AIR QUALITY MONITORING AND ANALYSIS BASED PREDICTIVE SYSTEM

2024-078

Project Proposal Report Mohamed Ismail Mohamed Inthikhaff

B.Sc. (Hons) Degree in Information Technology Specialized in Information Technology

Department of Information Technology

Sri Lanka Institute of Information Technology Sri Lanka

February 2024

AIR QUALITY MONITORING AND ANALYSIS BASED PREDICTIVE SYSTEM

2024-078

Project Proposal Report

Mohamed Ismail Mohamed Inthikhaff – IT21058028

Supervisor: Chathurangika Kahandawaarachchi

B.Sc. (Hons) Degree in Information Technology Specialized in Information Technology

Department of Information Technology

Sri Lanka Institute of Information Technology Sri Lanka

February 2024

DECLARATON

I hereby declare that this is my original work, and that, to the best of my knowledge and belief, it does not contain any previously published or written works by other people, unless such acknowledgement is made within the text. Nor does this proposal contain any previously submitted work for credit toward a degree or diploma from any other university or institute of higher learning.

Name	Student ID	Signature
M. I. M Inthikhaff	IT21058028	thursdrage

Signature of the Supervisor	Date	
(Ms. Chathurangika Kahandawaarachchi)		
	29/02/2024	

ABSTRACT

The well-being of the population of Sri Lanka is at risk due to the severe effects of air pollution on daily living, especially in major cities like Colombo and Kandy. The Air Quality Index measures how threatened the air is, which is essential to human survival (AQI). The primary offender, vehicle emissions, are in the spotlight and require immediate action as well as efficient monitoring methods. In light of this, this study explores the necessity of establishing a comprehensive air quality monitoring network that encompasses all of Sri Lanka. A forward-thinking predictive approach that uses machine learning to evaluate and forecast air quality levels in particular locations is the suggested remedy. With a focus on PM2.5 particle content, a crucial air quality indicator, the study attempts to determine the most accurate machine learning technique for forecasting the Air Quality Index, zooming in on Colombo.

The method uses four related air pollutant concentrations—SO2, NO2, PM2.5, and PM10—to predict the PM2.5 concentration in Colombo. To guarantee prediction accuracy, the dataset is subjected to thorough pre-processing. The goal is to develop a powerful prediction system that can forecast future air quality levels in addition to identifying the state of the air today. The ultimate objective is to reduce the health risks posed by air pollution and promote a healthier living environment for Sri Lankans.

Keywords: Air Quality Index, PM2.5 Concentration, Machine Learning, Ambient Air Quality, Pollution Monitoring.

Table of Contents

DECLARATION	
ABSTRACT	ii
LIST OF FIGURES	iv
LIST OF TABLES	
LIST OF ABBREVIATIONS	
1. INTRODUCTION	1
1.1 Background Literature	1
1.2 Research Gap	5
1.3 Research Problem	7
2. OBJECTIVES	9
2.1 Main Objectives	9
2.2 Specific Objectives	9
3. METHODOLOGY	10
3.1 System Architecture Diagram	11
3.2 Project requirements	12
3.2.1 Functional Requirements and Non-Functional Requirements	12
3.2.2 Software Requirements	13
3.2.3 Personnel Requirements	13
3.3 Software Solution	14
3.3.1 Requirement Gathering and Analysis	14
3.3.2 Feasibility Study	15
3.3.3 Design	16
3.3.4 Implementation	19
3.3.5 Testing (Track and Monitor)	20
4. GANTT CHART	21
5. DESCRIPTION OF PERSONAL AND FACILITIES	22
6. BUDGET AND BUDGET JUSTIFICATION	23
6. COMERCIALIZATION	24
6.1 Target Audience and Market Space	24
REFERENCES	25
APPENDICES	26

LIST OF FIGURES

Figure 3. 1 System Architecture Diagram	10
Figure 3. 2 Agile Methodology [9]	13
Figure 3. 3 Overall Use case Diagram	16
Figure 3. 4 Sequence Diagram	17
Figure 3. 5 Work Breakdown Structure	18
Figure 3. 6 High fidelity diagrams for information sharing to crowd	19
Figure 3. 7 Gantt Chart	21

LIST OF TABLES

Table 1.1 Comparison with existing applications	6
Table 3. 1 Functional and Non-Functional Requirements	12
Table 5. 1 Budget of the crowdsourcing component	23

LIST OF ABBREVIATIONS

PM	Particulate Matters
CEA	Central Environmental Authority
NBRO	National Building Resource Organization
AQI	Air Quality Indexing
API	Application Program Interface
CNN	Convolutional Neural Network
DB	Database
AI	Artificial Intelligence
IT	Information Technology

1. INTRODUCTION

1.1 Background Literature

Poor air quality affects about two thirds of the world's population. Air quality, or the quality of the air we breathe, is a highly valued commodity that everyone should be able to buy. The majority of people on the planet live in heavily air-polluted urban areas—over half, according to the most recent reports from the World Health Organization—but almost none of them are aware of how clean the air is where they live. The quality of the air in Colombo indicates that air pollution in the city is increasing at a very rapid annual rate. In the wee hours of the morning and evening, The Colombo Air Quality Index is higher above the 4th category standard set by the World Health Organization. Sadly, most people don't know what the local air quality index (AQI) level is, and as a result, especially in big cities like Colombo, they've become indifferent to the possible health risks that come with air pollution [1].

This work primarily focuses on using machine learning techniques to predict the air quality index, such as regression, support vector machines, k-nearest neighbor, and random forest. Many machine learning models have been tried and implemented in order to identify the best accurate model. Improving methods for forecasting the air quality index, deepening our knowledge of the index, and appreciating the effects of low air quality are the main objectives of this study.

In Colombo, the average unhealthy level is always the AQI score. With the aid of some reliable and precise AQI category predictions, the people of Colombo can take the necessary preventive measures, such as increasing indoor activities and decreasing outside activities, to protect themselves from the detrimental impacts of the city's poor air quality. This chapter begins with a brief explanation of the AQI, the rationale behind the inquiry, and some background data on the state of the air in Colombo before presenting the research's goal [2].

We must take action to lessen air pollution and ensure that everyone has access to clean air because human activity has led to an increase in global air pollution. The main causes of air pollution on Earth are human-caused revolutions like industrialization. The rising number of cars and other gear that emits carbon dioxide into the sky is another factor causing air pollution. Broadly speaking, air pollution is the result of large-scale releases of dangerous gases and chemicals into the atmosphere that contaminate the air in a particular geographic area. Emissions from production, transportation, and industry are the main sources of poor air quality. When compared to other rural areas, densely populated urban places like Colombo usually have the lowest air quality, mostly because of human activity. Threats to human health, including both immediate and long-term adverse effects on the health of living things and their surroundings, are strongly correlated with the value of the air quality index. When exposed to contaminated air, those who have previously had illnesses like asthma and pneumonia are more likely to develop heart- and lung-related conditions.

It has been noted that after breathing in contaminated air, the human immune system finds it extremely difficult, if not practically impossible, to self-purify PM2.5 and PM10 particles that have entered the body [3]. As several studies have shown, the main problem in Colombo is that people are unaware of the bad quality of the air. As a result, the air quality index (AQI) has become increasingly important in predicting and educating the public about the possible effects that air pollution (AQI levels) may have on human health and the environment, especially in light of the recent substantial increase in air pollution [4]. Furthermore, the negative consequences of air pollution on the ecosystem and human health can be significantly reduced or eliminated by anticipating the AQI value. Finding the best approach to predict the AQI value among the many approaches and the dataset that can be utilized to provide the most accurate forecast are two of the study's main accomplishments. The basis of human existence is air.

The Air Quality Index is the index value that most accurately represents the state of the air right now. The atmosphere and the wellbeing of all living things will be even more at risk from a substantial AQI. For this study, the Central Environment (CEA) Authority provided a historical air concentration dataset with hourly concentration levels of various air pollution factors & weather factors, such as PM10, PM2.5, SO2, NO2, CO, O3, wind speed, wind direction, average temperature, relative humidity, and solar radiation.

Current research on variations in air quality indicates that there have been some recent trends in Battaramulla and Kandy that negatively impact human life. The data indicates that the recent infrastructure and transportation work has had a major influence in both locations. Tiny particles are added to the local atmosphere as a result. As a result, it is evident that PM10PM10 is a significant factor that brings up several concerns that have a detrimental effect on quality of life.

This study was initiated as a first step in aiding the process of preserving air quality for the proper level of both natural and human quality of life. In preparation for a review of earlier studies on air quality forecasting, this work is the first to provide a comparative analysis of PM10 forecasting utilizing cutting-edge models [5].

Thus, this study analyzes the analysis of several machine learning models to forecast the PM10 concentration of two urban sites, namely Battaramulla in Colombo and Kandy. The concentration levels of PM10 were predicted using meteorological variables such as ambient temperature, relative humidity, solar radiation, rainfall, wind speed, and wind direction in addition to other air pollution components like the concentrations of O3, CO, NO2, SO2, and PM2.5. The comparison analysis was carried out by evaluating a number of performance indicators, including Mean Absolute Relative Error (MARE), coefficient of determination (R2), mean absolute error (MAE), mean squared error (MSE), and relative absolute error (RAE). This study thus extends our understanding of machine learning techniques for PM10 concentration forecasting with an emphasis on Sri Lanka (Mampitiya et al., 2023). The outcomes would be taken into consideration for formulating policies that ensure greater living conditions in urban areas [5].

The method for developing a machine-learning technique for PM10 value forecasting was suggested by this study. To put it briefly, this study uses eight machine learning models in conjunction with a special technique. The extensive literature review makes it evident that a specific methodology is needed for this study of environmental factors because a variety of factors are involved. This study recommended appropriate methods for processing data, such as data cleaning, correlation detection, and the use of machine learning models with loop-based hyperparameter optimization, which should be followed by a comprehensive comparative analysis.

Numerous techniques for data preparation were employed in order to ensure the accuracy of the forecast outcome. Using 20% of the preprocessed dataset for testing and the remaining 80% for model training allowed the cross-validation method to be used. Some of the machine learning approaches that have been applied as prediction models include Support Vector Machine, Random Forest, K Nearest Neighbors, and Multiple Linear Regression. On the basis of accuracy and performance, the optimal machine learning model for AQI prediction is selected.

'Air Pollution Level Prediction and Air Quality Indexing in Colombo' is becoming more and more necessary, according to the readings described above, in order to fill in the gaps in the current processes and systems. Predicting air pollution levels is beneficial for schoolchildren, those who are impacted by air pollution, extension workers, and others since it allows for real-time notifications when the Air Quality Index (AQI) is high. Moreover, the capacity to visualize air pollution levels through the comparison of past, present, and future situations is beneficial and provides a more intelligent way to deal with current problems.

1.2 Research Gap

For AQI prediction, there are several choices. Few of these forecasts, which depend on the atmospheric concentrations of PM10 and PM2.5 fine particles, provide comprehensive suggestions for reducing the detrimental impacts of air pollution on human health. A detailed examination of the literature was carried out in order to provide a concise overview of the research on AQI predictions made using different machine learning models, as there was no precise approach for AQI prediction within the study area. The study topic differs not only in terms of methodologies and strategies but also in terms of the datasets that are accessible, many of which are unique because of the traffic, climate, and environmental factors of the selected geographic area. The primary source of poor air quality is PM particles. For example, air pollutants like PM2.5 and PM10 are the primary cause of air pollution in certain cities, while COx, SOx, and NOx are the main causes of air pollution in other places [6].

Due to these limitations, a thorough review of the literature is done for this study in an effort to comprehend the breadth of AQI prediction research and identify relevant studies that achieve the same objective as this one. The literature review section of this paper compiles the most recent and relevant studies on AQI predictions. Some of the most recent techniques for AQI prediction are looked at here.

It is essential to consider the best approach for air pollution forecasting. The deep learning approach is one of the most often used techniques among the models that are currently in use for AQI prediction. The technique used most commonly to forecast AQI is machine learning. Machine learning techniques use large data sets and algorithms based on machine learning to train the model. Neural network methodology based on deep learning is another approach to AQI forecasting. A basic neural network can be used to provide accurate AQI predictions. The model can then be further tuned by varying the testing settings and input parameters [7].

The literature reviewed for this work highlights a few flaws in the methods currently used to forecast the air quality index, such as issues with dataset collecting. The low forecast accuracy of the air quality index prediction models that are currently available in Sri Lanka is caused by incomplete and inaccurate data. Data preparation is another element that makes the accuracy decline. Considering these elements, it is evident that the majority of Sri Lanka's present systems are unable to generate accurate projections. In light of the aforementioned in relation to other countries, they have succeeded in overcoming these obstacles and achieving exceptional precision.

Table 1.1 Comparison with existing applications

Product Reference	Research A	Research B	Research C	Proposed System
Mobile-based approach	✓	✓	×	✓
Geographical Information System	×	✓	×	✓
Real time notification regarding the high pollution Air Quality Index	~	✓	×	✓
Anonymous Data gathering based on the location.	×	×	×	✓
Make awareness to the affected people by connecting them via mobile platform	×	×	×	~

1.3 Research Problem

A major gap in the field of air quality prediction in Sri Lanka is the focus of the research challenge addressed in this work. This gap is particularly related to shortcomings in data preprocessing techniques and the general accuracy of the current Air Quality Index (AQI) prediction systems. The problems include shortcomings in the way that data preprocessing is currently done, such as how to deal with incomplete and erroneous data, which has a big effect on how reliable AQI forecasts are. Furthermore, difficulties in obtaining data pose a barrier to the high accuracy that the current prediction systems are capable of. Solving this research issue is critical to improving the accuracy of AQI forecasts, which is necessary for efficient air quality control. Additionally, in order to ensure accuracy on par with worldwide benchmarks, it is imperative that the current systems be improved in order to bring Sri Lanka's AQI prediction algorithms into compliance with international standards. It becomes necessary to close this gap in order to create reliable air quality forecasts that benefit human health as well as the environment. The main objective of the research is to create more sophisticated data preprocessing techniques and enhance current AQI prediction algorithms in order to provide a more sustainable and healthful living environment in Sri Lanka.

Data Preprocessing Challenges:

The initial facet of the issue concerns techniques for preparing data. Like many other developing countries, Sri Lanka struggles with problems like missing datasets, variable data quality, and regional differences in data gathering techniques. Predictive models are less effective as a result of these difficulties in precisely capturing the actual state of air quality. The predictive accuracy of AQI models is weakened in the absence of strong data preprocessing methods, which reduces their usefulness for decision-making and reducing the harmful impacts of air pollution on public health.

Accuracy of AQI Prediction Systems:

The overall accuracy of AQI prediction systems is the focus of the second component of the research challenge. Even with advances in predictive analytics and machine learning, Sri Lanka's current methods frequently fail to anticipate AQI levels with sufficient accuracy. This may be due to a number of things, such as inadequate model calibration, inadequate attention to local environmental conditions, and restrictions on the temporal and spatial resolution of the datasets that are currently available. Decision-makers thus have difficulties in organizing and carrying out focused initiatives to address air quality issues, which could pose health hazards and deteriorate the environment.

Implications and Significance:

The research topic has repercussions that go beyond academic investigation and include wider societal and environmental consequences. The public's health is seriously threatened by poor air quality, which is a major cause of respiratory ailments, cardiovascular disorders, and other negative health effects. It also makes environmental deterioration worse, which affects biodiversity, ecosystems, and the general balance of the ecosystem. Stakeholders can obtain practical insights into the dynamics of air quality and use this knowledge to inform policy formation, urban planning, and pollution control strategies by filling in the identified research need. Furthermore, raising the AQI forecast systems' accuracy is in line with international initiatives to tackle climate change and advance sustainable development, establishing Sri Lanka as a proactive player in the global environmental agenda.

2. OBJECTIVES

2.1 Main Objectives

This component's main goal is to provide a reliable model for estimating pollution levels. To effectively anticipate the concentration of contaminants in the air, this requires examining historical data, present environmental conditions, and other pertinent elements. By means of a thorough investigation, the project team determines which variables have the most influence. Then, feature engineering methods are implemented to improve the predictive power of the model. To extract more useful information and raise the predictive model's overall accuracy, this entails modifying and adjusting the selected features. Dynamic predictions are a novel element of the predictive model that allows for real-time modifications based on shifting environmental parameters. The model is now more responsive to quickly changing situations thanks to this improvement. Additionally, the integration of the pollution prediction model with the route generating component guarantees that the model actively participates in the creation of the best and least-polluted paths.

2.2 Specific Objectives

- Alerting the affected users who have registered once a high pollution level is discovered.
 - o Internal Server will be implemented to capture the event whenever an air quality index found high using the application.
- Data visualization and comparison of the degree of air pollution in Colombo areas
 - o Authorized users can view air pollution levels for past, present and future

3. METHODOLOGY

3.1 System Architecture Diagram

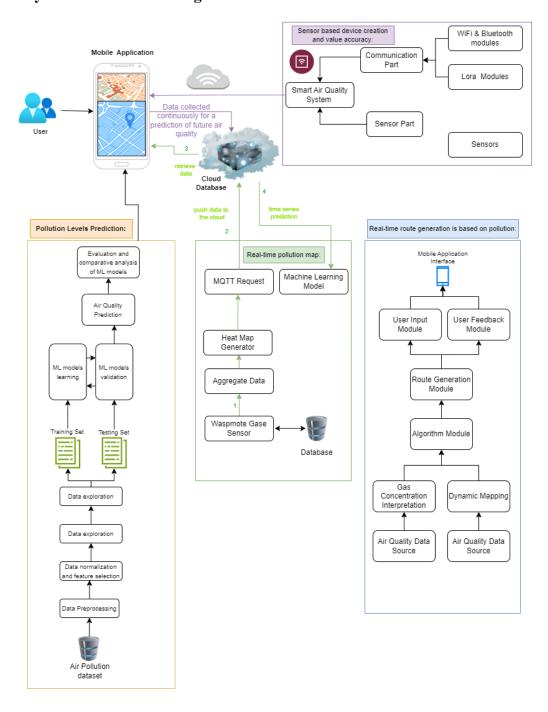
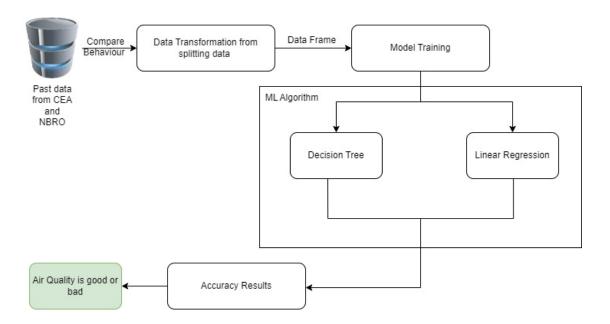


Figure 3. 1 System Architecture Diagram



The recommended approach for AQI prediction is a methodical procedure. The process entails obtaining the dataset from the Central Environmental Authority, pre-processing it, applying the appropriate machine learning algorithms, analyzing it to identify correlations between air pollution factors, and ultimately selecting the optimal machine learning strategy and assessing the prediction outcomes.

3.2 Project requirements

3.2.1 Functional Requirements and Non-Functional Requirements

Table 3. 1 Functional and Non-Functional Requirements

Functional Requirements	Non-Functional
	Requirements
1. Collecting polluted air data from the emission standard publication websites and Preprocessing, Cleaning, Transforming.	1.System should respond Realtime.
2. Feature Selection and Analysis.	2. The application should be reliable.
3. Model Development Implement machine learning algorithms for pollution level prediction. Explore and select the most suitable predictive model based on performance metrics.	3. Higher accuracy of results and more efficient results
4. Dynamic Predictions for Real-time Adjustments.	4.Interfaces should be user-friendly.
5. Real-time Pollution Data and Comparison and Visualization	5. Should properly work for android and IOS devices

3.2.2 Software Requirements

Software requirements are React Native, Flutter, Node Server, Firebase, Python, Jupyter Notebook, VS code, TensorFlow and IntelliJ Idea.

3.2.3 Personnel Requirements

The following people are needed in order to improve the research's integrity, quality, and knowledge.

Resources and Dataset Air pollution level in Colombo

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

3.3 Software Solution

Agile Scrum approach is the software development life cycle that is being suggested. Based on the agile principles of the Agile Manifesto, Scrum is an iterative methodology for software development [8]. Another definition of scrum is a lightweight development approach that offers complete transparency and quick adaptation [8]. Because of the aforementioned tendencies, agile allows for rapid component modifications during the implementation phase in response to changing needs. Agile scrum technique will be used to facilitate the changes in an effective manner. As a result, the suggested solution will be put into practice in accordance with the framework by facilitating quick adaptation and ongoing adjustments.



Figure 3. 2 Agile Methodology [9]

3.3.1 Requirement Gathering and Analysis

Gathering requirements from National Building Research Organization (NBRO) and Central Environmental Authority (CEA).

By requesting authority people who worked at above mention workplace collecting dataset of minimum of 5 years data of pollution level in Colombo areas. Discuss the problem in the datasets.

3.3.2 Feasibility Study

Feasibility Analysis is the process of determination of whether or not a project is worth doing.

Technical Feasibility

Members should have extensive knowledge of node.js, JavaScript, React Native, using external libraries, retrieving data from external APIs, event-driven architectures, and CPU optimization techniques in order to effectively finish the project. Project contributors also need to have specialized knowledge in software architectures and frameworks, mobile app development, and web app development.

Economic Feasibility

The limitations of the project required the employment of more dependable and reasonably priced resources in order to guarantee cost-effectiveness and dependability. As a result, budgetary constraints affected the choice of materials and essential elements. It is anticipated that the proposed sub-component will function flawlessly, with no problems or faults, stressing dependability, high performance, and affordability. The project strives for low costs related to component requirements and resources.

Schedule Feasibility

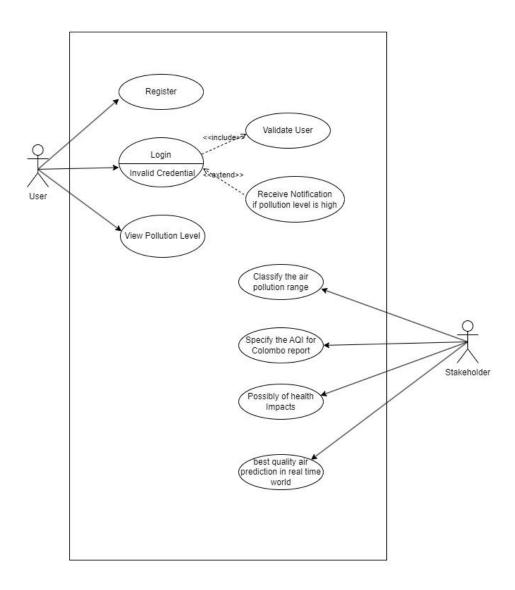
The proposed subcomponent needs to be completed in the allotted time frame, with higher precision in the results, and strictly adhering to the timeline that has been set forth. Last but not least, the final product presentation ought to happen on the scheduled deadline.

Operation Feasibility

A designated member is in charge of each stage of the software life cycle, with an emphasis on the requirement analysis phase in particular. Ensuring that the finished product meets the needs specified by the consumers is the ultimate goal, with a focus on user-centric design.

3.3.3 Design

3.3.3.1 Use Case Diagram



The Figure 3.3 overall use case diagram indicates the external and internal actor functional and technical overview with the node server system.

3.3.3.2 Sequence Diagram

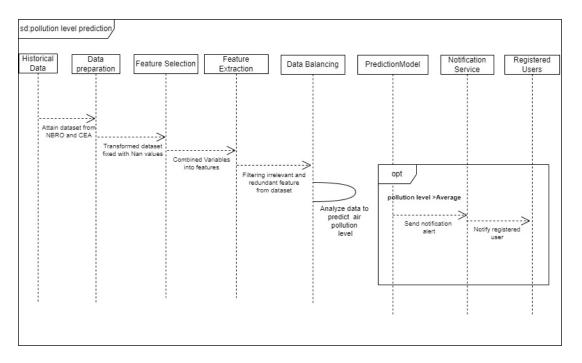


Figure 3. 4 Sequence Diagram

The figure 3.4 sequence diagram depicts the technical interactions with the system.

3.3.3.4 Dataset

The Central Environment Authority (CEA) and the National Building Research Organization (NBRO) have historical datasets with data on the hourly concentration levels of air pollution components in Colombo. The dataset includes average concentrations of air and meteorological components, including humidity, CO, SO2, NO2, PM10, and PM2.5, for the period of January 2020 to December 2023. Preprocessing techniques are being applied to the generated dataset in order to increase prediction accuracy and ensure the validity of the results. The distribution and type of the dataset, as well as the relationships between the variables related to air pollution, are ascertained using distribution charts and correlation matrices.

3.3.3.5 Work Breakdown Structure

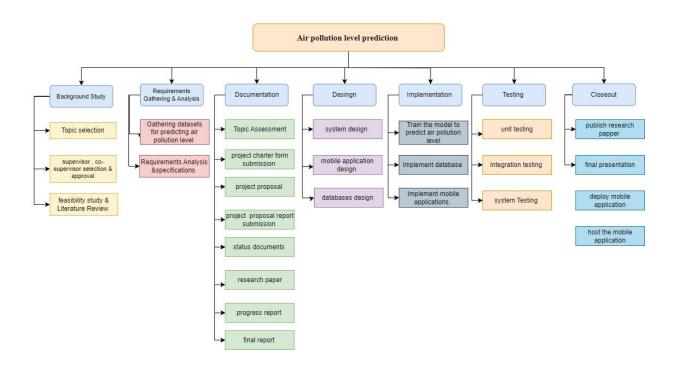


Figure 3. 5 Work Breakdown Structure

3.3.3.6 High Fidelity Diagram



Figure 3. 6 High fidelity diagrams for information sharing to crowd.

3.3.4 Implementation

The proposed approach consists of a sequence of phases for predicting the AQI. The sequence of phases includes collecting the dataset from Central Environmental Authority, pre-processing the collected dataset, analyzing the collected dataset to identifying the correlations among air pollution factors, applying appropriate ml algorithms, & ultimately selecting the most suitable machine learning approach

& analyzing the prediction results. as given below,

- Collecting polluted air data from the emission standard publication websites and Preprocessing, Cleaning, Transforming.
- Feature Selection and Analysis.
- Model Development Implement machine learning algorithms for pollution level prediction. Explore and select the most suitable predictive model based on performance metrics.
- Dynamic Predictions for Real-time Adjustments.
- Real-time Pollution Data and Comparison and Visualization.

We're bringing all these features together so that everyone involved can easily see and use them through a mobile app.

Mobile Application

The envisioned mobile application for predicting air pollution levels offers support in the English language. The application will be available in two versions: a commodity version and a premium version. These versions provide users with the flexibility to choose based on their preferences and needs, offering enhanced features in the premium version. To ensure users stay informed, the application is designed to send push notifications, delivering timely updates on air quality even when the app is not actively in use. This feature contributes to a seamless and user-friendly experience, keeping individuals updated on crucial information related to air pollution levels.

Sever Back-End

The server back-end is the most important part of the research, which is responsible for data storing and processing. MongoDB is the database of the system, it is a cross-platform, open source and document-oriented database.

3.3.5 Testing (Track and Monitor)

Every function in the application is tested for errors during this step. We have a plan to test it in different ways, like checking if users like it, making sure it works with other parts, and checking each part on its own. This step is not just about fixing mistakes; it's also about making sure the app is really good and works well.

4. GANTT CHART

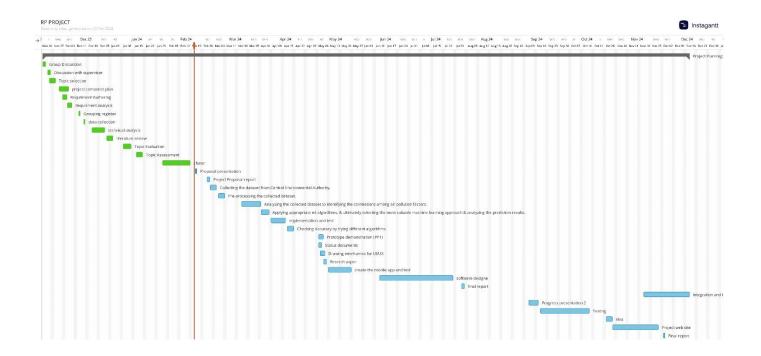


Figure 3. 7 Gantt Chart

5. DESCRIPTION OF PERSONAL AND FACILITIES

Facilitators:

Ms. Chathurangika Kahandawarachi - Sri Lanka Institute of Information Technology (SLIIT)

Ms. Pipuni Wijesera - Sri Lanka Institute of Information Technology (SLIIT)

Facilities:

National Building Resource Management (NBRO)

Central Environment Authority (CEA)

6. BUDGET AND BUDGET JUSTIFICATION

The estimated budget contains subscription costs, deployment costs, database costs and hosting costs.

Table 5. 1 Budget of the crowdsourcing component

Feature	Price
Travelling cost for collecting data	Rs. 4850.00
Deployment Cost	Rs. 5080.00 / month
MongoDB	Free
Mobile App -Hosting on Play Store	Rs. 4996.00
Mobile App -Hosting on App Store	Rs.18000.00 /annual
Total	Rs. 32926.00

6. COMERCIALIZATION

6.1 Target Audience and Market Space

Target Audience

- o Stakeholders
- o Public
- o School Childrens

Market Space

- O No need of prior knowledge in technology.
- o No age limitation for users.
- o No need of prior knowledge regarding Air Pollution.

REFERENCES

- [1] Mahanta, S., Ramakrishnudu, T., Jha, R. R., & Tailor, N. (2019, October). Urban air quality prediction using regression analysis. In *TENCON 2019-2019 IEEE Region 10 Conference (TENCON)* (pp. 1118-1123). IEEE.
- [2] Castelli, M., Clemente, F. M., Popovič, A., Silva, S., & Vanneschi, L. (2020). A machine learning approach to predict air quality in California. *Complexity*, 2020.
- [3] Zhong, S., Yu, Z., & Zhu, W. (2019). Study of the effects of air pollutants on human health based on Baidu indices of disease symptoms and air quality monitoring data in Beijing, China. *International journal of environmental research and public health*, 16(6), 1014.
- [4] Nandasena, Y. L. S., Wickremasinghe, A. R., & Sathiakumar, N. (2010). Air pollution and health in Sri Lanka: a review of epidemiologic studies. *BMC Public Health*, 10, 1-14.
- [5] Xayasouk, T., & Lee, H. (2018). Air pollution prediction system using deep learning. WIT Trans. Ecol. Environ, 230, 71-79.
- [6] Iskandaryan, D., Ramos, F., & Trilles, S. (2020). Air quality prediction in smart cities using machine learning technologies based on sensor data: a review. *Applied Sciences*, 10(7), 2401.
- [7] Samad, A., Garuda, S., Vogt, U., & Yang, B. (2023). Air pollution prediction using machine learning techniques—an approach to replace existing monitoring stations with virtual monitoring stations. *Atmospheric Environment*, *310*, 119987.
- [8] Anum, A., Rehman, M., & Anjum, M. (2017). Framework For Applicability Of Agile Scrum Methodology: A Perspective Of Software Industry. *International journal of advanced computer science and applications*, 8(9).
- [9] D. Iddo, "Agile Development," Medium, 04-Oct-2019. [Online]. Available: https://medium.com/moodah-pos/agile-development-95cad3573abf. [Accessed: 22-Mar-2021].

APPENDICES

· Plagiarism Report

