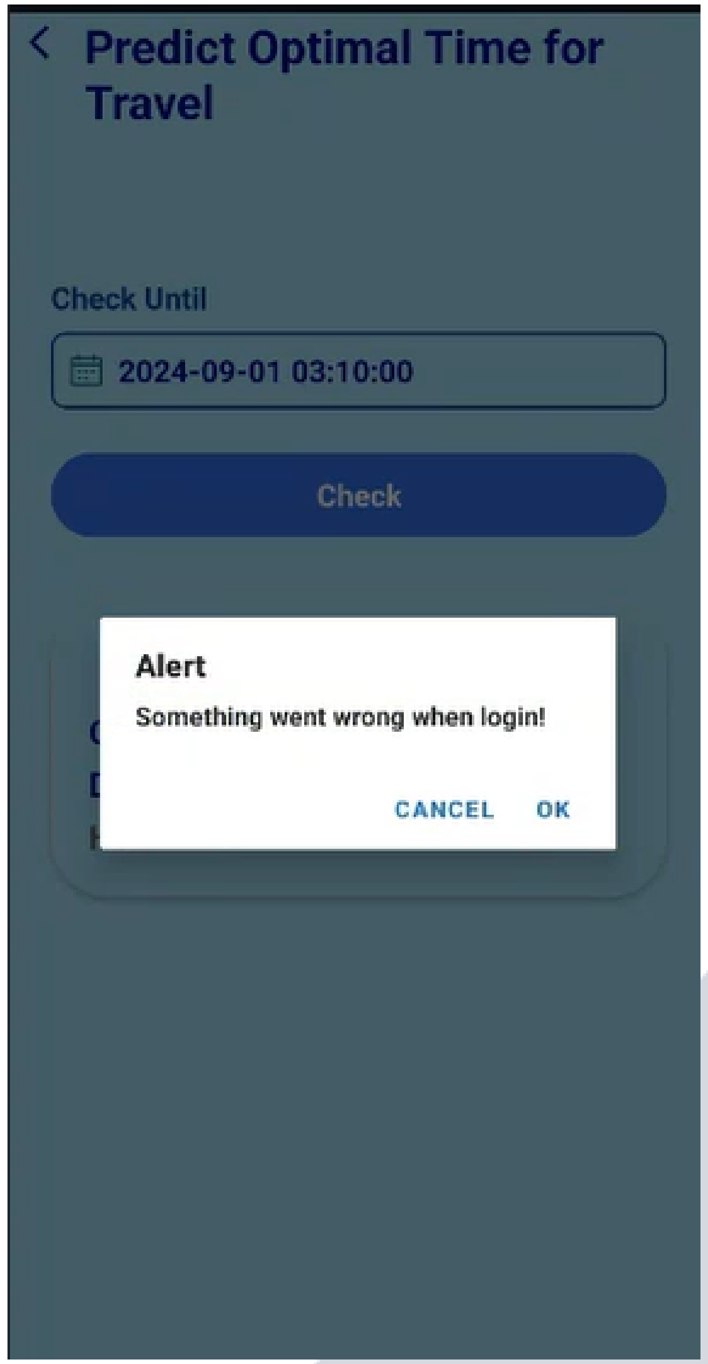




FRONTEND IMPLEMENTATION



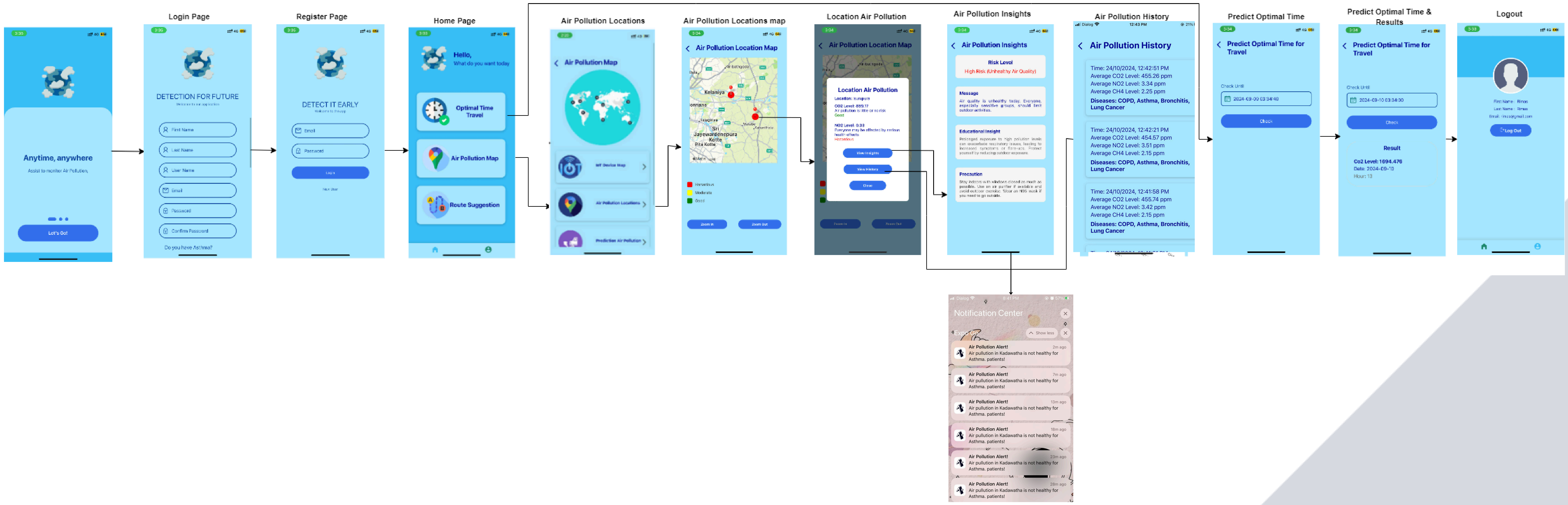




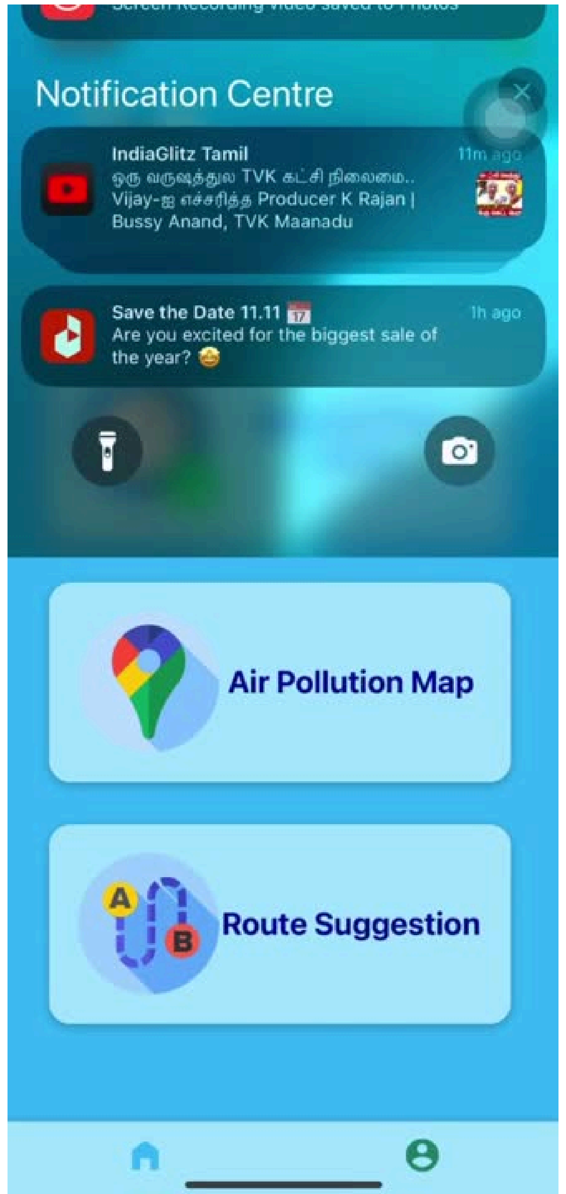

```
LOG  OptimalTimeResultCard
LOG  {"__v": 0, "_id": "6719f2745aa3f5030e19b8cc", "co2Level": 1783.7487552262141, "day": "2024-10-24", "hour": 13}
```

Capture the Lesser CO2 level until the End date and provide the Optimal time for the travel

Display the Final Results in Application

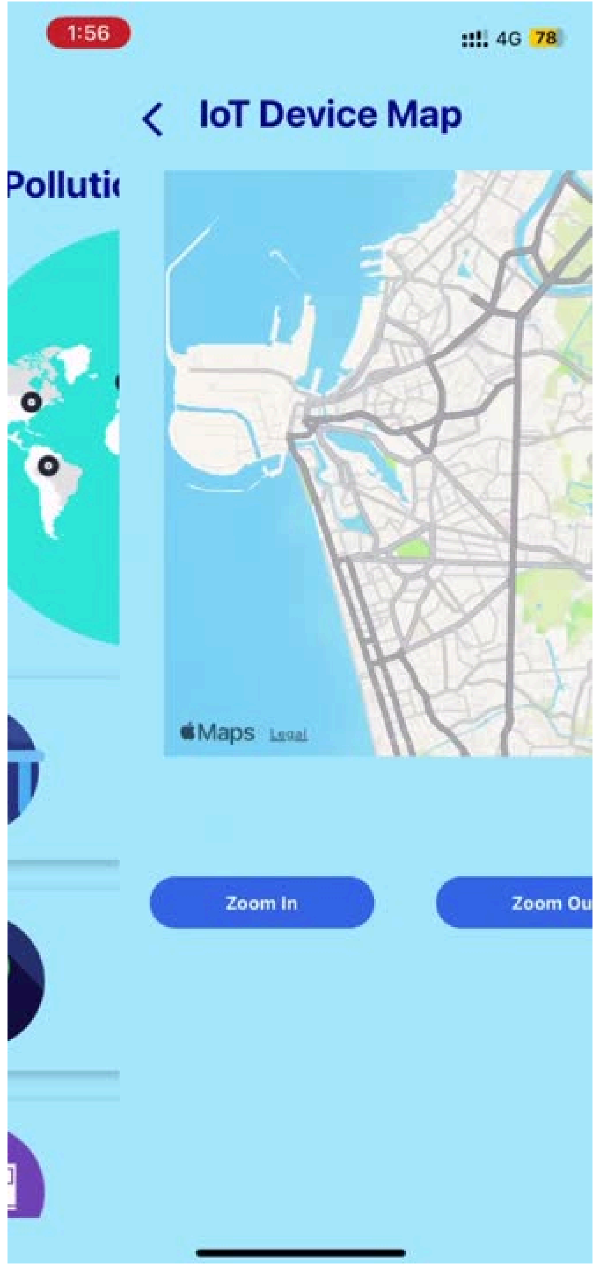


DEMO VIDEOS



Moments captured when Risk level is HIGH
and PUSH NOTIFICATIONS

DEMO VIDEOS



Location when Air pollution is high(HAZARDOUS)

Location when Air pollution is medium(MODERATE)

Location when Air pollution is low(LOW)

DEMO VIDEOS

if registered user is an asthma patient,push notifications in high exposure locations related to asthma

if registered user is a lung cancer patient,push notifications in high exposure locations related to lung cancer

if registered user is a COPD patient,push notifications in high exposure locations related to COPD

Completion of the component

- Display Air levels for CH4, CO2, and NO2 gases values on the heatmap based on different locations, show the history of air pollution and insights of air pollution
- Map integration with Google Maps API
- Data Visualization and Mapping in mobile application
- Connect with Server backend to integrate with Realtime pollution levels and data
- Recommendation of optimal travel time using current and forcasted air quality data
- Backend implementation of Risk exposure level, espcially those affected by health diseases
- Push Notification
- UI enhancement
- Increasing Performace
- Bug Fixing
- Testing & Validation

References

[1]Castell, N., Dauge, F.R., Schneider, P., Vogt, M., Lerner, U., Fishbain, B., Broday, D., and Bartonova, A. (2017). "Can commercial low-cost sensor platforms contribute to air quality monitoring and exposure estimates?" Environment International, 99, 293-302.

[2]Zhang, W., Vazquez-Canteli, J.R., and Sadiq, R. (2016).” A real-time air quality routing algorithm: Reducing human exposure to outdoor air pollution. “69, 281-300.

[3]Tsai, J.H., Lv, N., and Wang, Y. (2016). “A heatmap-based visualization framework for air pollution monitoring data. Journal of Ambient Intelligence and Humanized Computing,” 7(4), 499-512.

[4]“What are the WHO Air quality guidelines?,” WHO newsroom, Sep.22, 2021. [Online]. Available:https://www.who.int/news-room/feature-stories/detail/what-are-the-who-air-quality-guidelines

[5]A. Manuel, G. F. Sathyraj, R. C. Joseph, S. A. Philip, and S. M. Thomas, "AI-Integrated IoT-Enabled Smart Mask For SoS Alerting And Disease Prediction Based On Air Pollutants," in Proc. 2023 Int. Conf. Signal Process., Comput., Electron., Power, Telecommun. (IConSCEPT), 2023, pp. 1-6

IT21057588 | Pasan M.G.R

Specialization : Information Technoogy

Component 4 : Route Generation Based on pollution

Background

.

- Proposing the least air pollution shortest route when there are several routes to go between two locations within the Colombo district.

- A new algorithm has been developed by combining Dijkstra's algorithm and the dynamic mapping process to suggest the best way.

- Finally, the suggested shortest path and pollution levels for the route are sent to the Mobile application to display it on the map.

Research Questions

How data service for accessing and processing road network information.

How is the shortest route identified?

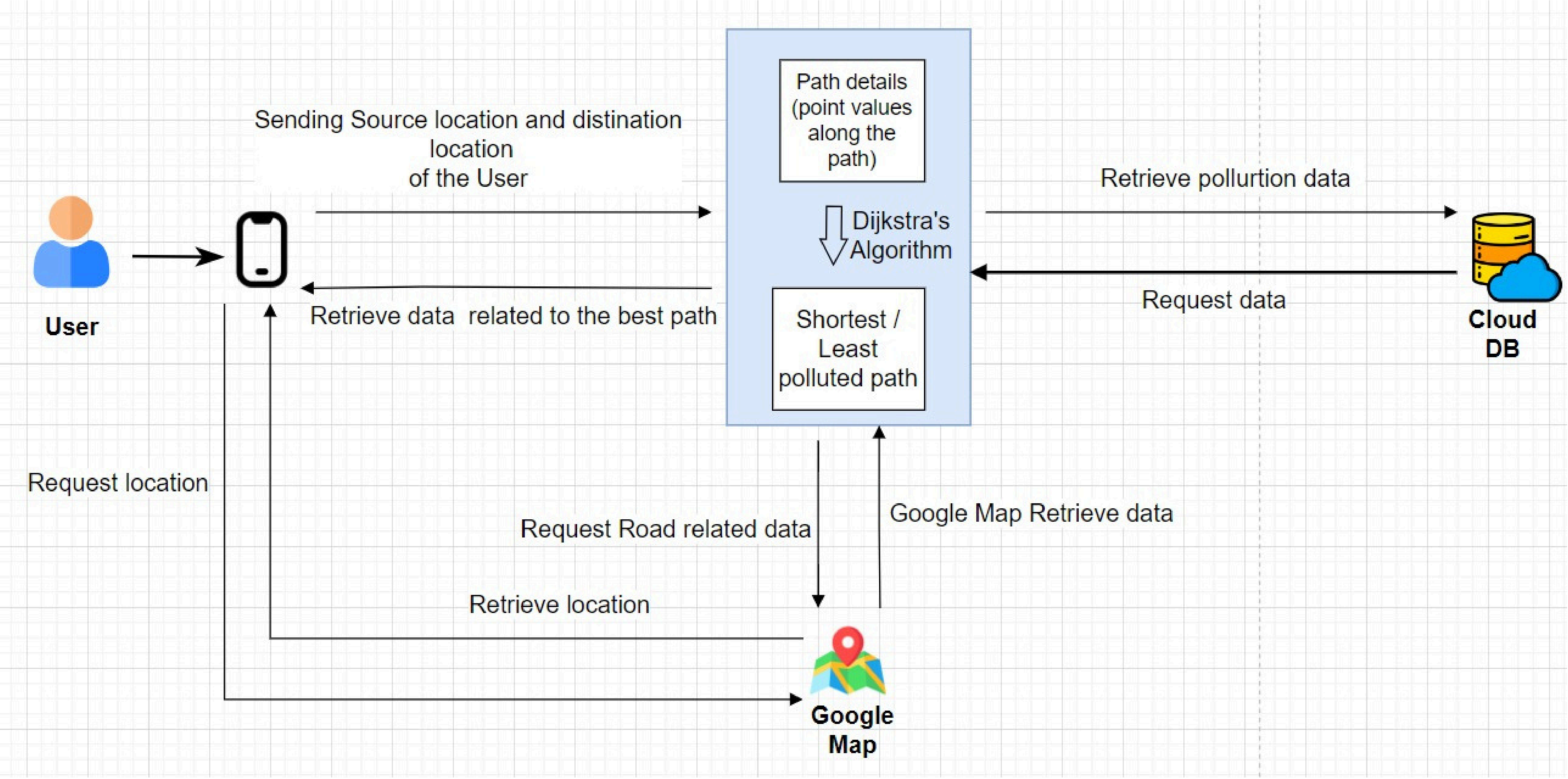
How is the Pollution Level Considered in the Route Map

Specific and Sub Objective

Route Generation Based on Pollution

- Identification of Polluted Route Areas
- Writing algorithm to calculate the near distance level low polluted route
- Finding best route with low pollution level

System Diagram



02

Used Techniques and Technologies

Technologies

- DataCollection
- Algorithm tuning

Algorithm

- Dijkstra's Algorithm

Techniques

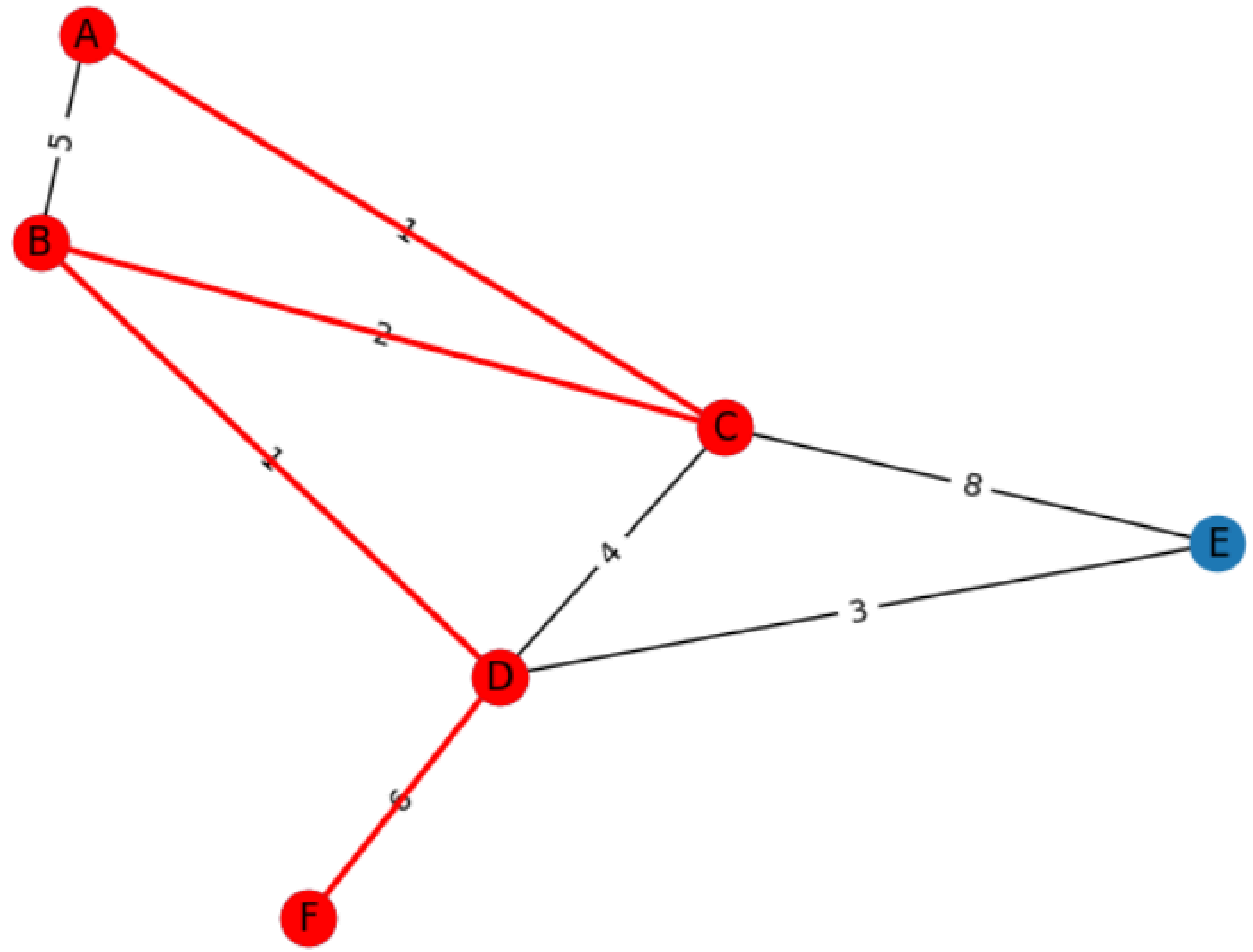
- React Native
- Python
- Fast API
- Node JS
- Google Collab
- VS code
- MongoDB

Methodology

Evidence of Completion

- Data collection
- Host the finalized model in the fast api
- Model Evaluation
- Deviloping backend calling Api
- Frontend development and Display the result in mobile application
- Model Testing

Output Graph For Dijkstra's Algorithm

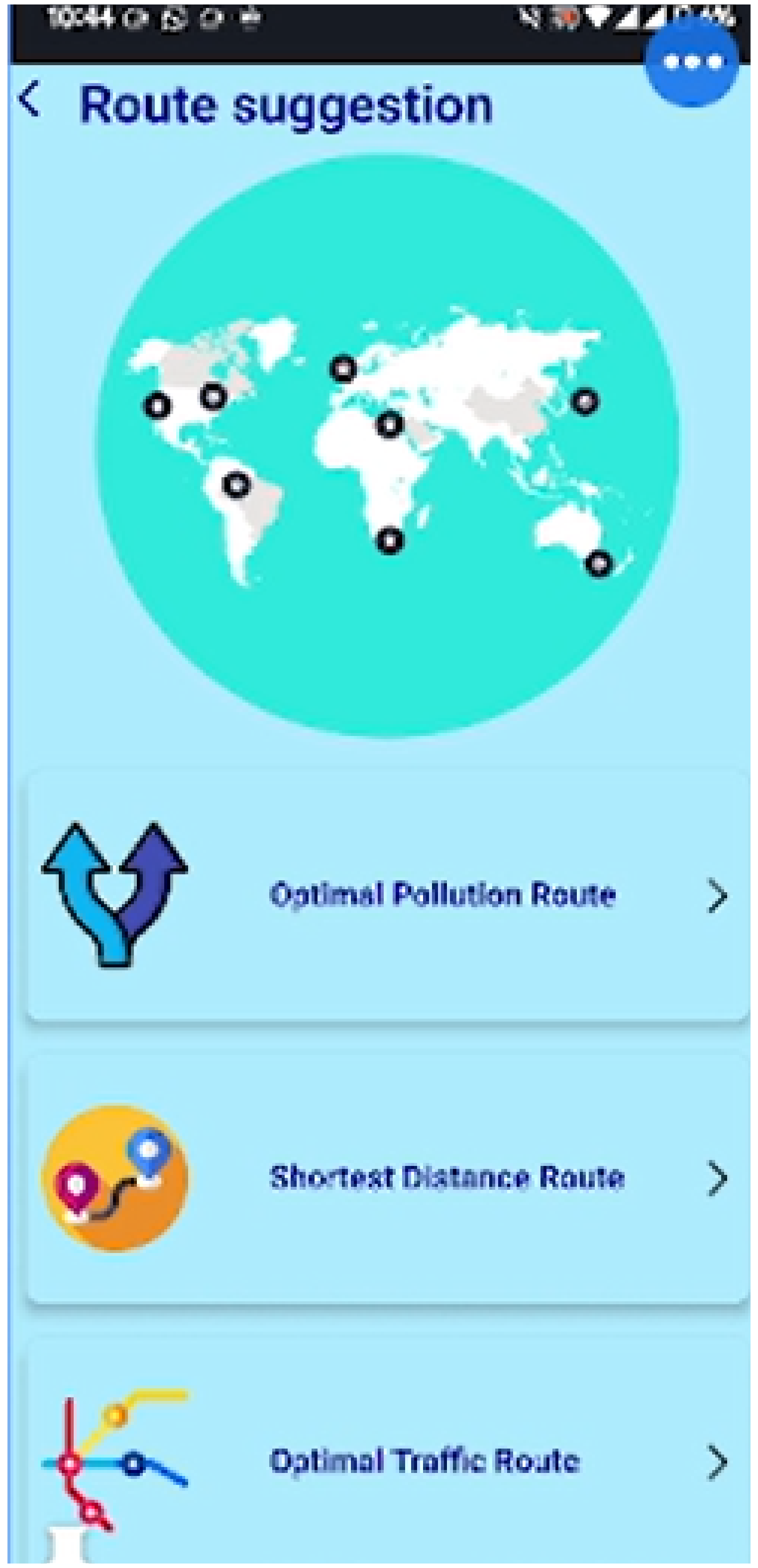


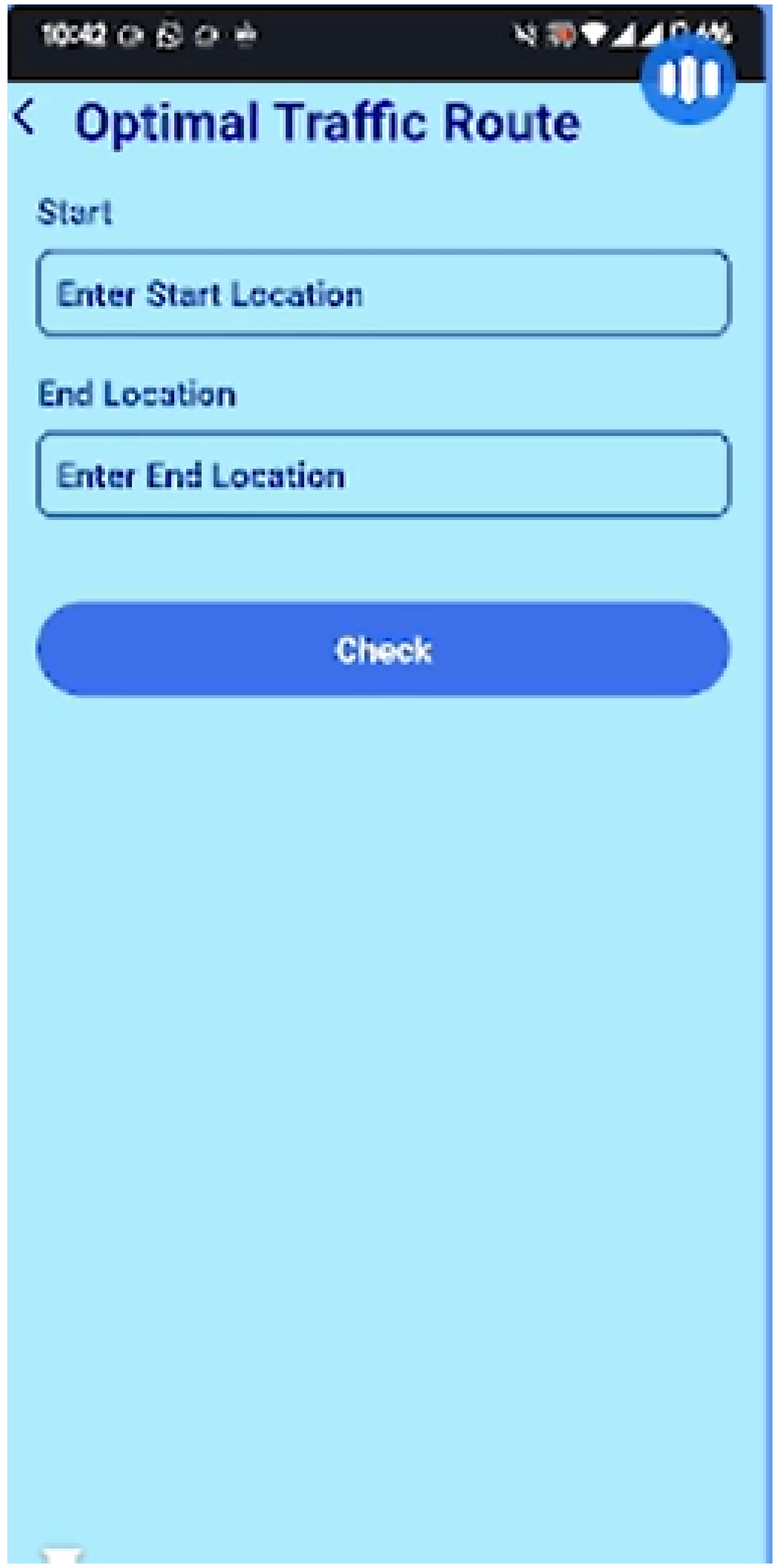
As example

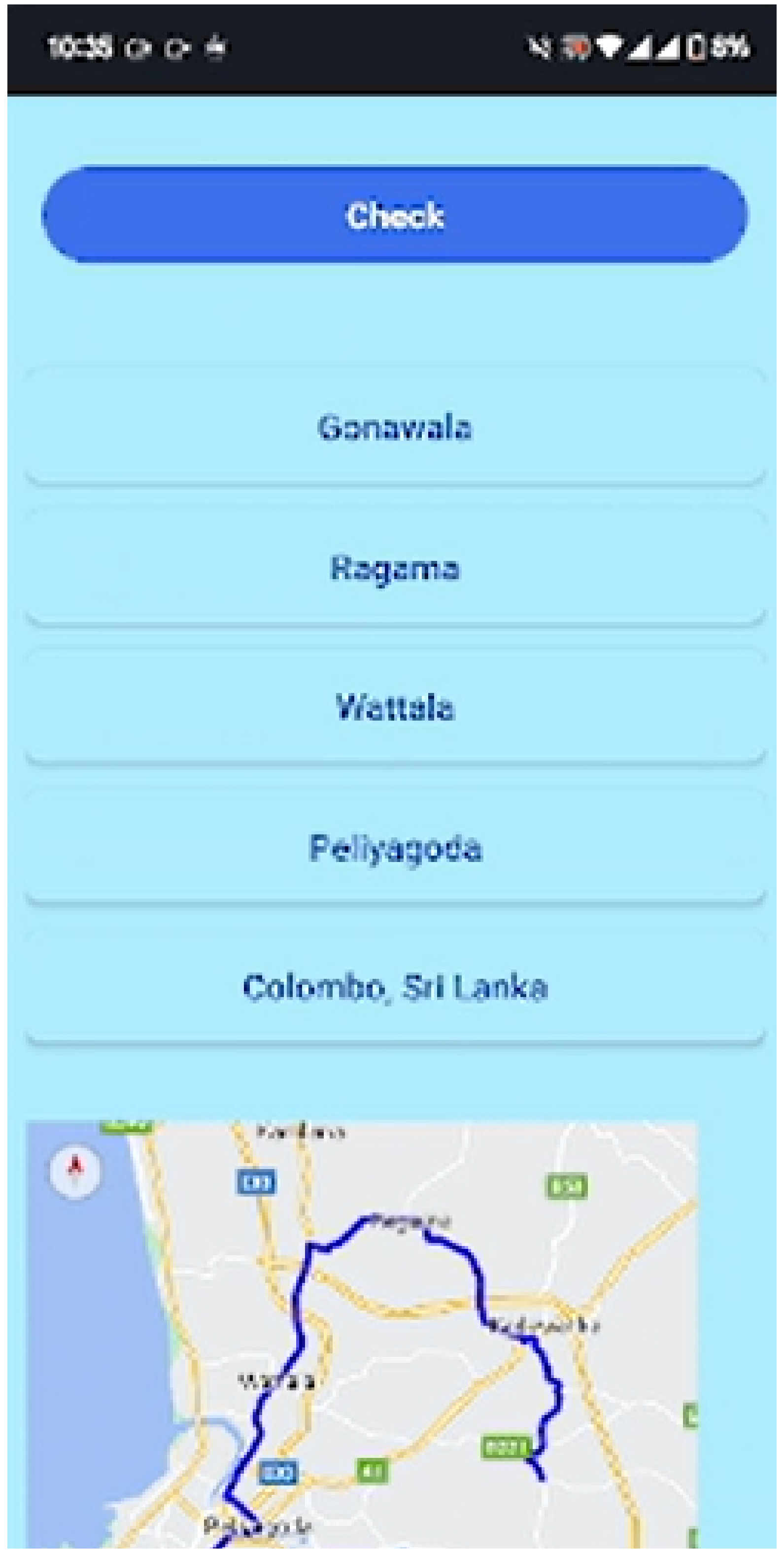
$A \xrightarrow{1} C \xrightarrow{4} D \xrightarrow{6} F = 11$

$A \xrightarrow{5} B \xrightarrow{3} D \xrightarrow{6} F = 14$

Display the result in Application







02

System, Personnel, and Software Specification Requirements

Functional Requirements

- When there are multiple routes between any two locations, display the non-polluted route on the map.
- Algorithms should be used to select the path with the least pollution level.
- The pollutant concentration data should be displayed above the minimum path of the Pottution level .

Non-Functional Requirements

- Interfaces should be User-friendly
- Should properly work for android and IOS devices
- The application should be reliable
- Higher accuracy of results
- Results should be more efficient

90%

Completion of the components

- path visulization using dijkarst Algoridum.
- The shortest distance route is shown on the map.
- The Optimal trapic route is shown on the map.
- The optimal pollution route is shown on the map.

10%

Future Implementations

- UI Enhancement
- Bugs Fixing
- Increasing the performance

References

- Wang, Z., Novack, T., Yan, Y. and Zipf, A., 2020. Quiet route planning for pedestrians in traffic noise polluted environments. IEEE Transactions on Intelligent Transportation Systems, 22(12), pp.7573-7584.

- Yang, J., Cai, B., Li, X. and Ge, R., 2023, May. Optimal path planning for electric vehicle travel time based on Dijkstra. In 2023 35th Chinese Control and Decision Conference (CCDC) (pp. 721-726). IEEE.

- Zheng, Z., Yao, S., Li, G., Han, L. and Wang, Z., 2023. Pareto Improver: Learning Improvement Heuristics for Multi-Objective Route Planning. IEEE Transactions on Intelligent Transportation Systems.

- Jayalath, K.G., Deeyamulla, M.P. and de Silva, R.C.L., 2023. An Assessment of Transboundary Pollution from Colombo to Kandy on the Atmospheric Deposition of Heavy Metals Using Moss (Hyophila involuta). Sri Lankan Journal of Applied Sciences, 2(01), pp.39-43.

02

Thank You !