Quality of ambient air prediction in Bangladesh - A Time Series Analysis and Machine Learning Approach

By

Habiba Chowdhury Akhi
211-15-14627

FINAL YEAR DESIGN PROJECT REPORT

This Report Presented in Partial Fulfillment of the Requirements for the **Degree of Bachelor of Science in Computer Science and Engineering**

Supervised by

Ms. Most. Hasna Hena
Assistant Professor
Department of Computer Science and Engineering
Daffodil International University

Co-Supervised by

Ms. Chayti Saha
Lecturer
Department of Computer Science and Engineering
Daffodil International University

DAFFODIL INTERNATIONAL UNIVERSITY
Dhaka, Bangladesh

January 12, 2025

APPROVAL

This Project titled **"Quality of ambient air prediction in Bangladesh- A Time Series Analysis and Machine Learning Approach"**, submitted by Habiba Chowdhury Akhi, ID No: **211-15-14627** to the Department of Computer Science and Engineering, Daffodil International University has been accepted as satisfactory for the partial fulfillment of the requirements for the degree of B.Sc. in Computer Science and Engineering and approved as to its style and contents. The presentation has been held on 12 January, 2025.

BOARD OF EXAMINERS

Dr. Fizar Ahmed (FZA) **Chairman**
Associate Professor
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

Mr. Abdus Sattar (AS) **Internal Examiner**
Assistant Professor
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

Ms. Zahura Zaman (ZZ) **Internal Examiner**
Lecturer
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

Dr. Ahmed Wasif Reza (DWR) **External Examiner**
Professor
Department of Computer Science and Engineering
East West University

DECLARATION

I hereby declare that this project has been done by me under the supervision of **Ms. Most. Hasna Hena, Assistant Professor**, Department of Computer Science and Engineering, Daffodil International University. I also declare that neither this project nor any part of this project has been submitted elsewhere for the award of any degree or diploma.

Supervised by:

12.01.25

Ms. Most. Hasna Hena
Assistant Professor
Department of Computer Science and Engineering
Daffodil International University

Co-Supervised by:

Chayti Saha
12.01.25

Ms. Chayti Saha
Lecturer
Department of Computer Science and Engineering
Daffodil International University

Submitted by:

Habiba
12.01.2025

Habiba Chowdhury Akhi
Student ID: 211-15-14627
Department of Computer Science and Engineering
Daffodil International University

ACKNOWLEDGEMENTS

This work would not have been possible without the support and contributions of many individuals over the past two semesters. I am deeply grateful to everyone who has assisted me in one way or another.

First, I express my heartfelt thanks and gratefulness to the almighty for His divine blessing making it possible for me to complete the **Final Year Design Project(FYDP)** successfully.

I am grateful and wish my profound indebtedness to **Ms. Most. Hasna Hena**, **Assistant Professor**, Department of Computer Science and Engineering, Daffodil International University, Dhaka, Bangladesh. Deep knowledge and keen interest of my supervisor in the field of **Machine Learning and Data Analysis** to carry out this project. His endless patience, scholarly guidance, continual encouragement, constant and energetic supervision, constructive criticism, valuable advice, reading many inferior drafts, and correcting them at all stages have made it possible to complete this project.

I would like to express my heartiest gratitude to the **Dr. Sheak Rashed Haider Noori, Professor and Head,** Department of CSE Daffodil International University, Dhaka, for his kind help in finishing my project and also to other faculty members and the staff of the Department of CSE, Daffodil International University.

I would like to thank my entire course-mates at Daffodil International University, who took part in this discussion while completing the coursework.

Finally, I must acknowledge with due respect the constant support and patience of my parents.

ABSTRACT

Air quality is a key environmental concern, especially in rapidly urbanizing countries such as Bangladesh, where air pollution poses substantial dangers to public health and the economy. This study investigates air quality index (AQI) patterns and trends across 13 main cities in Bangladesh, strives to understand the elements that determine air quality, and attempts to predict future AQI levels using three predictive models: Linear Regression, ARIMA, and LSTM. The study focuses on the ability to estimate AQI values in diverse urban environments by analyzing these models based on many performance measures including mean squared error (MSE), root mean squared error (RMSE), mean absolute error (MAE) and R-squared (R2). This collection, containing historical AQI data for cities such as Dhaka, Chittagong, Rajshahi and Sylhet in Bangladesh, demonstrates considerable differences in air quality between locations. The results demonstrate that LSTM outperforms other models with consistent forecasts and reasonably low error metrics across most cities. LSTM achieved the lowest MSEs and RMSEs in numerous urban regions, confirming its applicability for time-series forecasting of AQI. On the other hand, linear regression worked effectively in cities where AQI patterns were simpler and more linear, whereas LSTM, although a more advanced deep learning model, displayed some issues in managing non-linearity and seasonality in data. The relatively high MAE and low R2 of the model imply that more refining is required for its effective usage in AQI forecasting. The study also highlights seasonal and temporal tendencies, such as particular cities such as Dhaka and Narayanganj having higher pollution levels throughout certain months, underscoring the necessity for seasonal air quality management. This study shows the importance of predictive modelling in air quality monitoring and policy making. The results imply that specific efforts are needed for high-polluting cities, and future studies should consider hybrid models or more advanced machine learning approaches to boost the accuracy and usefulness of AQI predictions in Bangladesh all of this I display in my Website.

Table of Contents

Approval i

Declaration ii

Acknowledgements iii

Abstract iv

List of Figures vii

List of Tables viii

1 Introduction 1

1.1 Introduction 1
1.2 Motivation 2
1.3 Objectives 3
1.4 Methodology 3
1.5 Project Outcome 4
1.6 Organization of the Report 4

2 Background 6

2.1 Introduction 6
2.2 Literature Review 7
2.2.1 Similar Applications 7
2.2.2 Related Research 7
2.3 Gap Analysis 9
2.4 Summary 9

3 Research Methodology 11

3.1 Methodology 11
3.1.1 Overview 11
3.1.2 Proposed Methodology 11
3.1.3 Functional and Nonfunctional Requirements 13
3.1.4 Context Diagram 13

3.1.5 Data Flow Diagram Level 1 14
3.1.6 UI Design 15
3.2 Detailed Methodology and Design 23
3.3 Project Plan 30
3.4 Task Allocation 31
3.5 Summary 32

4 Implementation and Results **33**
4.1 Environment Setup 33
4.2 Comparative Analysis 33
4.3 Results and Discussion 34
4.4 Summary 40

5 Engineering Standards and Design Challenges **41**
5.1 Compliance with the Standards 41
5.1.1 Software Standards 41
5.1.2 Hardware Standards 41
5.1.3 Communication Standards 42
5.2 Impact on Society, Environment and Sustainability 43
5.2.1 Impact on Life 43
5.2.2 Impact on Society & Environment 43
5.2.3 Ethical Aspects 44
5.2.4 Sustainability Plan 44
5.3 Project Management and Financial Analysis 45
5.4 Complex Engineering Problem 47
5.4.1 Complex Problem Solving 47
5.4.2 Engineering Activities 48
5.5 Summary 49

6 Conclusion **50**
6.1 Summary 50
6.2 Limitation 50
6.3 Future Work 51

References **53**

List of Figures

3.1 Proposed Methodology. 12
3.2 Context Diagram of Website. 14
3.3 Data Flow Diagram Level 1 of Website. 15
3.4 UI Design of Website Home Page. 16
3.5 UI Design of PM 2.5 Page 17
3.6 UI Design of PM 10 Page 18
3.7 UI Design of CO Page 19
3.8 UI Design of SO2 Page 20
3.9 UI Design of NO2 Page 21
3.10 UI Design of O3 Page 22
3.11 AQI Desity Curves 13 Main City's in Bangladesh. 27
3.12 Outliers of Each City. 29
3.13 Dencity Distribution of AQI in 13 Cities. 30
3.14 Project Plan Time Frame 31

4.1 Comparisun of LR ARIMA and LSTM Model by MSE 35
4.2 Comparisun of LR ARIMA and LSTM Model by RMSE 36
4.3 Comparisun of LR ARIMA and LSTM Model by MAE 37
4.4 Comparisun of LR ARIMA and LSTM Model by R2 37
4.5 Lowest AQI of Different City In Bangladesh. 39
4.6 Avarage AQI of Different City In Bangladesh. 39
4.7 Highest AQI of Different City In Bangladesh. 40

5.1 Project Management Time Frame 46

List of Tables

2.1 Gap Analysis with Previous Work. 9

4.1 Comparative Analysis My Work with Other Study 34

5.1 Financial Analysis 46
5.2 Mapping with complex problem solving. 47
5.3 Mapping with knowledge Profile. 48
5.4 Mapping with complex engineering activities. 48

Chapter 1

Introduction

1.1 Introduction

Pollution of air is arguably one of, the most critical environmental problems in the world today with far-reaching effects on human lives, growth, and development of the economy and the entire natural systems. A commendable urban growth as well as an increase in the number of industries has further created air polluting issues in Bangladesh mostly in the large cities. AQI is a detailed index of air pollution that comprises density of primary pollutants such as Particulate Matter—PM 2.5 and PM 10, Sulphur dioxide-SO2, Nitrogen dioxide– NO 2, Carbon monoxide–CO Ozone- O 3. Study of AQI's trends and cycles is essential in case formulation and intervention to improve the health of the people. In the current research analysis, AQI data on 13 cities in Bangladesh were explored to analyze distribution, seasonal analysis, and episodes of air pollution. It had a total number of rows equal to 1173 and total number of features 27, which gives a strong background to look at the changes in air quality. This study is an attempt to understand the changes in AQI and its variation between cities to find out the places where is still problem of air quality and regular occurrence of dangerous level of Air Quality Index. Short-term and long-term studies are conducted because some events reoccur, for example, winter fog or industrial emissions, which contribute to air pollution. One of the significant aims of this study is to determine that part of the nation that necessitates policy action. Pollution and air quality problems in certain cities or areas require enhanced legal provisions or more facilities or mass sensitization programs. Besides, the study used linear regression, ARIMA, and LSTM to forecast the future AQI levels in order to assess the air quality. Optimal estimates can provide policymakers and other stakeholders with relevant knowledge to enable them to take preventable decisions. The main research topics of this study include: Which bangladeshi cities have the worst [Air Quality Index] on an index average? On how many occasions can cities be categorized as hazardous pollution like "extremely unhealthy"? Is there therefore any visible seasonality or temporal variation in the AQI data? And can AQI be accurately estimated by reference to the previous status? The answers to these questions are the goals of our project to add to knowledge on air pollution issues in Bangladesh and facilitate sound approaches to the alleviation of pollution and its impacts. In an attempt to address the existing literature gap in air quality research in Bangladesh, this work aims at identifying and exploring the quantitative measures required for sustainable urban and environmental improvement in the context of Bangladesh.

1.2 Motivation

Pollution of air is among the most significant global environmental and health problems, and the same is true about Bangladesh. As the industries, cities and populaces have grown, the air quality bad and has affected millions of lives in Bangladesh. Air pollution endangers not only people's health – it causes respiratory and cardiovascular diseases, premature death, health care costs, and decreased work productivity but is also adds to costs for the state as a result of worsening economic development. This dangerous situation calls for immediate intervention of the sources of air pollution and formulation of efficient measures. The background to this research is based on the increased need of responding to these concerns using empirical evidence on air quality trends in various cities of Bangladesh. However, there is a lack of comprehensive study on AQI variability and offering forecasting recommendations proportionate to Bangladesh's pollution, environment, and similar socioeconomic scenarios despite being among the most polluted countries on earth. Awareness of AQI distribution, its trends, and large pollution events is potentially useful in properly targeting policy interventions and investments. Moreover, the above-discussed emissions are particularly dangerous for sensitive groups of a population like children, elderly, and those who suffer from chronic diseases. The negative health impact of long-term exposure to polluted air further call for recognition of towns/regions with poor air quality which transgress the set standards. Harnessing these places can help catalyse targeted action in the form of better laws on industrial emissions, urban design, and information campaigns that will alter the behaviour of the inhabitants. From a scientific view, it is essential to predict future AQI level to be able to take relevant decisions on time. This insight makes it possible for authorities to predict future pollution levels, take the right measures, and study implementation outcomes in case of a policy shift. The use of sophisticated model identification and selection methodologies such as ARIMA and LSTM in this work not only extends methodological developments but also contributes greatly towards translating research findings into practice. Last of all, it is oriented by a goal to improve life satisfaction of millions of inhabitants in Bangladesh by addressing the roots of air pollution. Being a practical research, I aim as a result in contributing to the paradigm of sustainable development and construct a better, more habitable world for the future generation.

1.3 Objectives

The objective of this research is to establish the exploratory pattern of AQI in 13 cities of Bangladesh and report logically presented findings to tackle the challenges of air pollution. This paper aims at analyzing the distribution of AQI levels with the intention of determining which cities experience poor air quality most of the time longitudinally with a view of determining those in the 'very unhealthy' or worse hazardous pollution bracket. Knowledge of these patterns will help to prioritize the municipalities and regions that need fast changes from the government and active environmental management. One of the other important aims is for the first time recognizing the fluctuations in air quality depending on the seasons and time periods: cyclic patterns associated with weather conditions, or periods of industrial activity, etc. These tendencies may assist in shedding more light on what perhaps causes pollution increases over time and specific time that can be intervened to prevent occurrence of this kind of pollution.This research aims to develop accurate and reliable air quality prediction models using linear regression, ARIMA, and long short term memory network on the results of air quality. These projections are important to allow decision-makers, city specialists, and environmental nongovernmental organizations to make decisions, develop strategies, and evaluate the efficiency of their actions.

1.4 Methodology

This research applies a systematic approach in analyzing AQI data to achieve the laid down objectives. The overview of the investigation starts with data preprocessing in an attempt to deal with the missing components, duplicate and inconsistent data in the database. Descriptive statistical methodologies are used to study the frequency distribution of AQI in 13 cities of Bangladesh which will help to assess the mean level of air quality, and the incidence of days with severe air pollution risks. For this reason, time-series analysis is conducted, using decomposition methods in order to identify trends, seasonality and other residual factors. AQI high events, therefore, are examined to explain their frequency and effects on people and the environment. Probability analysis and stats as well as machine learning play important roles in the formulation of the predictive modeling plan. Data analysis and trends of AQI and other contributive factors call for linear regression analysis while AQI future trends are estimated using ARIMA models. LSTM networks were used to consider complex and nonlinear relationship and long-term structures in AQI. To ensure accuracy of the developed models other parameters including RMSE and mean of absolute errors are used. (MAE). For enhancement of graphical understanding and visual appeal, current research work employed various graphical interfaces comprising of line plots, bar charts and heat maps for depicting AQI with respect to its trend, dispersion and seasonal variations. The AQI not only indicates how much pollution a city should expect to have, but by looking at cities that consistently have high AQI values or numerous extreme pollution events, policy action areas are determined which are useful suggestions for amelioration. This approach hence enables the understanding of various AQI trends and the formulation of better AQI policies by the Bangladesh government and WHO.

1.5 Project Outcome

The following objectives are achieved with intent and a systematic method when assessing Air Quality Index data in this study. The first step of investigation is data cleansing which ensure the dataset is complete, consistent where issues such as missing values, duplications and inconsistencies are resolved. In this study, the descriptive statistical methods are employed to compute the average of AQI and the rate of hazardous pollution in 13 cities of Bangladesh. For the identification of seasonal and temporal patterns time series analysis is conducted, using decomposition to break down the data into trends, seasonal factors and residuals. The AQI index are studied at their highest levels to identify the occurrence sequence and the effects they have on health and the environment. Use of statistics and machine learning techniques constitutes the strategy for data modeling for prediction also make a website so that any one can see the findings. Simple linear regression is computed to assess the correlation between AQI and contributing factors whereas multiple linear regression is used in time series modeling to predict AQI for subsequent time by using ARIMA models. As for modelling, long short-term memory networks are used for identifying intricate, nonlinear relationships and long temporal dependencies of the AQI data. These models are validated for accuracy by using additional parameters like Root mean square error, Mean absolute error. For enhancement of the visualization and eager to understand easily the research used many graphical representations such as line plot, bar chart and heat plot for explaining trends, distribution and fluctuation of AQI. Thus, cities with high average AQI or with multiple cases of extreme pollution sustained over time are characterized; thereby, the areas for policy actions are defined and useful suggestions for environmental interventions are provided. It thus guarantees full comprehension of the AQI patterns, and fosters the development of sound policies that would improve air quality in Bangladesh.

1.6 Organization of the Report

The report is intended to provide a clear and thorough narrative of the study method, findings, and consequences. Each chapter serves a specific purpose, ensuring cohesion and depth in presenting the research.

Chapter 1 includes an overview of the report. It addresses the general overview of the topic of concern before proceeding to highlight on the rationale of the project. The purpose of this section is to provide an overview of the research goals of this study, the approach used in the study, the anticipated project's outcomes, and the layout of the report as a useful aid to readers. Chapter 2 forms the background and context of the research. It consists of an overview of the background and development of the particular research, secondary to the literature review examining previously conducted research in the subject area. This study need is further established given that a gap analysis is made in order to point out where previous studies lack. The last section of the chapter presents the conclusion of the study. Chapter 3 presents details regarding the approach used in the project. This writing commences with a brief description of the research methodology and the general plan of the system's design; it also provides a comprehensive description of the steps involved in the actualization of the proposed research methodology and system design. The chapter also contains the project plan, tasks to be assigned and the chapter summary. Chapter 4 is devoted to the description of the key practical aspects of the work. They include setting up of the environment, testing as well as evaluating the system. A performance analysis is presented, as well as a comparative analysis. The findings are further explored while placing critical focus on their status and relation to objectives. The conclusion of the chapter is a review of the findings of the study. Chapter 5 focuses on the engineering regulations of the project such as software and hardware engineering regulations. They articulate the social relation of the project to society and the environment, ethical question, and sustainability measures. Third, it measures the efficiency of project management and assesses the financial aspect of the project and offers an approach towards solving engineering issues for conduction the job. The last section of the chapter is a conclusion. In Chapter 6 the author articulates conclusion that in essence reflects all the findings and accomplishments of the report. The report is closed out with a discussion on the limitations and a listing of potential future work that can be done on the study.

Chapter 2

Background

2.1 Introduction

Pollution has emerged as a major global problem for human health, environment, and the economy. AQI is an aggregated index adopted by many countries to describe ambient air quality using concentration data of hazardous pollutants including particulate matter (PM 2.5 and PM10), sulphur dioxide (SO2), nitrogen dioxide (NO2), carbon monoxide (CO) and ozone (O3). Experience has shown that different economic activities such as fast industrialization, urban development and increased population density in developing societies like Bangladesh has resulted to significant deterioration in air quality especially in cities. Based on current research, Bangladesh often tops the league table of countries with very poor ambient air quality, highly dangerous AQI values, inclusive of populated areas like Dhaka. Air pollution contributes negatively in many ways in Bangladesh. Physical health effects involve higher incidences of respiratory diseases, cardiovascular diseases, and premature death while the other effects have a toll on the health of children, the elderly and other at risk groups. The economic impacts of poor air quality include greater health care costs, lower work efficiency and add a significant burden to the country's capacity to afford and supply resources. Nevertheless, there is a lack of extensive, empirical research on the AQI status and its changes in Bangladeshi cities. Previous literature often assesses individual facets of air quality separately based on their type of pollutants, which has major gaps considering the spatial-temporal distribution of AQI in China. It is crucial to fill these gaps in order to persist in developing the right policies and what should be done to minimise the effects of air pollution. This study attempt to fill these gaps by assessing AQI data of 13 cities of Bangladesh. Policies can be informed by assessing distribution patterns, seasonal fluctuation and on the significant episodes of pollution, according to the study. Further, using analytical prediction models to determine the future AQI values, the present investigation offers potent tools for effective environmental planning. In view of these, the current research aims to shed more light on the dynamics of air quality with a view to supporting larger propagate abroad health and the general environment in Bangladesh.

2.2 Literature Review

2.2.1 Similar Applications

Several kind of similar applications and projects are now in operation to monitor and manage air quality in Bangladesh with potentially useful data and awareness applications for the public and policymakers. Some of the notable examples are IQAir; it provides the current air quality index and Particulate Matter with size of 2.5 micrometers for cities of Bangladesh [1]. The U.S. Embassy in Bangladesh also provides opportunity for the general public to know more about the air quality in Bangladesh by the help of the monitoring stations set which measures PM2.5 [2]. Also, the Bangladesh Air Quality Management Project launched by the World Bank focuses on reducing car emissions and implementing the air quality standard on large cities [3]. The NSF-IRES Project and researchers from Bangladesh have partnered to promote higher monitoring with aid of low-cost sensors and satellite data [4]. The National Ambient Air Quality Standard of Pakistan 2024-2030 also deals with the guideline for the real-time air quality index and its integration with the department of environment systems is listed below MoEF Portal Gov BD [5]. In addition, the 'Ambient Air Quality in Bangladesh' report outlines current trends on air quality as well as information about the observed meteorological conditions [6]. The Air Quality Monitoring of Bangladesh is composed of acquired atmospheric maps and of AQI forecasts from the SCITEPRESS platform, whereas Real-time Air Quality Monitoring System IoT based concentrates on pollution risks in cities from the IEEE Xploreooks platform [7, 8]. The presented applications show how both technology and partnerships matter when it comes to addressing the problem of air pollution in Bangladesh and promoting sustainability.

2.2.2 Related Research

Condition of the air in Bangladesh is worsened especially in Dhaka city due to the unhealthy index of the air quality index of the country that regularly falls among the lowest in the world. A number of works have been devoted to the analysis of different factors and issues related with air pollution, their effects on health, as well as ways to minimize them. The cost of health in Dhaka due to air pollution is very high and there is high level of association between the pollutant concentrations and respiratory and cardiovascular ailment. Industries, automobiles and constructions dust are considered as main sources that affect the air quality adversely [9]. The reviews of air quality of Dhaka highlighted that socio anthropogenic factors play a significant role in air pollution due to urbanization and industrialization. These reviews motivate models focused on refining the clean technologies as well as enforcing more rigid laws to enhance air quality [10]. High technology using sensors for air quality monitoring has been recommended as affordable methods for detecting spatial and temporal differences in pollution. These systems are especially appropriate in places such as Dhaka where monitoring frameworks are not well established,

[11]. According to the cross-sectional findings of activities conducted on the temporal changes in the concentrations of pollutants during a 17 year period (2003 to 2019) there is a relatively increase in the concentrations of the pollutants which includes PM2. A relationship between pollution and climatic variables including surface temperature has also been confirmed. High exposure to pollution has been found to have made young people in Dhaka sensitive to respiratory difficulties [12]. Air pollution is reported to cause about three million deaths in low and middle income countries every year, according to WHO. However a number of monitoring efforts continue to face problems in Bangladesh such as high cost implication as well as substandard physical facilities. Many research works have been carried out to use satellite-aided systems and IoT as additional resources in place of the conventional ground-oriented observation systems [13]. The work uses Deep Convolutional Neural Networks (DCNNs) to predict AQI from atmospheric imaging and pollutant concentration levels. These models have shown fairly good improvement over baseline methods of analysis thus showing that data driven models can be of much utility in research related to air quality [14]. Fluctuations of the PM2.5, PM10, SO2, NOx, and CO concentrations in Dhaka have been established with this season being the most polluted. The non-monsoon seasons have regularly crossed the national and international air pollutant standards. The increase in industrial and transportation emissions are demonstrated in trends of pollutants like SO2, CO, etc. [15]. The health impacts of air pollution in Dhaka are severe, with high correlations between pollutant levels and respiratory and cardiovascular diseases. The main contributors to poor air quality are industrial pollution, automobile emissions, and construction dust [16]. Comprehensive reviews of air quality in Dhaka emphasize the socio-anthropogenic drivers of pollution, including urbanization and industrialization. These reviews call for strategic models to improve air quality through clean technologies and stricter legal measures [17]. Low-cost air quality monitoring systems employing sensors have been suggested as viable solutions to track spatiotemporal pollution variability. These systems are particularly suited for regions like Dhaka with limited monitoring infrastructure [18]. The research works also show how wildly different the levels of air pollution are in urban and rural areas of Bangladesh. PM2.5 levels are linked with respiratory diseases including asthma and bronchitis which calls for strict pollution control measures and awareness campaigns [19]. As of now, there are some challenges in the monitoring of air quality in Dhaka including the scarcity of functional monitoring stations and competent personnel. To address these concerns, the use of enhanced technologies and the development of the capacity of environmental NGOs have been suggested. Such a plan involving governmental and non-governmental institutions is thus necessary for the management of air quality in the long run [20]. Finally, the impacts of pollution on human health especially in urban centers such as Dhaka has been well captured. This paper finds that pollutants such as PM2.5, NOx and SO2 are highly linked with morbidity and mortality hence calling for proper air management to protect lives [21]. When all the studies are put together, one gets a good picture of the state of air quality in Bangladesh and the importance of advanced monitoring techniques,

predictive modelling, participation of the public, and strict policies in the management of the problem.

2.3 Gap Analysis

Table 2.1 presents the comparison of exiting researches to the proposed system with focusing on the main difference. Some of the studies, for example Rahman, M. H. and Rahman, take a general approach looking at AQI for so many cities while the others only focus on certain areas of study thus, creating a void for a comprehensive multi-city study which the proposed system will fill. The work done in analysing seasonal and temporal trends is found in Bhuiyan, A and Hossain but the proposed system extends this research to a larger set of data. A novel feature included in the proposed approach is health impact assessment, which, as mentioned in Rahman,M. H., can be used to make links between AQI levels and health effects. Long-term trend analy- sis, discussed in Hossain, is the trend magnitude of which is further enhanced in the proposed system using sophisticated predictive models as ARIMA and LSTM. Not a single paper emphasizes on the nature of low-cost monitoring technology or public awareness and behavioral research that the presented approach aims to fill through the assessment of inexpensive monitoring options and integrating public education strategies. Therefore, the proposed methodology gives a comprehensive framework for identification and management of air quality concerns in Bangladesh, which can address the research gaps of this field.

Table 2.1: Gap Analysis with Previous Work.

Featurs	**Bhuiyan, A**	**Rahman, M. H.**	**Hossain**	**Rahman**	**My Work**
AQI Analysis Across Multiple Cities	No	Yes	No	Yes	Yes
Seasonal and Temporal Trends	Yes	No	Yes	No	Yes
Health Impact Assessment	No	Yes	No	No	Yes
Long-Term Trend Analysis	No	No	Yes	No	Yes
Integration of Predictive Models	No	No	No	Yes	Yes
Low-Cost Monitoring Technologies	No	No	No	No	Yes
Public Awareness and Behavioral Study	No	No	No	No	Yes

2.4 Summary

The background chapter contributes greatly towards putting into perspective the material and relevance of existing research on trends in air quality in Bangladesh. This highlights that air pollution is a disaster in the country more so in big cities for instance Dhaka since

AQI is continually exceeding legal limits. This chapter assesses current uses and studies relevant to air quality monitoring and prediction and associated health risks. While many interventions are directed toward local approaches to air quality assessment or towards particular pollutants, few endeavors are devoted to multi-city AQI trends analysis, seasonal and temporal changes, and to the use of a low-cost monitoring tool. Moreover, less effort has been devoted to the relationship of air quality indexes with possible health effects and the popularization of health-related activities. Hence, this chapter reveals the importance of adopting a multi-dimensional approach in analysing data for air quality and also in offering possible solutions to the problem for decision makers within policy domains. The background chapter provides the logical foundation for an extensible system that leverages state-of-art prediction models, low-cost mind-devouring technologies, and engaging stakeholders' approaches to prevent and overcome the air quality problems in Bangladesh by pointing out the research and application lacunas.

Chapter 3

Research Methodology

3.1 Methodology

3.1.1 Overview

The chapter methodology affords a description of the method used with regards to identifying changes in air quality, and forecasting future AQI trends in Bangladesh. This research is intended to employ data gathering and acquisition, data analysis, and prediction model which will provide a comprehensive insight into the nature of air quality. The geographical locations involved 13 cities over several years and data was collected from reliable sources and underwent proper pre-processing to enable missing values, outliers and inconsistent data to be dealt with appropriately. The nature of AQI distribution is analyzed using statistics and machine learning techniques to investigate seasonal and temporal characteristics and the frequency of high AQI events. Hence, the ARIMA, Linear Regression, and LSTM models were used for predicting AQI levels. The effectiveness of each of the models was evaluated with reference to the RMSE and accuracy in order to make sound predictions. State of the art methods were used in creating trends and patterns to enable easier conveying of results. In order to solve low-cost monitoring systems' problems as well as further research obstacles, the study includes low-cost monitoring systems and explores the potential of public awareness campaigns as a part of a comprehensive strategy to cope with air pollution. This technique provides a four-dimensional approach to an assessment that incorporates findings in conjunction with recommendations for policy makers as well as stakeholders. By so doing, the research aims to provide a real life based study on how one can define and handle facets impacting air quality in Bangladesh hence enhance the health of the society.

3.1.2 Proposed Methodology

The methodology 3.1 of this type of research technique is to analyze and forecast the air quality index (AQI) in several cities of Bangladesh. Data collection is the first step of the procedure, and AQI data is collected from monitoring stations and source data obtained from the web. Data requires preprocessing of data cleaning and data normalization of missing numbers, outliers, and inconsistent-values.

This included the empirical distribution of AQI quality around the city, distribution

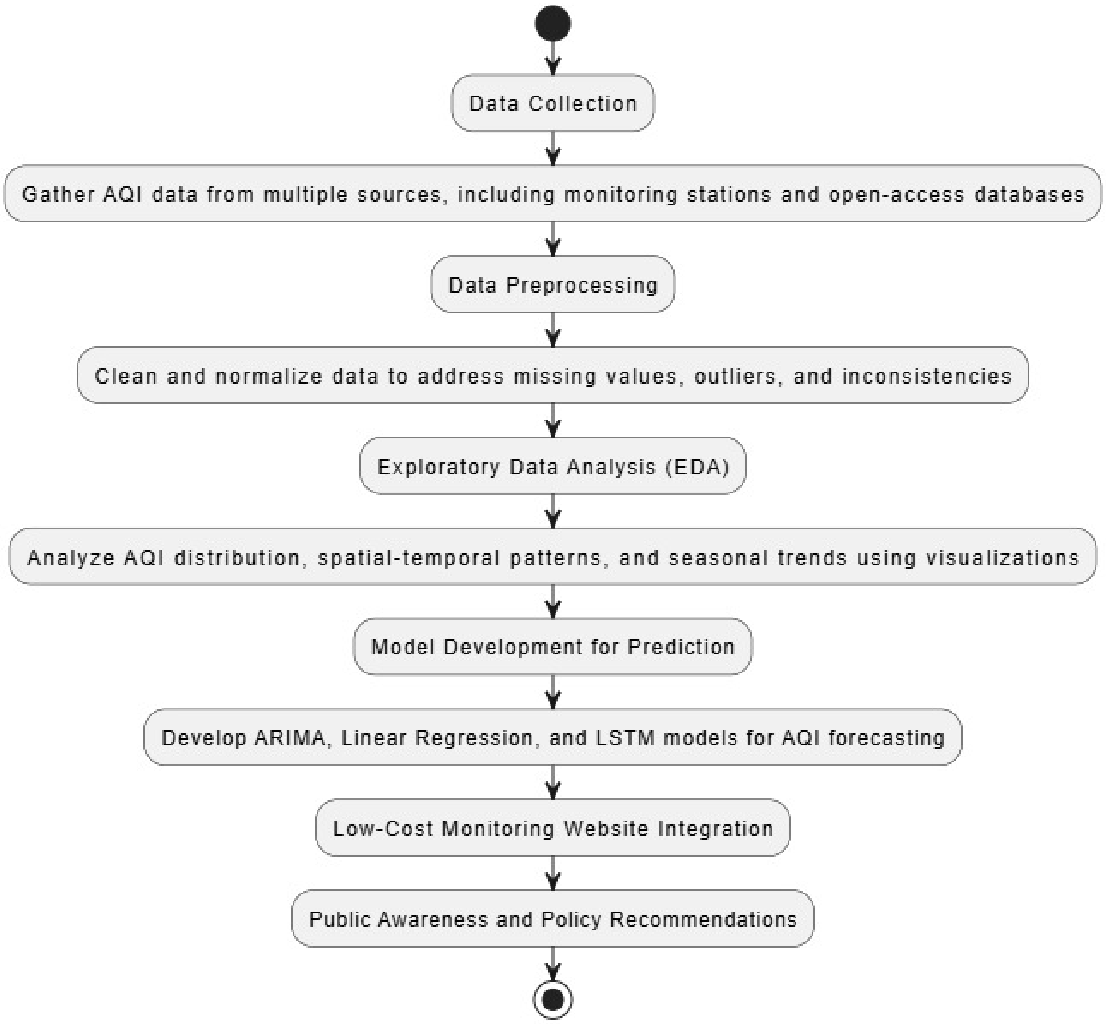

Figure 3.1: Proposed Methodology.

pattern based on the emporial zones and AQI quality distribution based on the seasons.. One of them is constitution of the model where we have selected predictive models like ARIMA, Linear Regression and LSTM for predicting AQI level based on historical details. There also a method of examining the characteristics of variation associated with time by time series decomposition and correlation to identify if there are seasonal and temporal variations in the pollution levels. The project also looks at the method of incorporating low cost monitoring technologies to provide affordable ways of constant AQI monitoring. Besides, advertisement campaigns as well as policy recommendations from the corporates regarding health consequences and anti-pollution measures. Finally, the results are presented in a form of live web-based visualization and complete report for better understanding from the policy makers, researchers and the public. This is the overall type of approach is intended to provide the overall picture of AQI fluctuations, health risks, and opportunities for policy action that would improve air quality in Bangladesh.

3.1.3 Functional and Nonfunctional Requirements

The functional requirements of the proposed system are as follows in order to offer an additive measure for AQI monitoring, analysis as well as predicting. Some features involve the automated extraction of data from many cities, AQI computation, AQI prediction services based on models such as ARIMA, Linear Regression, LSTM, and the visualizations of collected data in interactive dashboards. The system must also provide the analysis of historical data to view trends by seasons and times of day, forecast the air quality in the future, and select cities with the worst air quality. Also, it must incorporate an affordable technology for the live monitoring of AQI and spread public consciousness through recommendation for improvements in health and activities. The non-functional requirements are centred on the reliability of the system and the way it will perform when deployed. These include high accuracy in data processing and forecasting, ensuring there is low delay for providing the AQI data in real-time, and flexibility of capacity in relation to the amount of new data when the number of cities or sensors is added. It should also be easy to use: the home page should have clear links to the most important sections so that a first-time visitor will be able to quickly find what they're looking for without needing to click through and navigate a complex hierarchy of sub-pages It should also be easy to access and should appeal to a variety of users such as scholars, policymakers, and citizens in general. The security and privacy of individuals must be maintained especially when processing data from monitoring instruments proper technique like encryption and data anonymization must be employed. The system should also be able to; The system should also be very robust and capable to handling very large amounts of data and still perform optimally despite this burden.

3.1.4 Context Diagram

The context diagram 3.2 for the AQI Monitoring and Forecasting System presents an illustration of the relationships that exist between several external entities and the system. Public Users and admin are the major actors in the diagram and they have specific roles as well as data access rights. Public Users access the system in order to view actual AQI data of various cities. Administrators input prior AQI data for further analysis, whereas Policymakers use the system to access comprehensive reports with information that supports decisions about the management of air quality and formulation of air quality policies. The system itself integrates several external components: Sensor Network Monitoring Stations for Real-time AQI data with GPS coordinates, Historical AQI data for pattern analysis, Weather Data web-service for weather conditions affecting the air quality, Health Data web-service for comparing impact of Air Pollution with the Health of people. This figure shows how these organizations synchronously function to develop an adequate monitoring and forecasting mechanism in Bangladesh for air quality, aiming to encourage evidence-based decision making for air quality.

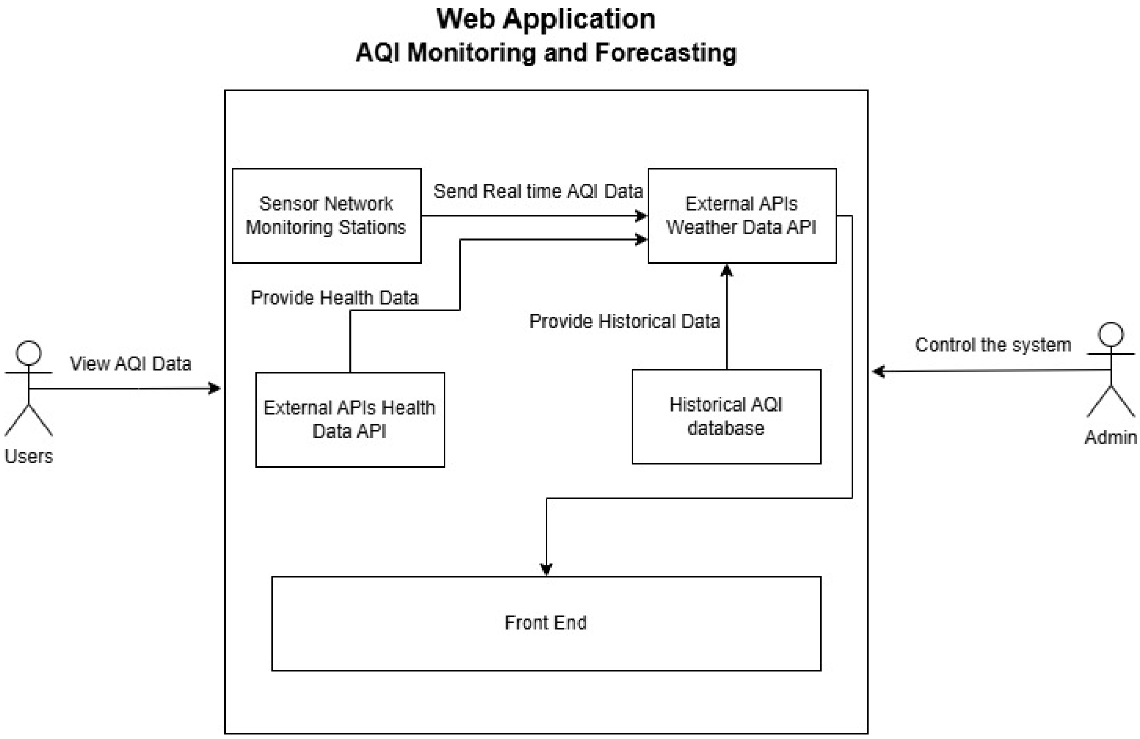

Figure 3.2: Context Diagram of Website.

3.1.5 Data Flow Diagram Level 1

The Data Flow Diagram 3.3 Level 1 of the AQI M&F System (DFD) showcases the inputs and output, as well as the data flows that are central in the real-time AQI monitoring and forecasting. The major tasks of this graphic are as follows: Data acquisition of AQI Data Processing of historic data Integration of weather data Analysis of health impact Data forecasting Data reporting. The AQI data collection method is a process of supplying actual signal data received from the sensor and storing it in the actual real-time AQI database. This information is made easily retrievable to the community to ensure that anyone is able to view the current data regarding the air quality. The historical data processing method obtains the past AQI data and store them in historical AQI database for future comparison. Weather data integration strategy entails the inclusion of actual weather data from the Weather Information API that is in turn used to analyse the effects of meteorological factors on air quality. Likewise, the Health Impact Analysis approach uses the data from Health Information API to establish the relationship between air quality and health issues, especially those pneumonia-related diseases associated with low quality air. The extended forecasting technique incorporates complex models of forecasting including ARIMA, Linear Regression, and LSTM, for predicting the future levels of AQI for long-term planning of air quality. In the last step of the process, reporting & insights in its final form synthesizes all the collected data, rejects, and health analysis into complete reports and recommendations then forwarded it to policy makers and educators to enhance the decision making & public health activities. This flow of information assist in facilitating competent responses to the air quality challenges.

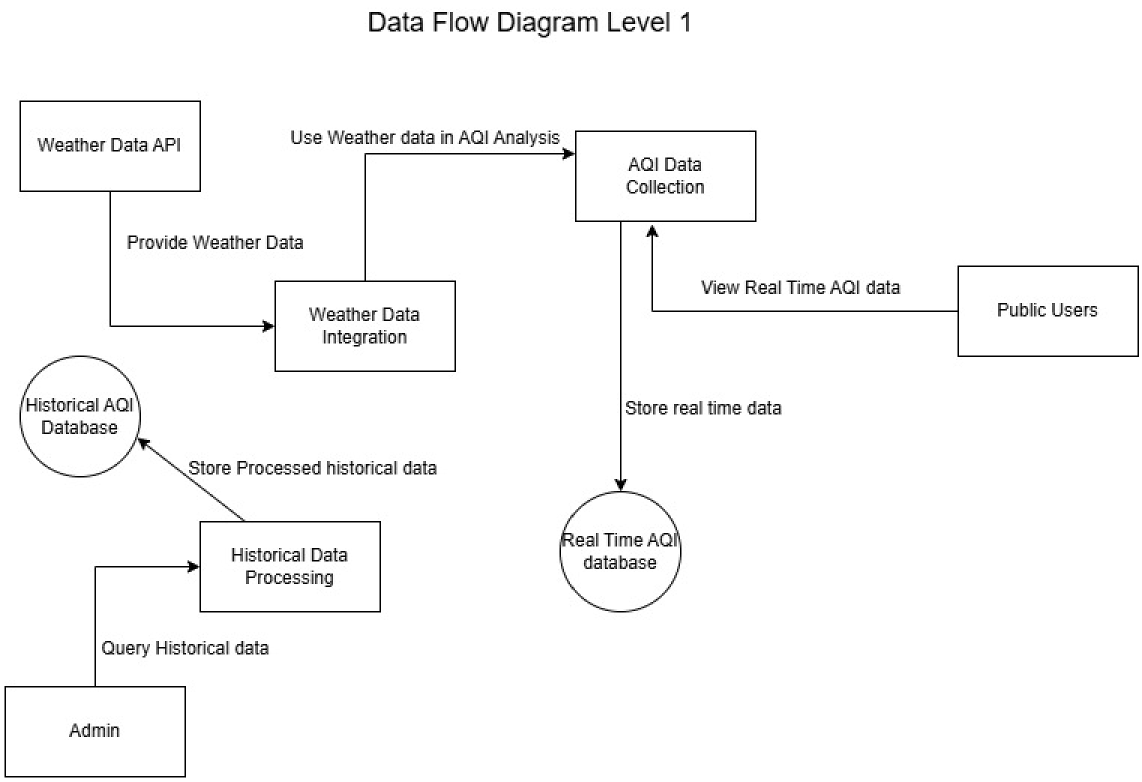

Figure 3.3: Data Flow Diagram Level 1 of Website.

3.1.6 UI Design

UI design plays a significant part in the success of this AQI monitoring and forecasting research as it directly effects user experience, engagement, and accessibility of the system. A well-designed user interface ensures that the data and insights linked to air quality are presented in a clear and straightforward manner, making it easy for users to interpret complex information. For example, a dashboard providing real-time AQI values, historical trends, and anticipated data should be visually appealing and easy to browse, allowing users to rapidly comprehend the status of air quality in different cities. Interactive elements include filtering options, city selection, and data visualization (graphs and charts) promote user interaction, enabling users to investigate individual data points or trends. Additionally, accessibility is a critical concern; the UI must be responsive and user-friendly across different devices, ensuring it's usable by a large audience, including people with low technical expertise. By providing a seamless and efficient interface, the UI design boosts the value of the study, making air quality data more accessible and actionable for policymakers, researchers, and the general public.

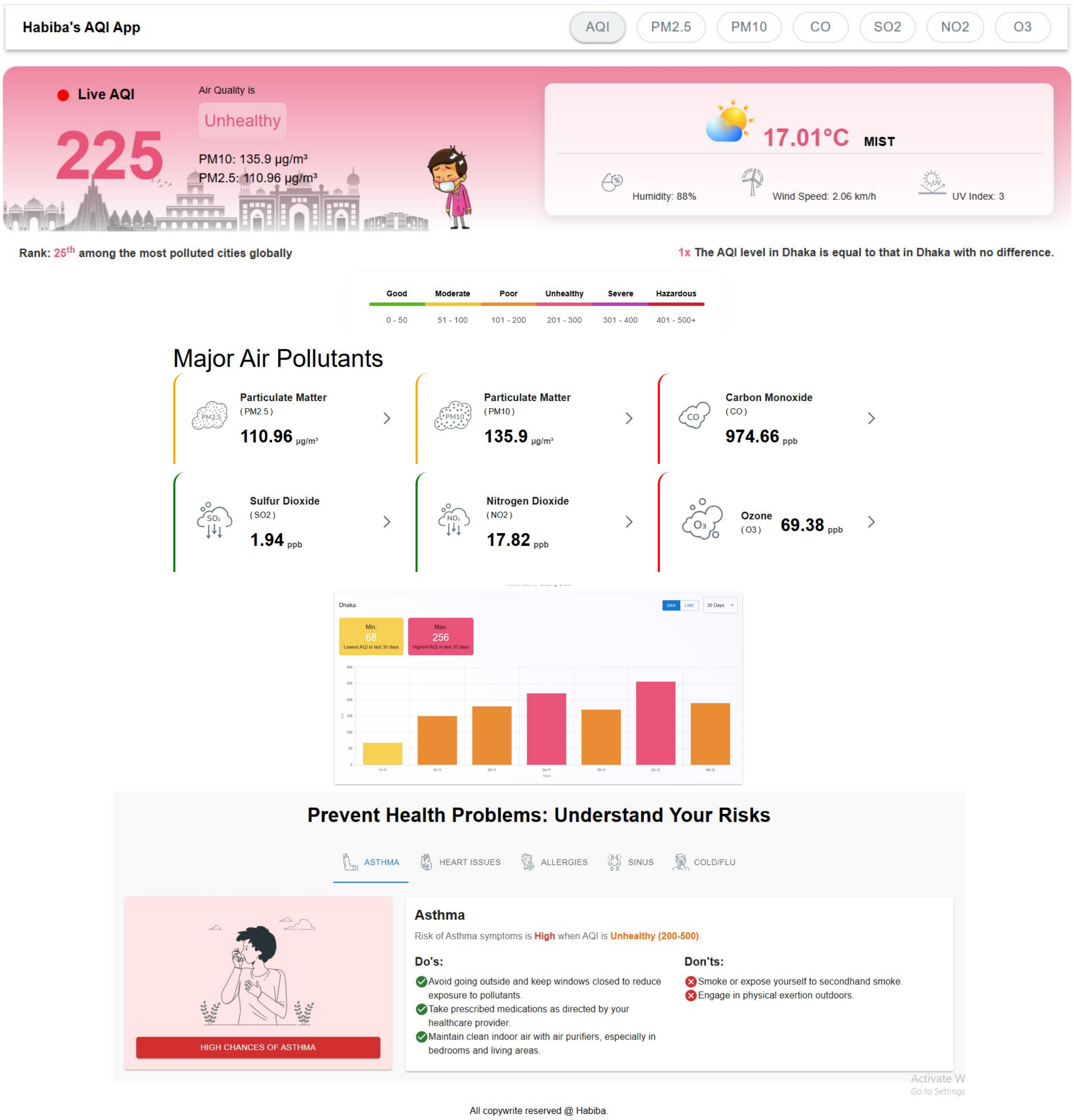

Figure 3.4: UI Design of Website Home Page.

The 3.4 of Habiba's AQI app is well-structured and visually appealing, effectively providing air quality information across four key categories: tags include headlines, real-time AQI, key pollutants, archive, and health advice. The current place and AQI another instance is also shown on the header as the name of the application, the current, and AQI level, which is in this case 225, is exhibited in a bar with colors showing the levels of air quality ranging from green, which represent good quality to red depicting dangerous quality. Analysing the live AQI segment it highlights the current AQI, its ranking globally and a colour coded scale of intensity in real time In case of the main pollutant segment, key pollutants including PM 2.5 and NO2 with directional trend arrows- up or downwards. A bar chart of historical data area shows AQI readings over the last 30 days so that users understand the pattern. Last but not least, the list of precautions that can be taken with regard to health contains measures that would help preventing possible effects of pollution in case such situation occurs, particularly in case of asthmatic patient. It pairs clear images and icons.

The 3.5 design for Habiba's AQI App is thoughtfully structured into five essential sec-

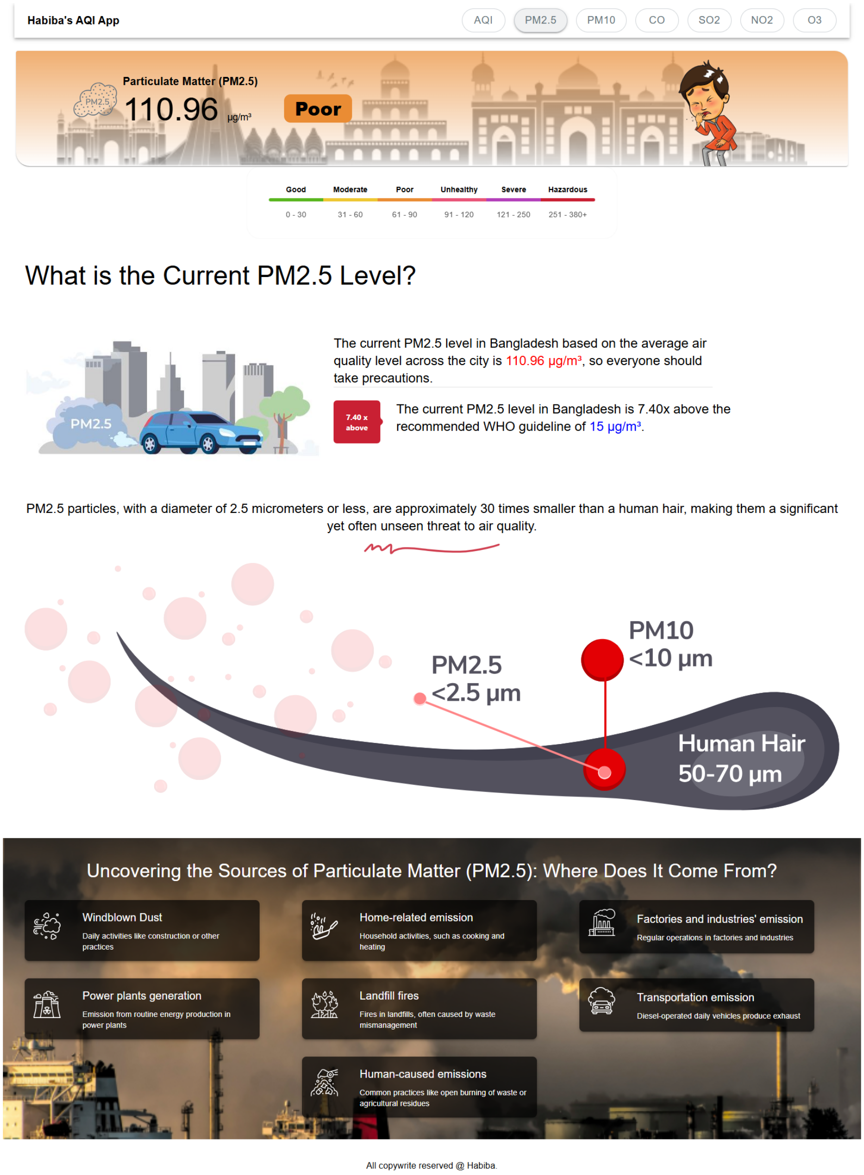

Figure 3.5: UI Design of PM 2.5 Page

tions: the AQI, current values of each AQI, major air pollutants, past days' data, and health warning. The header section then presents the app name, location, for instance, Dhaka, and the Total AQI, for instance, 225 which the bar will show depending on its severity that ranges from green for good quality to red for dangerous quality. The live AQI area displays the current AQI number, the rank out of all other global cities, and a bar indicating its severity level. In major pollutants section, key pollutants including PM 2.5, PM10 and NO2 are provided with arrow indicating whether the concentration is rising or falling. The AQI is displayed in a bar chart with data presented for the past thirty days so as to show trends. Last but not the least, under health advice the measures, essential to minimise the effects of Air Pollution have been provided particularly about the persons suffering from Asthma. In general, the design combines comprehensible images, crucial information, and suggestions for improving knowledge and performance.

In the 3.6 Habiba's AQI app, the UI design comprises of five key areas and air quality

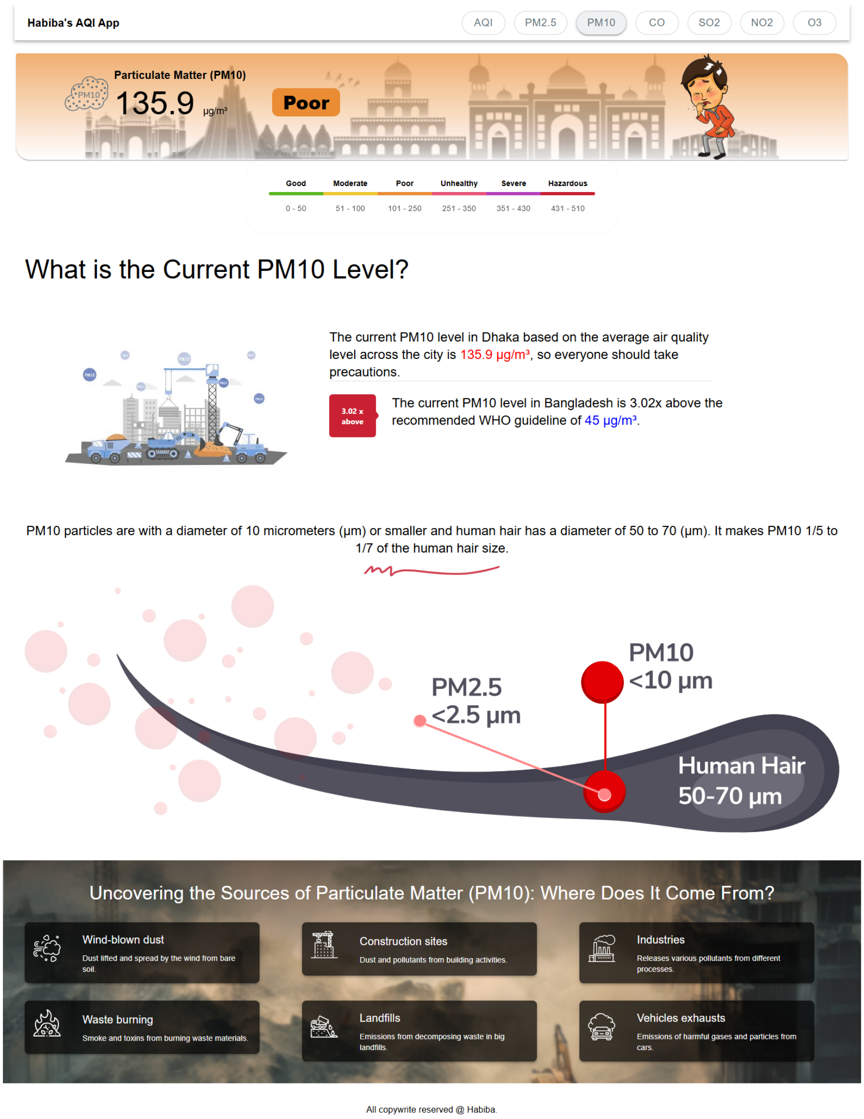

Figure 3.6: UI Design of PM 10 Page

details are well arranged and well aligned in an aesthetical manner. The index reveals the app name, the place (for example, Dhaka) and AQI number (for instance 225) using a bar to demonstrate the level of air quality that could range from green as a sign of good air quality to red signifying bad air quality. The current AQI section provides a coloured bar along the scale of the current AQI value, the rank in the world and aggression. The Major Air Pollutants category show the amount of the pollutant like PM 2.5, PM10 and NO2 and up or down arrows to indicate that these levels are rising or falling. The previous AQI values are presented by 30-bars where users can see historical trends as well. Last, the Department of Health Issues Prevention outlines measures to prevent the effects of air pollution towards patient's suffering from asthma. Including a number of graphical elements based on its simplicity, necessary information which may be important to be indicated in the process of using the application, as well as proper advice, the application optimizes and improves the condition of users and their choices.

The design of 3.7 Habiba's AQI app has a neat and clean UI and it delivers the air

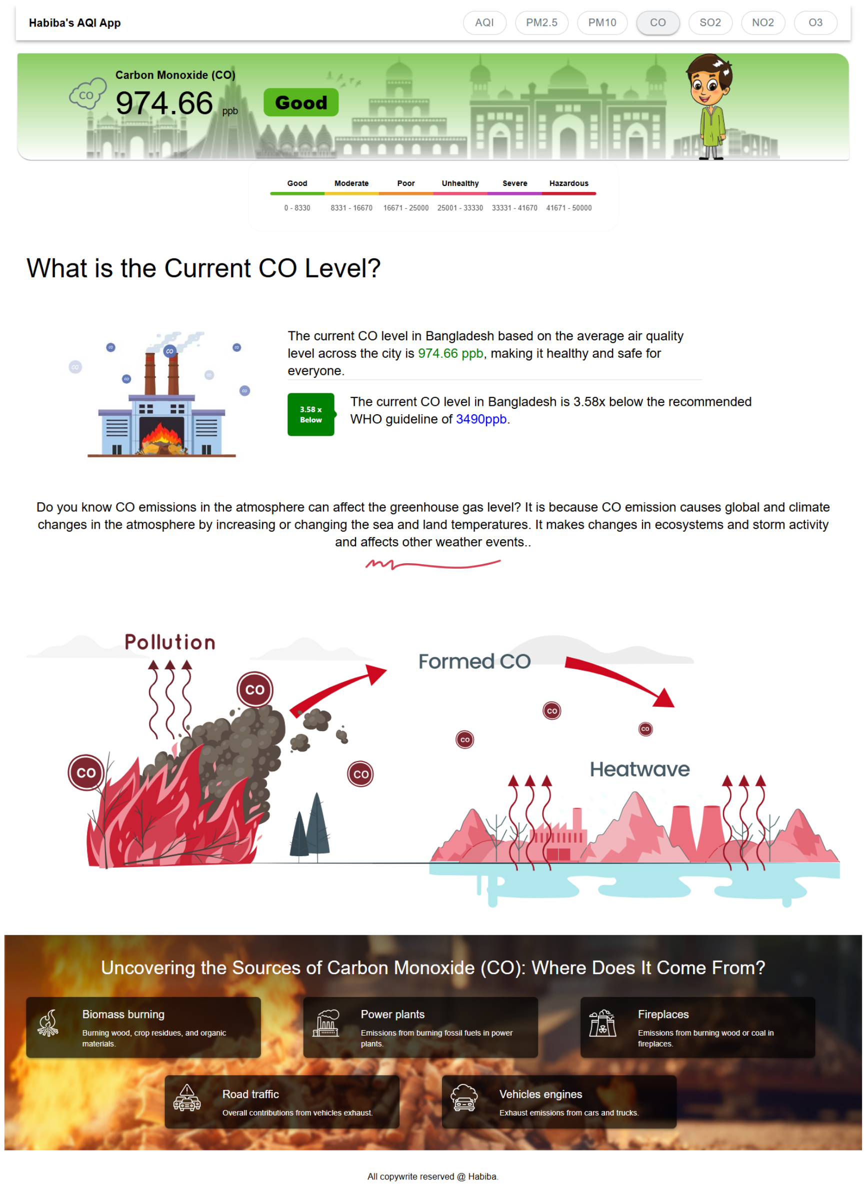

Figure 3.7: UI Design of CO Page

quality indices and factors in five main sectors in an orderly manner. At the top of the header, the current city (e.g., Dhaka), and AQI (e.g., 225) in a colored bar showing the intensity of Air quality from green (good) to red (danger). The currently featured live

AQI highlights the current AQI, its worldwide ranking, and graphic rating. Major Air Pollutants shows the current air quality in basic pollutant as PM 2.5, PM10, NO2 etc with increasing or decreasing bar shown through arrows. On this interface, the historical AQI data is represented by a bar chart which shows the trend of the AQI of the last 30 days with more convenience. Asthma UK's Department of Health Issues Prevention offers information to reduce the impact of polluted air on the community and particular step-by-step advice for the asthmatic. Making use of clear and comprehensible graphics, relevant information, and useful and informative health tips, the application gives a clear reference and assistance in decision-making.

The figure 3.8 Habiba's UI design AQI app extremely informative and easy to understand

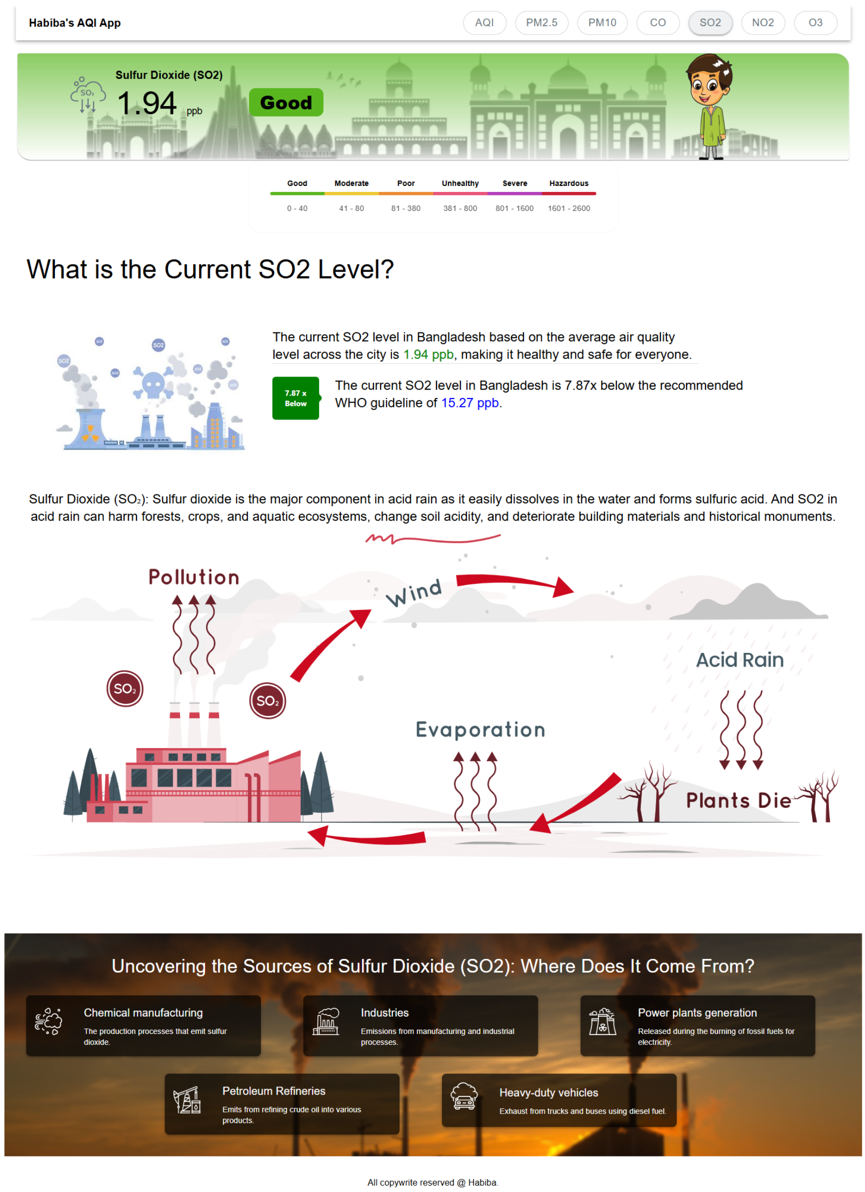

Figure 3.8: UI Design of SO2 Page

because the air quality information is broken into five important segments. The header contains the name of the app, the current location, i.e. the city (e.g., Dhaka), and the AQI level (e.g., 225) accompanied by a bar that may be green (good), yellow (sensible), orange (unhealthy), red (dangerous), etc. The major air pollutants contains the bars of the pollutant levels, including PM 2.5, PM10, and NO2, and the arrows showing the rising or decreasing trends. Month AQI data is depicted from a bar chart representing trends of that data over the past thirty days. The Department of Health Issues Prevention provides tips on reducing the effects of poor quality air especially on asthmatic people. Simple graphics and picture messages containing important information and effective tips let the users to get acquainted with the app as a complex tool to recognize and adapt to the changes in air conditions.

The figure 3.9 of Habiba's AQI app UI design is also successfully separated into five core

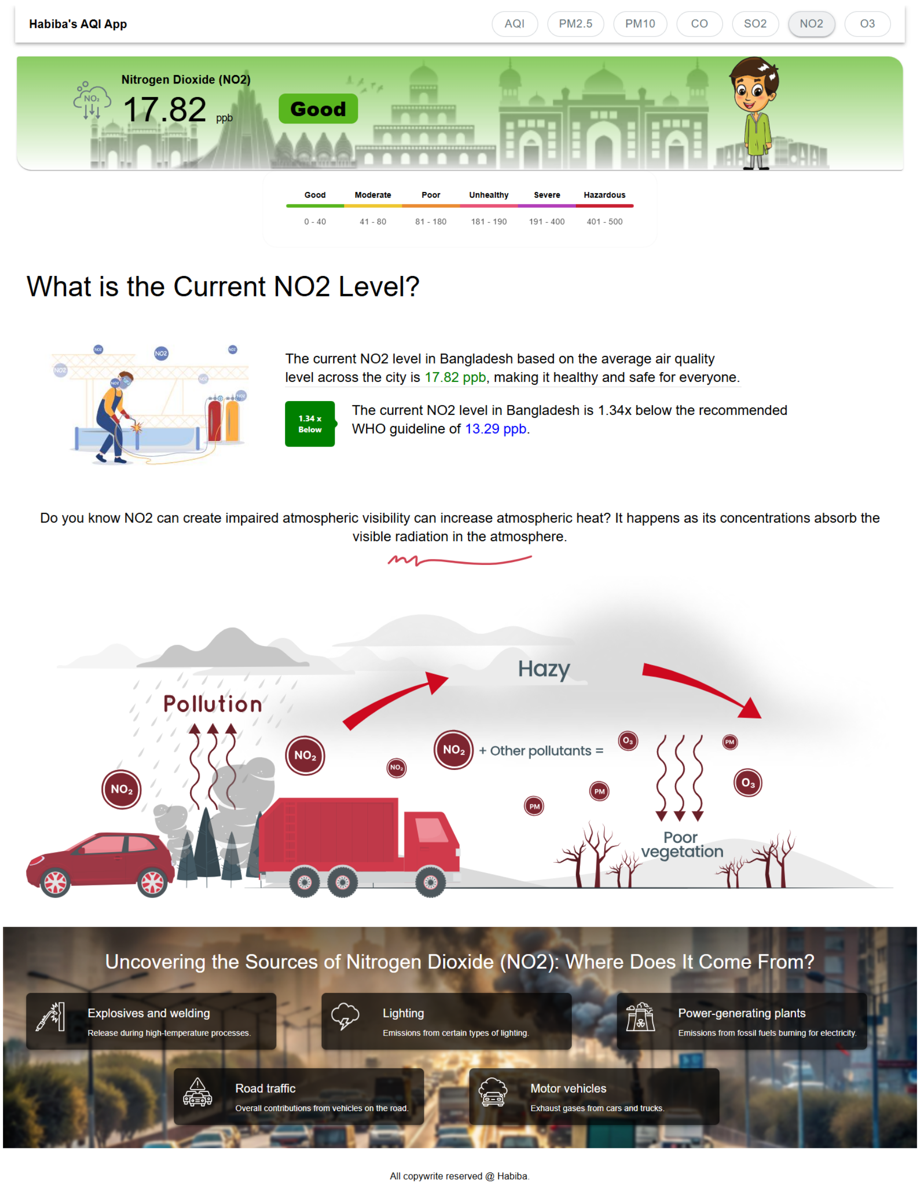

Figure 3.9: UI Design of NO2 Page

sections, which delivers air quality data effectively and aesthetically. The header of each page displays the name and the logo of the app, as well as the AQI level of the current location (e.g., '225' for Dhaka) or AQI intensity bar to show the extent of the air quality- green, orange, red for good, moderate to dangerous. The major air pollutants category shows the concentration of pollutants such as, PM 2.5, PM10 and NO2 with up or down arrow to depict the high or low pollutant level respectively. The historical AQI data can be represented by a bar chart to display the monthly trend of the last one month. Finally, the Department of Health Issues Prevention offers ways on how to reduce the impact of air pollution to the people particularly those suffering from asthma which are doable how-to's. By employing shared intuitive images, critical information, and straightforward recommendations.

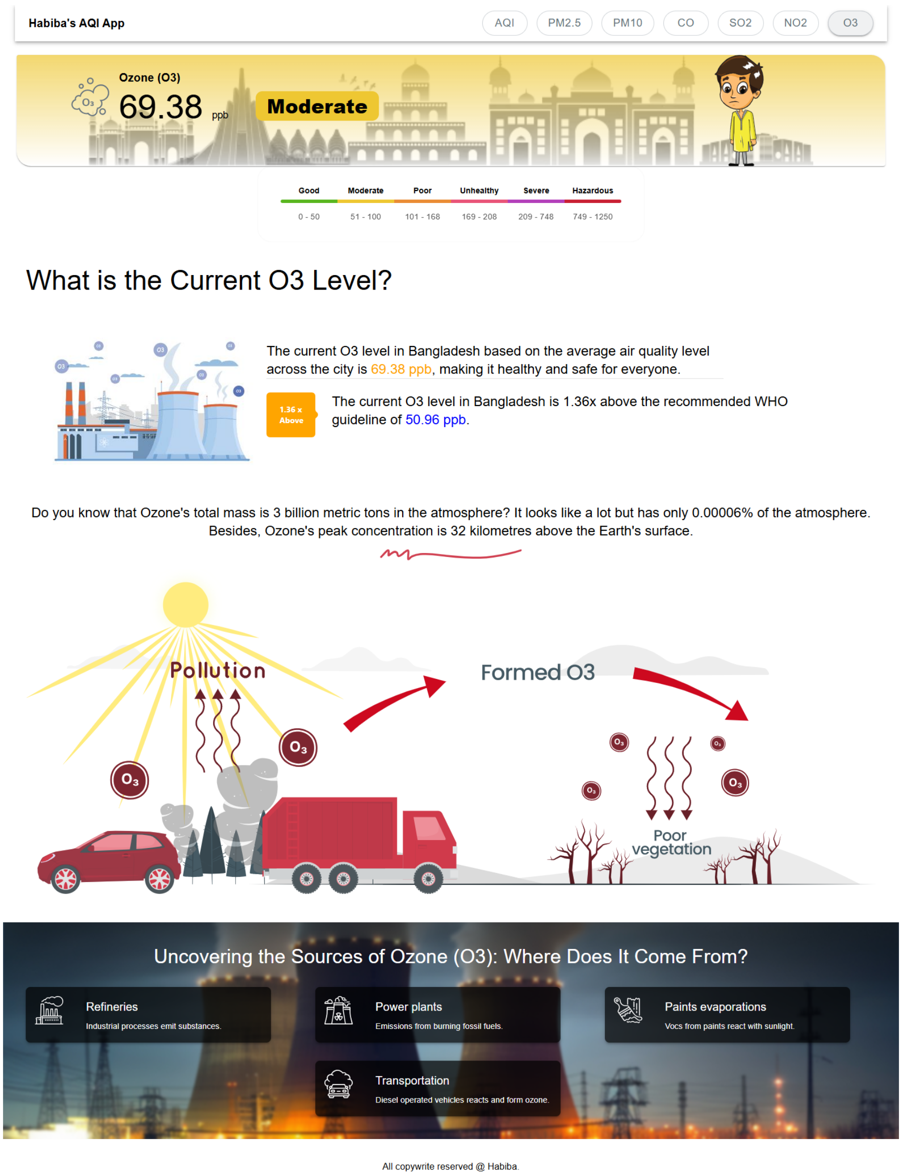

Figure 3.10: UI Design of O3 Page

The 3.10 UI design is well organized and aesthetically pleasing making it easy to find the air quality. It consists of a coloured bar header that shows the place (Dhaka) and AQI (225), the real-time AQI category with the intensity pointer, polluter levels with arrow that shows direction, a historical 30-day AQI chart, and effective health precaution advice for pollunter protection particularly for asthmatic persons.

3.2 Detailed Methodology and Design

Data Collection Procedure: Data collection strategies for AQI monitoring and forecasting systems involve systematic and accurate gathering of air quality data from multiple sources. Initially, appropriate datasets are identified from public and private environmental agencies, such as the Department of Environment (DOE) or open datasets such as the World Air Quality Index or Kaggle repositories. Historical AQI records have been collected throughout 13 cities in Bangladesh, confirming numerous year-round data to capture seasonal and long-term trends. Information such as pollutant concentrations (e.g., PM 2.5, PM10, CO, NO2, SO2), weather information, and timestamps are needed parameters. Once sourced, the raw data goes through pre-processing stages to handle missing numbers, discrepancies, and outliers, assuring data quality. The cleansed data is then standardized and arranged into a structured way consistent with the predictive models. Additional information, such as geographic and population density statistics, can be integrated to aid in the research. This method stresses ethical concerns, ensuring that data use complies with acceptable regulations and laws. Finally, the processed dataset is properly stored and backed up for further analysis and modeling. This methodical technique ensures the integrity, quality and relevancy of the data to achieve the project objectives.

Data Preprocessing: Data preparation is a key stage in the AQI monitoring and forecasting system, which ensures that the data obtained is clear, consistent and ready for analysis. The procedure begins by addressing missing values, where techniques such as average estimation, interpolation, or deletion of incomplete records are utilized, depending on the size and nature of the missing data. Outliers, which possibly distort the data, are recognized using statistical methods or eye inspection and are dealt with appropriately, either restricting or modifying the maximum number. The data is then normalized or scaled to put various variables into a common range, especially when machine learning models are utilized to maintain uniformity in the study. Fake records, if any, are found and destroyed to avoid repetition. Incompatible data formats, such as distinct date or time representations, are standardized for easy integration. In addition, categorical data, such as pollution intensity levels, are encoded in numerical representations using techniques such as one-hot encoding or label encoding for interoperability with analytical models. The data is also separated into training and test sets in preparation for predictive modeling. This comprehensive pre-processing guarantees that the dataset accurately reflects real-world air quality patterns for dependable and effective monitoring and forecasting.

Feature Extraction: Feature extraction is a crucial activity in AQI monitoring and forecasting systems, which enables the discovery and exploitation of significant properties from raw data to boost model performance. Key techniques applicable to this project include Principal Component Analysis (PCA) which minimizes the scale while keeping crucial information, allowing to focus on the most significant features. The statistical feature extraction is applied to calculate metrics such as mean, variance, standard deviation, and trends throughout time from the AQI data, capturing the relevant temporal patterns. Domain-specific factors, such as pollutant concentrations (e.g., PM 2.5, PM10, CO, NO2, SO2) and weather conditions (e.g., temperature, humidity, wind speed) are crucial as they affect AQI levels. Time-series feature engineering is utilized to construct lagged features, moving averages, and seasonal indicators, which reflect historical and periodic variations in air quality. Correlation analysis is used to locate and keep variables of high relevance to the AQI forecast, reducing redundant or poorly matched features. For predictive modeling, techniques such as Fourier transform or wavelet transform are useful for decomposing time-series data into frequency components, which can highlight patterns that are not obvious in the raw data. In addition, text feature extraction can be performed if public health reports or social media data are employed, techniques such as TF-IDF or embedding are used for context analysis. These methods guarantee a comprehensive display of data to support accurate AQI analysis and forecast.

Model Training: In the model training phase, several machine learning and time-series forecasting methods are used to predict the air quality index (AQI) and its associated trends. The key algorithms used in this study are Linear Regression, ARIMA (Auto Regressive Integrated Moving Average) and LSTM (Long Short-Term Memory) networks. Linear regression: It is a statistical method used to predict a continuous dependent variable (AQI) based on one or more independent variables. (such as pollutant concentrations, temperature, and humidity). The linear regression equation is:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_n X_n + \epsilon \tag{3.1}$$

Where Y is the dependent variable (AQI), $X_1, X_2, \ldots, X_n$ are the independent variables (pollutant concentrations and other environmental factors), β_0 is the intercept, $\beta_1, \ldots, \beta_n$ are the coefficients for the respective independent variables, and ϵ is the error term. ARIMA: ARIMA is a time-series forecasting technique that combines the elements of Autoregressive (AR) Differentiating (I) and Moving Average (MA) to present temporal data. The ARIMA model is presented as follows:

$$Y_t = \alpha + \sum_{i=1}^{p} \phi_i Y_{t-i} + \sum_{j=1}^{q} \theta_j \epsilon_{t-j} + \epsilon_t \tag{3.2}$$

Where Y_t is the AQI at time t, α is a constant, ϕ_i are the autoregressive coefficients, θ_j are the moving average coefficients, and ϵ_t is the error term. LSTM (Long Short-

Term Memory) LSTM is a form of recurrent neural network (RNN) designed to handle time-series data with long-range dependencies. LSTM networks use memory cells to store information over long periods of time, which is important for predicting AQI based on prior data. The LSTM model is presented using the following equation for cell conditions.

$$c_t = f_t \cdot c_{t-1} + i_t \cdot \tilde{c}_t \tag{3.3}$$

Where c_t is the current cell state, f_t is the forget gate, i_t is the input gate, $\tilde{c}_t$ is the candidate memory cell, and c_{t-1} is the previous cell state. Each of these algorithms is trained using past AQI data and environmental parameters to generate predictions. Models are tested and fine-tuned to increase predictive accuracy, and their performance is evaluated by various evaluation systems such as mean squared error (MSE) and R-squared.
Model Selection: At the model selection stage, we examine the performance of various algorithms (Linear Regression, ARIMA and LSTM) to identify the best accurate model for forecasting the air quality index. (AQI). This method assesses the performance of each model based on specified evaluation criteria that measure the model's capacity to correct prediction errors and generalize to new data. The most prominent performance metrics used for regression activities such as AQI forecasting are Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Mean Absolute Error (MAE) and R-Squared (R2). The mean squared error (MSE) represents the average of the squared discrepancies between the projected values and the actual values. This is a commonly used statistic for regression work, where smaller values suggest stronger model performance.

$$\text{MSE} = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \tag{3.4}$$

Where n is the number of data points, y_i is the actual value of AQI, and $\hat{y}_i$ is the predicted value. The root mean squared error (RMSE) is the square root of the MSE, which yields an error metric with the same units as the target variable. (AQI). RMSE gives an insight of how divergent the predictions of the model are from the average actual values.

$$\text{RMSE} = \sqrt{\text{MSE}} \tag{3.5}$$

The mean absolute error (MAE) assesses the average of the absolute differences between expected and actual values. It is less susceptible to significant errors than MSE and is often utilized when extraneous ones are not the major concern.

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \tag{3.6}$$

R-square quantifies the fraction of the variance of the dependent variable (AQI) that is predictable from the independent variable. It goes from 0 to 1, where a value near to 1 denotes a better fit of the model to the data.

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \bar{y})^2}{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \tag{3.7}$$

Where $\bar{y}$ is the mean of the actual values. o determine the best suited model for AQI forecasting, the effectiveness of each algorithm is evaluated using the above measurements. Linear regression is useful for fundamental linear relationships, but may function less well for extremely non-linear data. ARIMA models are designed for time-series forecasting and are ideal for AQI forecasting based on prior data. However, ARIMA may struggle to capture non-linear linkages or complicated patterns. On the other hand, LSTM models, being a form of recurrent neural network, excel in capturing sequential long-term dependencies and non-linear relationships, which makes them more robust for time-series prediction in AQI forecasting. After reviewing the models, the one that minimizes the MSE, RMSE and MAE, maximizes the R2 value, is picked as the final model for deployment. In this circumstance, the LSTM model can generate the most accurate predictions for AQI considering its capacity to manage time-series data and capture non-linear correlations.

Model Deployment: The Model Deployment stage involves integrating the trained model into a real-world context to deliver real-time or batch forecasts for Air Quality Index (AQI) forecasting. This comprises numerous key steps: first, the model is packaged using formats like Pickle or Joblib for Python models, or HDF5 for LSTM-based models. Next, the deployment infrastructure is put up. The model is then incorporated into the AQI monitoring application via APIs, allowing for seamless communication between the application and the model for real-time predictions. Once implemented, the system's performance is regularly monitored, focusing on forecast accuracy and computing efficiency. It also needs ongoing model retraining to guarantee the forecasts remain relevant as environmental conditions develop. Additionally, the model is integrated into a user-friendly web interface that allows users to monitor current and anticipated AQI data, track trends, and make informed decisions based on the forecasts. The model implementation assures that the AQI forecasting system is scalable, dependable, and capable of giving significant insights to users, helping to enhanced public health and environmental decision-making.

Statistical Analysis: Statistical analysis in this AQI monitoring and forecasting research is fundamental for extracting meaningful insights from the collected data and drawing accurate conclusions about air quality patterns and trends. It employs multiple statistical tools to understand the distribution, central tendency, variability, and interactions among the various components that determine air quality. Descriptive statistics, such as mean, median, mode, standard deviation, and range, are used to characterize AQI data and to highlight similar patterns across different cities. Time-series analysis, which encompasses techniques such as moving averages, seasonal breakdowns and trend analysis, aids in comprehending the temporal changes in AQI and discovering seasonal or cyclical trends.

Figure 3.11provide a complete study of many variables using histograms with overlay

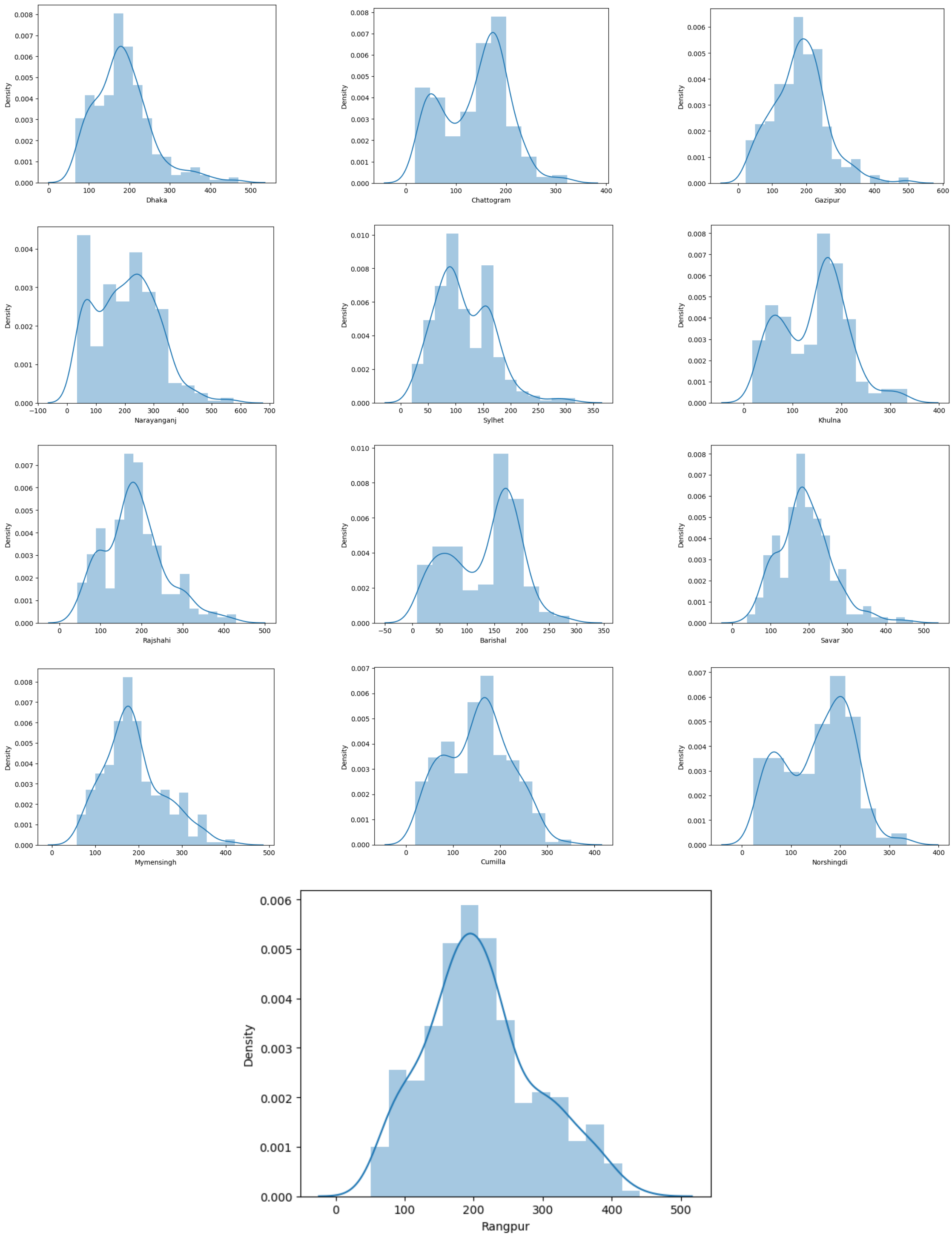

Figure 3.11: AQI Desity Curves 13 Main City's in Bangladesh.

density curves, indicating that most distributions with center peaks and symmetrical tails are virtually normal.It has 13 values.They are: Rangpur, Comilla, Narsingdi, Chittagong, Mymensingh, Savar, Barisal, Rajshahi, Khulna, Sylhet, Narayanganj, Dhaka and Gazipur. The average value of variables such as Rangpur, Comilla, Mymensingh, Savar, Barisal, Rajshahi, Khulna, Sylhet, Narayanganj, Dhaka and Gazipur is normally around 200, with the range varying between 0 and 600. These distributions show that the data may represent measurements such as temperature, rainfall, or population size, with the normal distribution indicating a central tendency and fewer observations at extremes. For further insight, descriptive statistics such as mean, median, mode, standard deviation, and quartile can be produced to summarize the core trend and distribution of the data. Statistical experiments can compare these variables across multiple categories or populations, while visualizations such as box plots and scatter plots can explore inter-variable interactions. If a variable denotes a category, histograms can depict the numerical data between those groupings. The research emphasizes that understanding the probability density functions of these variables is necessary to explain the possibility of observing specific values. For example, the distributions of Rangpur and Khulna are centered at 200, with ranges of 0-500 and 0-400, respectively. Similarly, a large range (0-600) can be found in Gazipur while keeping the regular size. The study underlines the need of applying statistical techniques to extract insights, with an emphasis on the impact of distribution for practical knowledge in demographic, climatic, or categorization contexts. Data from several regions (Rangpur, Narsingdi, Comilla, Mymensingh, Savar, Barisal, Khulna, Narayanganj, Gazipur, Chittagong, Rajshahi, Sylhet and Dhaka) Figure 3.12 show a wide variety of distributions, most of which exhibit a wide range and slight right-diagonal patterns. Boxplots reveal data spanning from a minimum of 50 to a maximum of between 250 and 450 across regions with a noticeable interquartile range (IQR) of 50 to 100 units, which in most cases indicates a medium to large spread. The regions of Rangpur, Savar and Gazipur show a higher prevalence, with more consistent figures in some areas such as Barisal and Khulna. Outliers are often detected in all boxplots, often above the upper whisker, showing exceptional or inconsistent values in certain areas, especially in Rangpur and Gazipur. The diagonal distribution suggests that most data points are focused on the bottom edge, but occasionally there are higher values in the tail stretch. This shows that while most data is grouped within a certain range, extraneous and extreme values have a considerable impact on determining the distribution. The absence of outliers in some areas (such as Barisal and Khulna) suggests a more regular trend in those areas. Overall, there seems to be some normality in the data in several regions, but in others the presence of outliers and skewed structures indicates the possibility of anomalies or above-average occurrences. Understanding the unit and context of this information will provide greater insight into its significance and relevance. This Kernel Density Estimates (KDE) 3.11 image depicts the density distribution of air quality index (AQI) data in 13 cities in Bangladesh. The colour gradient, paired with the contour line, illustrates areas of varying AQI intensity, with darker patches and tightly packed contours showing hotspots of increased AQI concentration. These hotspots possibly correspond to prominent cities in Bangladesh, including Dhaka, Chittagong, Gazipur, and Narayanganj, which are noted for their high population density and industrial activities. On the other hand, lighter areas with scattered outlines may reflect cities such as Sylhet or Rangpur, where AQI concen-

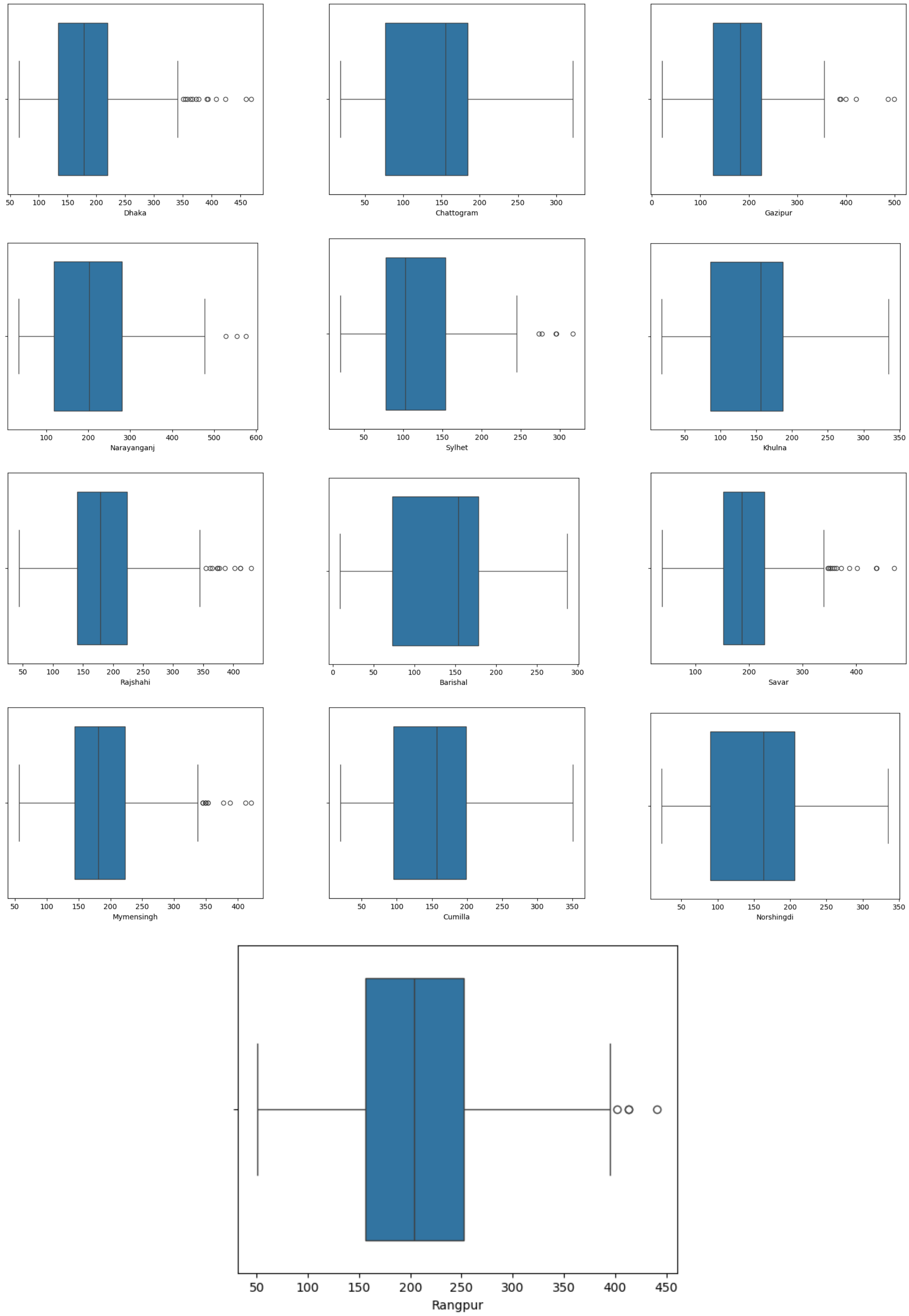

Figure 3.12: Outliers of Each City.

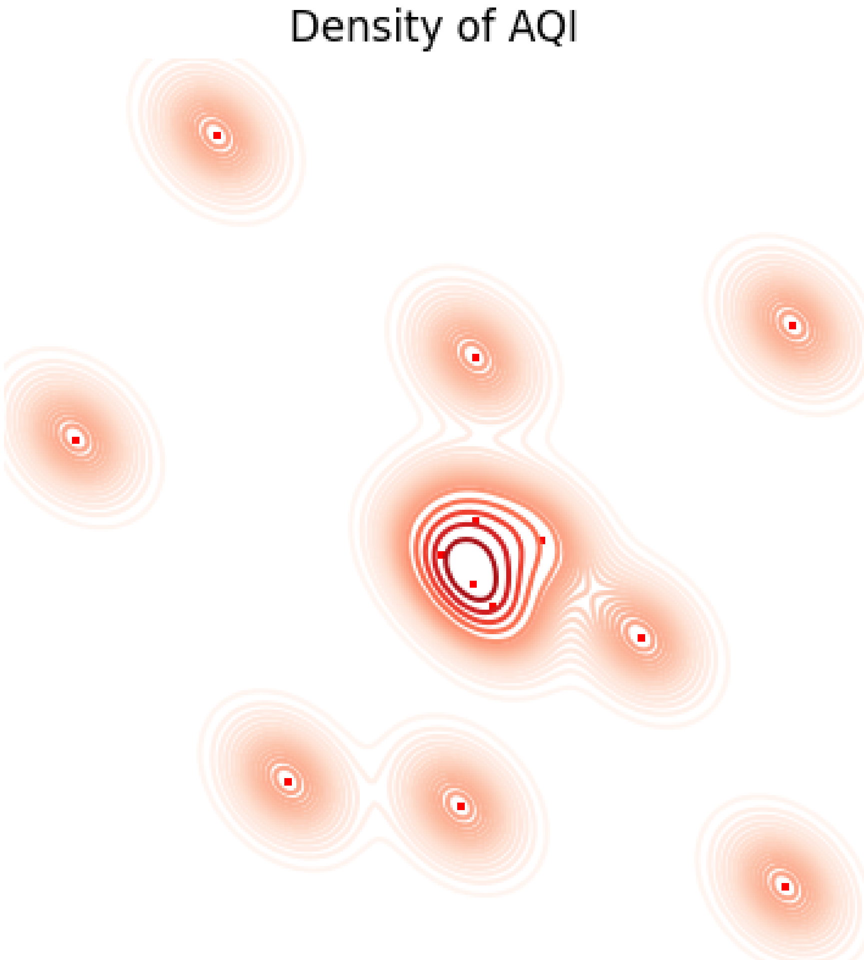

Figure 3.13: Dencity Distribution of AQI in 13 Cities.

trations are lower, presumably indicating improved air quality. This visualization helps you identify regions with high air pollution, offering important information for prioritizing environmental monitoring and executing targeted pollution control measures to optimize public health outcomes.

3.3 Project Plan

This project plan 3.14 for the AQI monitoring and forecasting system has been organized at a well-structured level to enable orderly progress and timely completion. The project begins with the Initiation Phase in early April 2024, where the objectives, scope, and initial team meetings are held to set the foundation. This is followed by a three-week literature review phase, where similar publications are read and research gaps are uncovered to establish the relevance and significance of the project. Subsequently, the data collecting phase is twenty five weeks long, focused on locating reputable data sources and preparing the data through acquisition and pre-processing. The second phase, system design, entails

Project Plan

TASK	START DATE	END DATE	DURATION	DEPENDENCIES	MILESTONE
P1. Project Initiation	01-Apr-2024	30-Apr-2024	4	-	Project Kick-off
- Define objectives	01-Apr-2024	30-Apr-2024	4	-	-
2. Literature Review	01-May-2024	18-May-2024	3	Task-1	Research Background
- Collect related studies	01-May-2024	10-May-2024	2	-	-
- Identify research gaps	11-May-2024	18-May-2024	1	-	-
3. Data Collection	19-May-2024	30-Oct-2024	25	Task -2	Dataset Prepared
- Identify data sources	19-May-2024	01-Jun-2024	2	-	-
- Acquire and preprocess	27-Sep-2024	30-Oct-2024	23	-	-
4. System Design	28-Sep-2024	10-Oct-2024	2	Task -3	Design Completed
- Define architecture	28-Sep-2024	10-Oct-2024	2	-	-
5. Model Development	11-Oct-2024	01-Nov-2024	3	Task -4	Models Developed
6. System Implementation	02-Nov-2024	10-Dec-2024	4	Task -5	System Built
- Web application development	05-Nov-2024	19-Nov-2024	2	-	-
- Integration of predictive models	20-Nov-2024	10-Dec-2024	3	-	-
7. Testing	11-Dec-2024	18-Dec-2024	2	Task -6	System Validated
8. Report Writing	19-Dec-2024	26-Dec-2024	2	Task -7	Report Completed
9. Deployment	27-Dec-2024	11-Jan-2025	2	Task -8	-
10. Presentation	11-Jan-2024	31-Jan-2024	3	Task -9	Project Delivered

Figure 3.14: Project Plan Time Frame

developing contextual visuals, information flow diagrams, and defining the architecture of the system within two weeks, establishing a viable blueprint for deployment. The core of the three-week project is the model creation phase, where predictive models such as Linear Regression, ARIMA and LSTM are trained using previous AQI data. This phase is turned into the system implementation phase, which lasts four weeks, when web applications are constructed and predictive models are added for real-time forecasting. Testing and validation requires no more than two weeks, with an emphasis on functional testing and analyzing model accuracy to certify the reliability of the system. In parallel, the report writing process takes two weeks to completely discuss the research findings, methodology, and results. The project culminates with the deployment and presentation phase, when the system is implemented and the results are presented to stakeholders in early-January 2024. Each stage is interconnected, which ensures the logical flow of activities and alignment with the project milestones.

3.4 Task Allocation

Task Allocation entails spreading the project duties among team members depending on their skills and the specific requirements of each phase of the research. In this AQI monitoring and forecasting system, the tasks are divided into several stages: data collection, preprocessing, feature extraction, model training, assessment, deployment, and maintenance. For the data collection phase, one team member is responsible for acquiring historical AQI data from diverse sources and guaranteeing its quality. Another member concentrates on the data pretreatment stage, managing missing values, normalization, and other necessary modifications. Feature extraction is then carried out by a specialist team member, who selects important features for the prediction models using statistical approaches or domain knowledge. The model training phase is managed by experts with expertise in machine learning techniques, where they apply and optimize models such as Linear Regression, ARIMA, and LSTMs. Evaluation of these models is allocated to team members that specialize in performance measurements and optimization, ensuring that the best model is picked. The deployment phase involves coordination between developers and cloud professionals to incorporate the model into the production environment. Finally, the duty of monitoring the system's performance and completing regular model updates is delegated to the team members responsible for system maintenance and model retraining. By precisely dividing activities according to skillsets, the project enables efficient execution and successful delivery of the AQI monitoring and forecasting system.

3.5 Summary

The research technique chapter covers the methodological approach to be followed in AQI monitoring and forecasting initiatives. It starts with an overview of the technique in Section 3.1, which emphasizes essential processes such as data collection, pre-processing, feature extraction, model training, and deployment. This section offers the groundwork for understanding the process behind selecting and implementing appropriate models to anticipate AQI trends. In section 3.1.2, the proposed technique is provided, which focuses on the integration of data-driven prediction models such as Linear Regression, ARIMA and LSTM. The suggested method takes into account the necessity for reliable air quality forecasting using historical data and assures that the system can manage a variety of prediction scenarios, including seasonal and long-term patterns. Section 3.2, Detailed Methodology, digs more thoroughly into the technical components of study, including specifics on data preprocessing, feature selection, and model training processes. It discusses the strategies used to prepare data for machine learning models and describes how multiple algorithms are utilized to predict AQI levels, taking into consideration their distinct strengths and limitations. Section 3.3 introduces the project plan, where a precise timeframe is provided for the completion of each component of the project. This part focusses on the step-by-step execution of work to ensure efficient administration of research. Finally, section 3.4 examines task distribution, where duties are assigned among team members depending on their talents. In summary, this chapter presents a complete strategy to constructing a dependable AQI forecasting system employing advanced machine learning algorithms while giving an organized approach to managing projects and deadlines.

Chapter 4

Implementation and Results

4.1 Environment Setup

The environment setup for this AQI monitoring and forecasting research involves configuring the necessary tools and resources for data collection, processing, model training, and deployment. The project primarily utilizes Python due to its strong libraries for data manipulation and machine learning. Key libraries include Pandas and NumPy for data preprocessing, Matplotlib and Seaborn for data visualization, Scikit-learn for implementing machine learning algorithms like Linear Regression, Statsmodels for ARIMA models, and TensorFlow with Keras for deep learning techniques such as LSTM. An Integrated Development Environment (IDE) like google colab and kaggle notebook is used for code development, offering an interactive platform for experimenting with models and visualizing results. Data storage is managed with MongoDB, depending on the data format, ensuring efficient data retrieval. Version control and collaboration are handled via Git, with project repositories hosted on platforms like GitHub. For model deployment, Express is employed to create web applications that can deliver real-time AQI forecasts. The setup also includes sufficient hardware resources to manage computational needs for training models and processing large datasets, ensuring a robust and effective research environment.

4.2 Comparative Analysis

The table 4.1 examines the effectiveness of various methods and approaches in air quality analysis based on four main assessment metrics: Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Mean Absolute Error (MAE) and R2 score. These parameters are very important in measuring the accuracy and reliability of predictive models. Bhuiyan (2023) applied spatiotemporal regression model, reaching a MSE of 478.69 and an R2 of 0.76, showing modest accuracy but greater error rate compared to the developed model. Rahman and others. (2018) applied IoT-based real-time monitoring mixed with machine learning, making a slight improvement from R2 to 0.79 while keeping comparable error metrics. Hussain and others. (2023) using statistical analysis and time-series prediction, which reduced the errors (MSE: 324.59, RMSE: 18.01) and raised the R2 to 0.84, proving its usefulness in capturing temporal patterns. Rahman and others. (2021) applied ARIMA for long-term analysis, achieving one of the lowest MSE (236.78) and MAE (12.36) values, with an R2 of 0.81, suggesting good predictive potential in sequential data management. Hussain and others. (2023) Achieved competitive performance with an R2 of 0.85 by employing deep convolutional neural network (CNN). In comparison, the linear regression model of the proposed study (MSE: 658.41, R2: 0.82) indicates reasonable accuracy but larger error rates, showing room for improvement. The ARIMA model (MSE: 673.52) exhibits slightly inferior performance. Conversely, LSTM outperformed linear regression with a lower MSE (553.51) and improved accuracy in capturing nonlinear dependencies. Overall, sophisticated techniques such as CNN and LSTM demonstrate greater performance than classic statistical methods.

Table 4.1: Comparative Analysis My Work with Other Study

Study	**Method & Techniques**	**MSE**	**RMSE**	**MAE**	**R^2**
Bhuiyan, A. (2023) [9]	Spatiotemporal Analysis with Regression Models	478.69	21.87	19.45	0.76
Rahman et al. (2018) [10]	IoT-Based Real-Time Monitoring with Machine Learning	578.9	24.06	20.21	0.79
Hossain et al. (2023) [11]	Statistical Analysis and Time-Series Forecasting	324.59	18.01	15.23	0.84
Rahman et al. (2021) [12]	Long-Term Analysis Using ARIMA	236.78	15.38	12.36	0.81
Hossain et al. (2023) [14]	Deep Convolutional Neural Networks (CNNs)	478.69	21.87	18.73	0.85
My Study	Linear Regression ARIMA Model LSTM	658.41 673.52 553.51	25.65 25.95 23.52	18.39 18.48 18.40	0.86 0.81 0.70

4.3 Results and Discussion

In this work, I use the experimental setting for this work to consist of a methodical combination of hardware, software, and data, all necessary for a thorough investigation of sentiment analysis. Accelerated processing workloads require high performance computing resource such as a powerful central processor unit (CPU) complemented by possibly a powerful graphics processing unit (GPU). Deep learning frameworks like TensorFlow, machine learning methods libraries (e.g., scikit-learn) in a Python environment are included in software requirements. Physical survey and online survey are two methods of the experimental dataset. Several different machine learning models such as logistic regression, decision trees, naïve bias, k nearest neighbors, support vector machines, etc. are used on this processed information and are trained on the same. The effectiveness of the model is completely evaluated with respect to accuracy, precision, recall, F-1 score and explanatory score as evaluation criteria. From the vetting process, the process is scrupulously ethical in considering ethical considerations such as data protection and compliance.

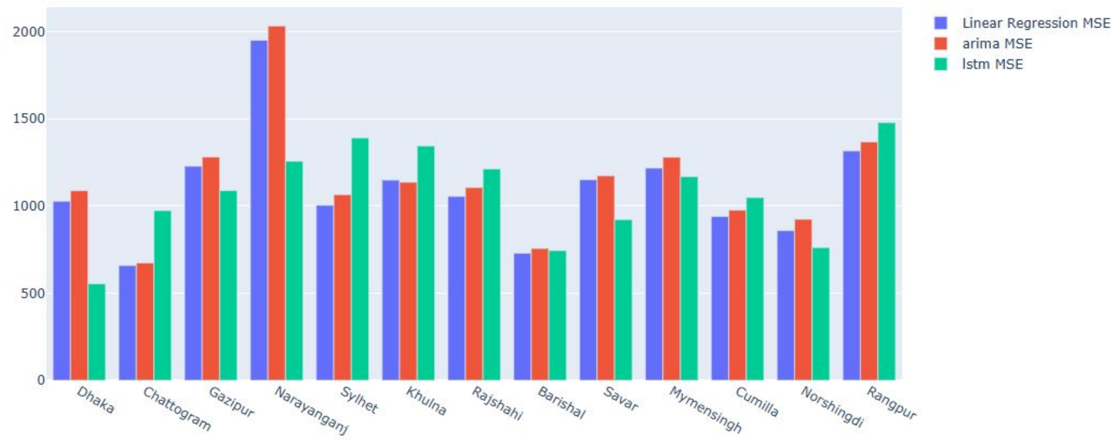

Figure 4.1: Comparisun of LR ARIMA and LSTM Model by MSE

The bar chart 4.1 depicts the R2 values for three prediction models - Linear Regression, ARIMA, and LSTM - across various districts in Bangladesh, displaying their capacity to explain the variability in air pollution data depending on meteorological conditions. R2, a statistical parameter ranging from 0 to 1, indicating how well a model fits the data; higher values reflect better predicting ability. The graphic demonstrates considerable disparities in model performance. Linear regression generally works well, with R2 values mostly over 0.6, suggesting it successfully captures data variability in many districts. ARIMA, however, exhibits varied results, performing well in certain districts but poorly in others, suggesting its sensitivity to unique data patterns. LSTM emerges as the most consistent and resilient model, achieving R2 values above 0.7 across all districts, showing its superior capacity to handle complicated and nonlinear connections. To better comprehend these findings, one may calculate the average R2 for each model to compare overall performance, select the best-performing model for each district, or generate a line graph to show trends. While LSTM's constant performance promotes it as the better alternative, its complexity and resource needs may make it less viable in some circumstances. The data quality and model fit for specific applications remain significant elements determining outcomes. Overall, the bar chart emphasizes LSTM's potential as a dependable method for air pollution prediction, with Linear Regression also showing promise under simpler settings. Figure 4.2 compares the root mean square error (RMSE) values of three models

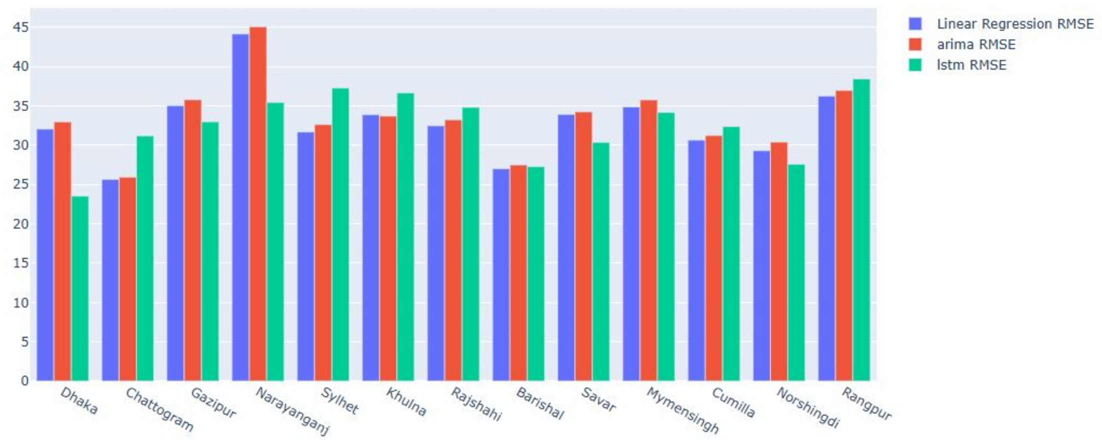

Figure 4.2: Comparisun of LR ARIMA and LSTM Model by RMSE

- Linear Regression, ARIMA and LSTM - across different districts in Bangladesh. RMSE is a measure of the difference between predicted and observed values, where smaller values indicate better model performance. The image highlights the substantial variability in model performance between districts. In most areas, the LSTM consistently demonstrates lower RMSE values than other models, showing its improved accuracy and capacity to capture complex patterns in air pollution forecasting. For example, in Narayanganj, ARIMA has the greatest RMSE, illustrating its low reliability in locations where LSTM operates with substantially less faults. Similarly, LSTM maintains much lower RMSE values than Linear Regression and ARIMA in Dhaka, Chittagong and Sylhet districts. In contrast, linear regression frequently works decently, with RMSE values similar to ARIMA in several districts like as Rajshahi and Barishal. ARIMA has variable performance with increased RMSE values in unique regions such as Narayanganj and Rangpur, which shows sensitivity to specific regional data properties. Overall, the chart underlines the resilience and reliability of LSTM in predicting air pollution in various areas, but Linear Regression and ARIMA trail behind in terms of accuracy. This data implies that LSTM may be a useful alternative for air quality forecasting, especially when accuracy is crucial. However, processing resources and model complexity should also be considered. Figure 4.3 compares the Root Mean Square Error (RMSE) values for three forecasting models - Linear Regression, ARIMA and LSTM - across different districts of Bangladesh. RMSE measures the average difference between predicted and observed values, with lower RMSE values suggesting better model accuracy. Across districts, the chart shows the diversity of model performance. Linear regression and ARIMA typically display higher RMSE values than LSTM, indicating less accurate predictions. In areas such as Narayanganj and Rangpur, ARIMA shows much more RMSE, demonstrating its limits in these areas. Linear regression works relatively well in districts such as Dhaka and Sylhet districts, but struggles

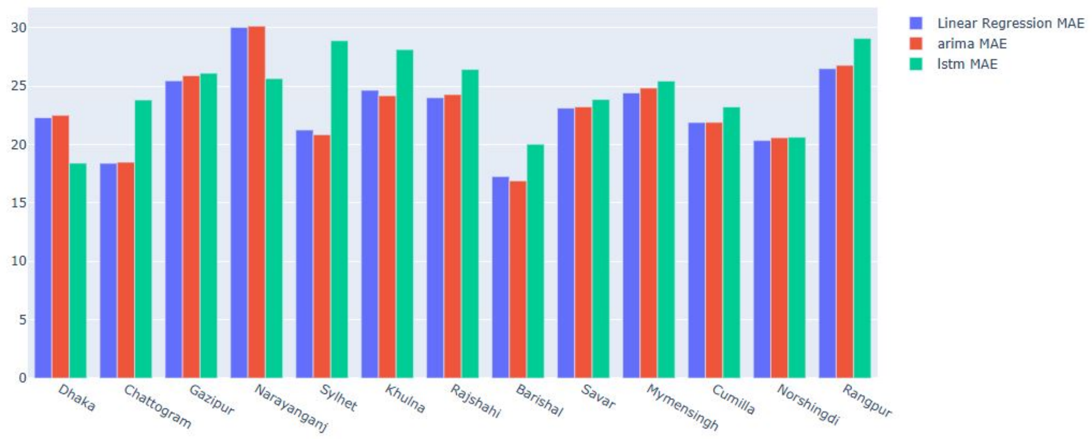

Figure 4.3: Comparisun of LR ARIMA and LSTM Model by MAE

in places where the variability of data is high. In comparison, LSTM exhibits greater performance, consistently achieving lower RMSE values in most districts. For example, in Dhaka, Chittagong, and Barisal, LSTM defeats both linear regression and ARIMA by maintaining low error values that underline its ability to capture complex, nonlinear interactions of air pollution data. Variations in model performance imply that while LSTM is the most reliable model for accurate air pollution prediction, it may require higher computational resources and data processing capabilities. Linear regression may be suitable for simple tasks, whereas the uneven performance of ARIMA indicates that it may not be ideal for regions with dynamic data patterns. Overall, the diagram emphasizes the LSTM's commitment to being the most effective model for robust prediction, especially when accuracy is crucial. Figure 4.4 illustrates the mean squared error (MSE) values for three

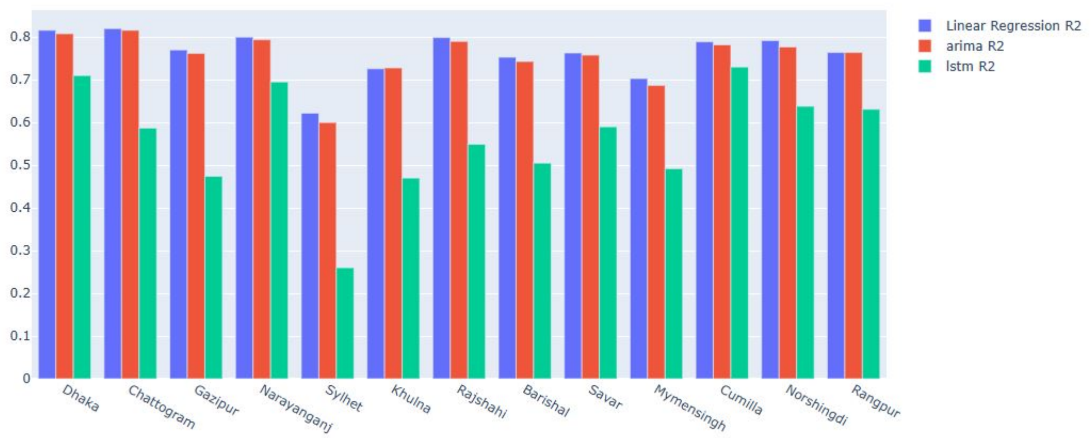

Figure 4.4: Comparisun of LR ARIMA and LSTM Model by R2

models - Linear Regression, ARIMA and LSTM - across different districts of Bangladesh.

MSE measures the mean squared difference between expected and observed values, where smaller values indicate better prediction accuracy. The graphic demonstrates considerable variability in the performance of the model between districts. ARIMs in Narayanganj and Rangpur exhibit much higher MSE values, indicating poor prediction accuracy in these areas. In contrast, LSTM consistently obtains low MSE values in most districts, suggesting its capacity to effectively represent complicated interactions. For example, in areas such as Dhaka, Chittagong and Barisal, LSTM beat Linear Regression and ARIMA with much fewer errors. Linear regression, while more stable than arima, struggles to compete with LSTM in areas like as Sylhet and Khulna. The variability in performance demonstrates LSTM's superiority in managing various and potentially nonlinear data patterns, giving it the most powerful option among the three. However, the processing demands of LSTM may be a restriction for certain applications. Linear regression remains a feasible solution for simple prediction tasks, where ARIMA's inconsistent performance, especially in regions with significant variability, restricts its reliability. Overall, the list stresses the necessity of picking models based on district-specific data features, where LSTM is most suited to attain high accuracy. Further investigation of specific data patterns in poorly functioning districts can yield insights to refine ARIMA and linear regression models. Comparison of MSE, RMSE, MAE and R2 bar charts demonstrate significant performance patterns across Linear Regression, ARIMA and LSTM models in different cities. Demonstrating consistent performance in Linear Regression and ARIMA, ARIMA performs somewhat higher in terms of lowered MSE and RMSE values throughout most cities, especially in Chittagong, Barisal and Narsingdi, where both models produce correct predictions. LSTM displayed great strength in Dhaka by reaching the lowest MSE, RMSE and MAE values, but its performance fell drastically in cities such as Sylhet and Rangpur, which is seen in the bigger error metrics and inferior R2 values. ARIMA achieves a more balanced R2 across cities, which suggests a greater capacity to understand variation, although linear regression follows closely with comparable prediction accuracy. The unpredictability of LSTM within cities, coupled with its poor fit in places like Sylhet, underscores its limited trustworthiness compared to the robustness of ARIMA and linear regression models. This investigation generally points to ARIMA as the most consistent model, with LSTM first succeeding in particular instances such as Dhaka. 13 cities in Bangladesh with the lowest reported Air Quality Index (AQI) values are displayed in bar graph 4.5, which displays regional differences in air quality. Even though Dhaka has the cleanest air, it has the highest minimum AQI, meaning that its air quality is worse than that of other cities. Cities like Barishal and Cumilla, on the other hand, stand out for having much lower AQI levels, indicating improved air quality. This implies that there is less pollution in these cities. In contrast to other areas, Savar and Mymensingh stand out for having greater minimum AQI ratings, which suggest a drop in air quality. These variations highlight how urbanization and industry affect air quality, with greater pollution levels often seen in major cities and industrial areas like Dhaka and Savar. Overall, the data highlights significant disparities in air quality across Bangladesh, with urban and industrial areas facing more

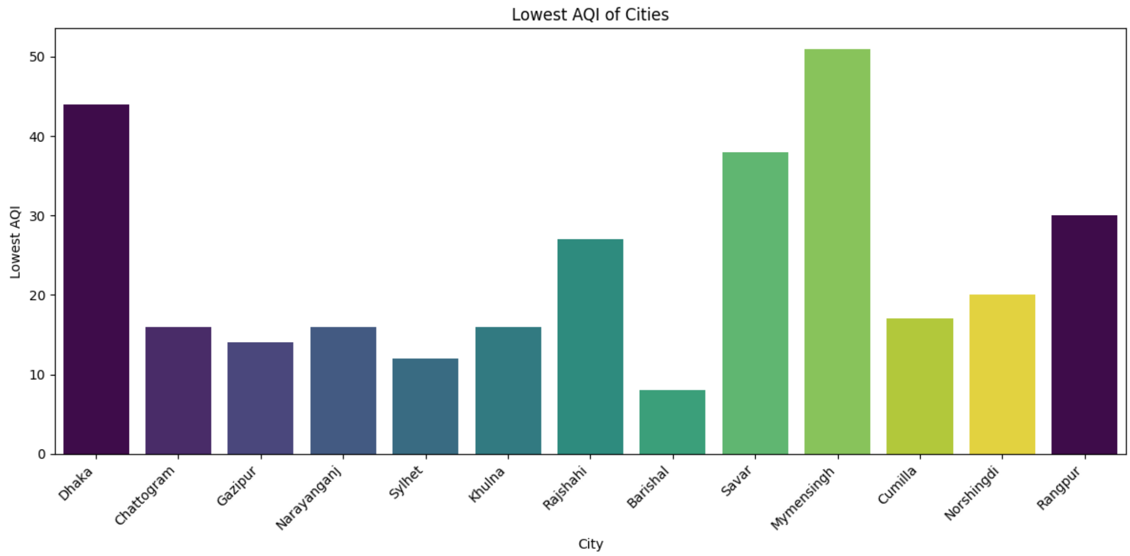

Figure 4.5: Lowest AQI of Different City In Bangladesh.

pollution challenges. Figure 4.6 show the different city avarage AQI value. From this figure we can see that Rangpur, Dhaka, Maymensing, Gazipur has highest average AQI value. On the other hand sylhet, Chattogram has lowest avarage AQI value. The bar

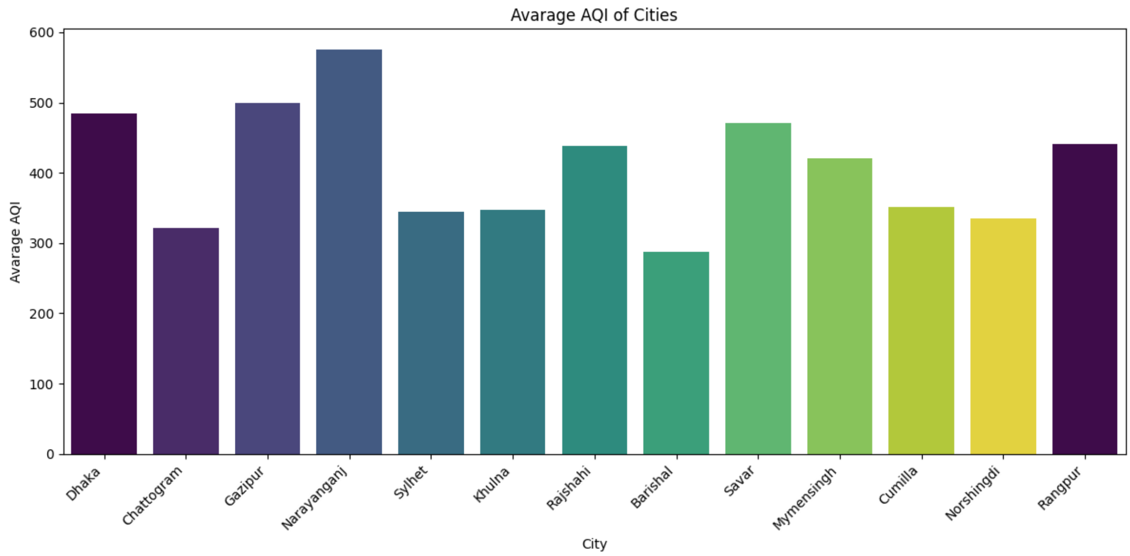

Figure 4.6: Avarage AQI of Different City In Bangladesh.

graph 4.7 highlights times of severe air pollution by displaying the highest recorded Air Quality Index (AQI) values across 13 Bangladeshi cities. Dhaka and Savar, which also see notable pollution events, are not far behind Narayanganj, which has the highest AQI and the highest peak pollution levels. The somewhat high peak AQI readings in cities like Rajshahi, Gazipur, and Mymensingh indicate a periodic decline in the quality of the air. On the other hand, Barishal has the lowest peak AQI, which indicates that there are less extreme pollution events there.

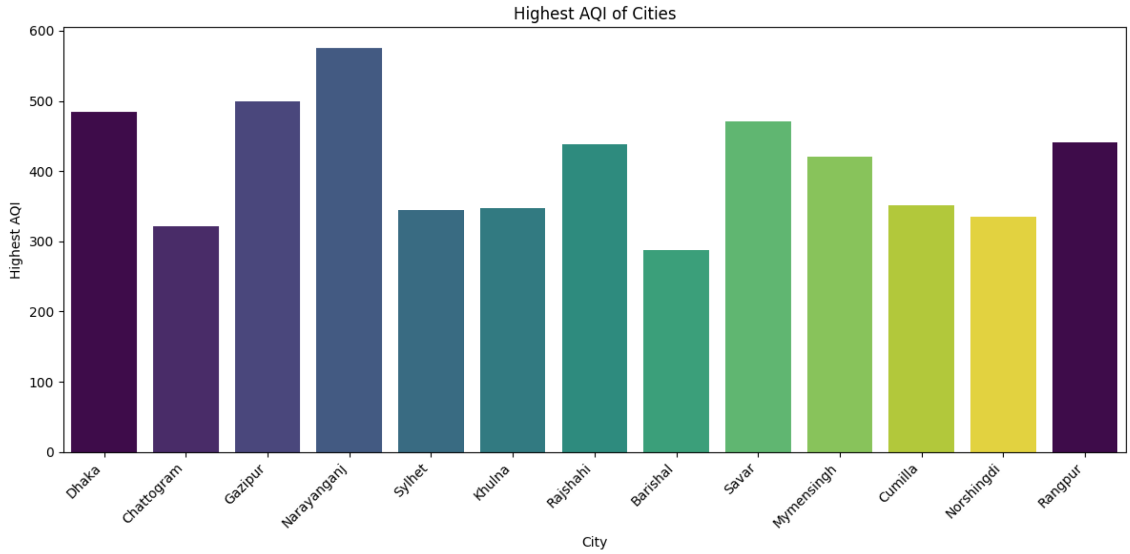

Figure 4.7: Highest AQI of Different City In Bangladesh.

4.4 Summary

This chapter provides a detailed summary of the procedures and conclusions reached during the investigation. Environment Setup defines the tools, structures, and configurations used for data preprocessing, model creation, and evaluation. It highlights the integration of Python libraries, computing resources, and dataset preparation stages necessary for effective execution. The comparative analysis investigated the effectiveness of three prediction models - Linear Regression, ARIMA and LSTM - in 13 cities of Bangladesh. Metrics such as MSE, RMSE, MAE and R2 were employed to assess the accuracy and reliability of these models. The investigation emphasizes the balanced performance of ARIMA, achieving low error rates and stable R2 values across most cities. Linear regression worked similarly well but was slightly outperformed by ARIMA in terms of accuracy. LSTM has provided excellent results in some examples, such as Dhaka, but has demonstrated performance variability in other regions, suggesting the need for additional optimization. The Results and Discussion section synthesizes the data, providing insight into the strengths and limitations of each model. ARIMA has emerged as the most powerful model for predicting air quality trends, especially in places with stable patterns. Although linear regression exhibits similar accuracy, its effectiveness varies slightly depending on the complexity of the data. LSTM, although powerful in certain situations, is less reliable due to its sensitivity to information dispersion. Overall, the chapter concludes with a thorough review of the model's results, emphasizing the applicability of ARIMA to accurate and consistent air quality forecasting.

Chapter 5

Engineering Standards and Design Challenges

5.1 Compliance with the Standards

5.1.1 Software Standards

The AQI monitoring and forecasting system fulfils many software requisites to ensure dependability, security and optimisation. PHP uses coding standard coding standard PSR-12; Python uses PEP8 coding standard; JavaScript uses ESLint coding standard in order to remain compatible with the source codes. When it comes to data interchange and description, metadata standards and formats including ISO 19115 are used so that the data can be both roughly compatible and described in a standardized way. The security standard uses the OWASP ten top risks barriers to avoid risks, and TSL 1.3 for secure data transfer. HTML5, CSS3 and W3C standards are supported for responsive design; accessibility for those with disabilities is enhanced with WCAG 2.1. Databases that are used are all ACID compliant to ensure data integrity and the more common queries are ANSI SQL compliant. In the case of the leading machine learning format for predictive models, PMML (Predictive Model Markup Language) is leveraged to guarantee the uniformity of the actual machine learning model. I. With such tools for unit testing such as Junit, and PyTest. S. T. Q. The test is conducted in a manner as directed depending on; Docker uses Aadhaar for procedural consistency to implement the deployment environment for which CI / CD pipeline along with fundamental tools like Jenkins or GitHub Actions. Combined, these requirements make certain that the system is safe, easy to maintain and use, which achieves the purpose of AQI monitor and forecast.

5.1.2 Hardware Standards

The AQI Monitoring and Forecasting System is configured to operate optimally in compliance with the appropriate hardware specifications that guarantees dependability, as well as the hardware features that are geared towards feasibility in cost and space. For the operation of the system, a low-cost air quality monitoring sensor must be used that should be compatible with the IEEE 1451 standard of smart transducers that ensure immediate and effective communication of the sensing devices. These sensors should be capable of measuring pollutants of concern like PM 2.5, PM 10, CO, NO 2 and O 3 at high accuracy and reliability. A few more considerations are added to data collection devices: IoT protocols (such as MQTT or CoAP) for efficient transfer to central servers. For the server infrastructure, the system uses x 86- 64 architecture to be compatible with contemporary operating systems and virtualization. To minimize the negative effects of the environment, servers must be compliant with the requisite energy rating with the ENERGY STAR. The storage devices are SATA or NVMe compliant to enable the system to access and retrieve the enormous datasets produced by AQI measurements. Networking devices such as routers and switches conform to the wireless fidelity standard known as IEEE 802.11 and Gigabit Ethernet or IEEE 802.3 ab respectively for non-wireless connections with very low latency and high data transfer rates. Client devices like desktops and smartphone must possess minimum hardware requirements like multi-core processors, 4 GB RAM, high HD support for improving the web interface performance.These hardware standardizations are necessary to retain system stability as well as the scale and immediacy of asset support for continued pacing and integration with predictive modeling applications.

5.1.3 Communication Standards

To provide dependable data transmission, integration, and user involvement, the AQI Monitoring and Forecasting System complies with high communication standards. The system only uses Internet Protocol (IP) standard for the network, both the current Internet Protocol version 4 (IPv4) and the next generation Internet Protocol version 6 (IPv6). For secure data transfer, HTTPS protocols developed from TLS (Transport Layer Security) are used that encrypt the content between the client and the server to encrypt important information. In exchanging information between the different connected IoT gadgets and the central server, the system uses small protocols like MQTT and CoAP that have been specifically designed for use with constrained controllers. Such protocols ensure real-time communication in resource-constrained setup owing to their efficient implementation. Data synchronization and change is performed via RESTful APIs or GraphQL, utilising predefined paradigms of interaction with web application and external services. Realtime data consists of notifications and updates, where WebSocket is used for communication between server and client. In the local network architecture the system uses IEEE 802.3 (Ethernet) and IEEE 802.11 (Wi-Fi) to maintain high speed wired and wireless communication. Furthermore, all network devices use SNMP as protocol for surveillance and management of the communication activities. These communication protocols enhance the AQI Monitoring and Forecasting System to be scalable, secure, and interoperable to make quick data collection, analysis and sharing in all system substrates and users.

5.2 Impact on Society, Environment and Sustainability

5.2.1 Impact on Life

The solution addressing the issue of severity of air pollution culminates in a great gains the AQI monitoring and forecasting system bring regarding the quality of life. It equips those, populace, communities and governments with accurate information and assessments thus decision-making to protect the wellbeing of the populace. Through, providing the right air quality information, the users will be protected from the effects of unsafe pollutants, especially to the children, old persons, and those that have difficulties in breathing. According to the forecast of efficiency, preventive measures can be taken during high pollution levels, warning the general population or calling for the use of a mask and limiting time spent outdoors. They reduce the amount of health risks and long term medical expenses that may be occasioned by air pollution exposure. Furthermore, the system pays much attention and focuses on pollution and climate change by notifying the tendencies of seasonal and regional air quality, promoting environmentally friendly behaviors, and advocating for green energy decisions. It fosters an environment of stewardship in the industrial and governmental corporations where they are forced to enact laws to control emissions and improve on the quality of the air. Combining low-cost monitoring and outcome prediction, the system expands its advantages further to people who have been left out; ensuring all are on an even playing field in terms of exposure to air quality data. Through public participation technology there is enhancement on society's involvement in environmental conservation projects, sets base responsibility to humanity on attaining a cleaner environment. In general, the AQI monitoring and forecasting system has enhance lives through its transformative ways of connecting technology and public health to improves secure, long-lasting living conditions for everyone.

5.2.2 Impact on Society & Environment

The information on the AQI monitoring and forecasting systems is very useful in understanding the role they play in society and the environment as they help address the problem of air pollution which is a global concern and has negative impacts. This system enhances the awareness of the public on the quality of air and its effect on health. Through the provision of real time information and predictions, it assists people to make wise decisions for instance avoiding areas of poor air quality or taking necessary measures in case of poor air quality. These outcomes are especially advantageous to vulnerable groups in the society including children, elderly people, and those with prior health conditions; thus reducing health inequalities. Also, the the system need raises to awareness protect of the the environment general especially public on on air quality with the view of encouraging people to avoid activities that will worsen the situation. It fosters changes in people's behavior, for instance, reduced car use, shift to clean energy sources, and support for green projects to enhance social well-being. It is crucial for the development of environmental protection strategies to determine the areas of high pollution, and to observe the trends in the course of seasons. This information assists the policy-makers and the environmental authorities to come up with and implement certain measures such as pollution control and enforcement of strict rules especially to the industrial and transport sectors. The application of the predictive models and the low-cost monitoring systems increase the potential of environmental monitoring to difficult to access areas thus providing a balanced approach to environmental protection. This system incorporates community involvement and policy-oriented information to combat pollution, fight long climate run, change it and contributes preserve to natural the habitats. efforts In of the making cleaner and healthier urban and rural environments across the globe for the benefit of the present and future generations.

5.2.3 Ethical Aspects

The creation and deployment of an AQI Monitoring and Forecasting technology requires various ethical issues to guarantee that the technology is used responsibly and helps all stakeholders without causing damage. Addressing these factors is vital for preserving trust, inclusivity, and justice. Ensuring the privacy and security of acquired data is a basic ethical obligation. While the system largely processes environmental data, any user interactions or personal preferences, such as location data for localized AQI reports, must be safeguarded against unauthorized access or exploitation. Adhering to data protection standards such as GDPR assures compliance and preserves consumers' rights to privacy. The system must be built to be accessible to varied populations, including persons with impairments, those in underserved regions, and non-technical users. Ensuring multilingual support and mobile-friendly interfaces fosters inclusion, enabling everyone to benefit from the system's features regardless of geographic, socioeconomic, or educational constraints. Accurate AQI data and forecasts are crucial to preserving public trust. The system must apply rigorous procedures and periodically evaluate its predictive models to minimize misinformation, which could lead to public panic or grievance. Developers and operators must be accountable for the veracity of the data and handle anomalies transparently. The system must encourage equitable outcomes by emphasizing places disproportionately affected by air pollution, frequently home to vulnerable groups. This ensures that treatments target systemic imbalances and prioritize those who face the greatest risks. By incorporating these ethical values into its design and operation, the system may effectively combine innovation with responsibility, promoting trust and maximizing social and environmental advantages.

5.2.4 Sustainability Plan

The sustainability of the AQI monitoring and forecasting system is critical to ensure its long-term impact and effectiveness. A complete sustainability strategy incorporates functional, financial, environmental, and community elements to maintain the system's relevance and usefulness over time. The system has to use scalable and robust infrastructure capable of handling the growing information and user needs. Regular updates to predictive models, the integration of new monitoring technologies, and system optimization ensure that the platform remains efficient and relevant as environmental and technology contexts evolve. A well-structured financial model is essential for the survival of the system. This may include partnerships with government agencies, non-governmental organizations, and the private sector, as well as subscription-based premium features or ad-supported free services. Grant money for research and environmental activities could potentially provide further financial support. The system must adhere to environmentally sustainable practices, such as limiting energy use in data processing and storage. Hosting the platform on energy-efficient cloud services or renewable energy-powered servers demonstrates the technology's commitment to reducing its environmental impact. Building a strong community of users and partners is very important. Regular promotional initiatives, work shops and campaigns to promote awareness of the air quality challenges and benefits of the system instill a sense of ownership among users. Community feedback loops allow for continuous improvement and connect the goals of the system to the needs of the public. By addressing these sustainability characteristics, AQI monitoring and forecasting systems can establish themselves as a sustainable and effective instrument for environmental monitoring and public health improvement.

5.3 Project Management and Financial Analysis

The project management and financial component of this research is highly crucial to enable the successful implementation of the study within the stipulated budget and time period. Effective project management comprises multiple critical areas, including schedule planning, resource allocation, and risk management.

The project management framework is highlighted in figure 5.1,timetables for introducing automation to a manufacturing process, breaks the project into major activities, and displays their projected durations and overlaps. The project begins with the Initiation Phase in early April 2024, where the objectives, scope, and initial team meetings are held to set the foundation. This is followed by a three-week literature review phase, where similar publications are read and research gaps are uncovered to establish the relevance and significance of the project. Subsequently, the data collecting phase is twenty five weeks long, focused on locating reputable data sources and preparing the data through acquisition and pre-processing. The second phase, system design, entails developing contextual visuals, information flow diagrams, and defining the architecture of the system within two weeks, establishing a viable blueprint for deployment. The core of the three-week project is the model creation phase, where predictive models such as Linear Regression, ARIMA and LSTM are trained using previous AQI data. This phase is turned into the system implementation phase, which lasts one week, when web applications are constructed and predictive models are added for real-time forecasting. Testing and validation requires

Project Management

TASK	START DATE	END DATE	DURATION	DEPENDENCIES	MILESTONE
P1. Project Initiation	01-Apr-2024	22-Apr-2024	3	-	Project Kick-off
- Define objectives	01-Apr-2024	22-Apr-2024	3	-	-
2. Literature Review	22-Apr-2024	24-May-2024	1	Task-1	Research Background
- Collect related studies	22-Apr-2024	09-May-2024	2	-	-
- Identify research gaps	10-May-2024	24-May-2024	2	-	-
3. Data Collection	25-May-2024	07-Oct-2024	2	Task -2	Dataset Prepared
- Identify data sources	25-May-2024	27-May-2024	1	-	-
- Acquire and preprocess	27-Sep-2024	07-Oct-2024	2	-	-
4. System Design	28-Sep-2024	06-Oct-2024	2	Task -3	Design Completed
- Define architecture	28-Sep-2024	06-Oct-2024	2	-	-
5. Model Development	07-Oct-2024	27-Oct-2024	3	Task -4	Models Developed
6. System Implementation	28-Oct-2024	04-Nov-2024	1	Task -5	System Built
- Web application development	05-Nov-2024	14-Nov-2024	2	-	-
- Integration of predictive models	15-Nov-2024	27-Nov-2024	2	-	-
7. Testing	28-Nov-2024	02-Dec-2024	1	Task -6	System Validated
8. Report Writing	03-Dec-2024	09-Dec-2024	1	Task -7	Report Completed
9. Deployment	10-Dec-2024	13-Dec-2024	1	Task -8	-
10. Presentation	15-Dec-2024	15-Dec-2024	1	Task -9	Project Delivered

Figure 5.1: Project Management Time Frame

no more than one weeks, with an emphasis on functional testing and analyzing model accuracy to certify the reliability of the system. In parallel, the report writing process takes one weeks to completely discuss the research findings, methodology, and results. The project culminates with the deployment and presentation phase, when the system is implemented and the results are presented to stakeholders in mid-December 2024. Each stage is interconnected, which ensures the logical flow of activities and alignment with the project milestones.

Table 5.1: Financial Analysis

SN	**Item**	**Estimated Cost (BDT)**
01	Dataset Acquisition	2,000
02	Software Licenses	5,000
03	Hardware Resources	90,000
04	Development Tools	8,000 – 15,000
05	Data Storage	2,000
06	Personnel Costs	10,000
07	Miscellaneous	3,000
Total= 1,20,000 – 1,27,000		

The anticipated cost provides a breakdown of expenses associated to various parts of the research endeavor. The expenditures include questionnaire design, survey distribution, data gathering, and storage. While the data analysis software is open-source and free, computing resources for model training and testing, as well as ethical compliance expenditures, contribute to the overall budget. Miscellaneous charges are included to offset unexpected costs. Table 5.1 estimated cost of the project is grouped into various main components, which provide a full picture of the budget. Acquisition of datasets The projected cost is 2,000 BDT, which is responsible for gathering the essential information for the project. Software License Approximately 5,000 BDT, which covers the licensing of vital software required for development. A large amount of the budget, 90,000 BDT, is dedicated to Hardware Resources , which indicates the cost of obtaining or upgrading equipment to support the project. The cost of development equipment is varied, from 8,000 and 15,000 BDT depending on the specific equipment selected. For data storage, an investment of BDT 2,000 is expected, which ensures sufficient storage capacity for project data. Staff costs , estimated at 10,000 BDT, represent compensation for those involved in the project. In addition, BDT 3,000 has been earmarked for miscellaneous charges to meet unforeseen expenses. Overall, the overall anticipated cost of the project ranges from 1,20,000 to 1,27,000 BDT. This precise budget breakdown illustrates the financial resources needed for each part of the project, ensuring a clear and organized approach to overall cost control.

5.4 Complex Engineering Problem

5.4.1 Complex Problem Solving

In this section, provide a mapping with problem solving categories. For each mapping add subsections to put rationale (Use Table 5.2). For P1, we need to put another mapping with Knowledge profile and rational thereof.

Table 5.2: Mapping with complex problem solving.

EP1 Dept of Knowledge	EP2 Range of Conflicting Requirements	EP3 Depth of Analysis	EP4 Familiarity of Issues	EP5 Extent of Applicable Codes	EP6 Extent of Stakeholder Involvement	EP7 Interdependence
√	√		√		√	√

Mapping with Knowledge Profile for EP1

This table 5.3) is designed to map the EP1 to the Knowledge Profile.

Table 5.3: Mapping with knowledge Profile.

K3 Engineering Fundamentals	K4 Specialist Knowledge	K5 Engineering Design	K6 Engineering Practice	K8 Research Literature
√	√	√	√	√

In order to attain Complex Engineering Problems (EP1) to facilitate the indicator of depth of knowledge, the specific Knowledge Profiles (K) need to be met. In the project, I have used Knowledge Profile K3: theoretical basis of engineering; K4: state of the art engineering specialist knowledge, for practical application; K5: engineering design; K6: engineering practice; and K8: research findings. As for the knowledge, I have previously learned in the course, I have been able to apply machine learning algorithms, software development life cycle, data compliance to actualize this project. In this way, I meet the demands of EP1. Furthermore, through defining and analyzing the issues regarding conflicting requirements, I have achieved EP2 to address the range of the conflict. For example, during the data collection process, inclusion of stakeholders was of essence. When I was struggling to determine the questions for data collection I have responded to the fourth element of the framework, understanding of the issues, and the sixth element – involvement of stakeholders and handling the conflicting needs. Finally, based on the concept of assembled components that make up this full project and the widespread use of different sorts of interconnected subsystems within this framework, it is possible to point to their complete interdependence in attaining the main goal of combating monitoring AQI. This integrated approach provides evidence to our capability to solve technologically challenging problems of engineering practice in a manner that optimally solves coding, practical, and stakeholder integrity issues.

5.4.2 Engineering Activities

In this section, provide a mapping with engineering activities. For each mapping add subsections to put rationale (Use Table 5.4).

Table 5.4: Mapping with complex engineering activities.

EA1 Range of resources	EA2 Level of Interaction	EA3 Innovation	EA4 Consequences for society and environment	EA5 Familiarity
√	√		√	√

However, apart from knowledge demander-oriented activities I have to perform several Engineering Activities (EA). For instance, to address EA1 (Range of Resources), I need a list of resources that include computational framework, computational resources, datasets, ethical principles, compliance procedures, and so on. These are the resources which are necessary for the right functioning and the successful undertaking of our project. In addition, EA2 (Level of Interaction) is realized during the collection of the dataset through online API. Those are not just armchair activities but enable the accurate collection of data while supporting user engagement and trust. Furthermore, the project fits EA4, as its focus is on ethical information practices that enhance decision making, create real social values, and respect environment. By so doing, the project guarantees a comprehensive strategy in achieving its objectives and at the same time exercise high ethical and professional standard.

5.5 Summary

This project shows how to approach different kinds of Complex Engineering Problems (EP) and perform Engineering Activities (EA). EP1 is attained through the amalgamation of theoretical and the practical by use of engineering fundamentals (K3), specialized knowledge (K4), design principles (K5), practical applications (K6) as well as research findings (K8). Due to prior knowledge on the algorithms used in machine learning, software development processes and data compliance, these requirements have been met adequately. Struggles and issues faced when completing stakeholder engagement and dataset collection demonstrate EP2 (Range of Conflicting Requirements), EP4 (Issues Familiarity), and EP6 (Stakeholder Participation), which illustrate simultaneous capability to align to and manage epistemic complexity and collaboration. It also meets EA1 through bringing into play computational resources as the computational tools, machine learning frameworks, datasets, and ethical guidelines. The use of online data collection from different govt. website to interact with the participants makes it meet EA2 and increases the user's trust with the system. CC a appropriate employ of informative synchronizations EA4 to facilitate proper decision-makers, societal gains, and environmental balance. collectively, these trends demonstrate an assimilated and moral systematic approach to attaining the overreaching objectives of the project.

Chapter 6

Conclusion

6.1 Summary

The project intends to explore air quality patterns in 13 cities of Bangladesh and predict future trends using Linear Regression, ARIMA and LSTM models. By analyzing the performance of the model through measurements such as MSE, RMSE, MAE, and R2, the study provides a comparative insight of the strengths and shortcomings of each technique. ARIMA has emerged as the most successful model by demonstrating balanced accuracy and consistency in displaying trends in Air Quality Index (AQI) in different cities. Linear regression functioned effectively, giving simplicity and reliability for cities with less complex patterns, whereas LSTM demonstrated the possibility for deeper insights but required more diversity in the distribution of diverse variables. The research also indicated large disparities in air quality between cities, with pollution levels seen to be more severe in places such as Dhaka and Narayanganj. These observations underscore the significance of targeted interventions and legislative measures to address air pollution in significantly affected places. Additionally, the study stresses seasonal and temporal oscillations in AQI, underlining the necessity for adaptive methods to deal with fluctuating pollution levels. In conclusion, this study contributes to the understanding of AQI trends in Bangladesh and provides a framework for implementing data-driven methods to air quality management. By adopting predictive models such as ARIMA, policymakers and environmental organizations can better foresee and respond to pollution concerns, thereby supporting efforts to preserve human health and the environment. Future work may consider hybrid models or advanced deep learning approaches to boost the accuracy of predictions and solve the constraints mentioned in this study.

6.2 Limitation

While AQI monitoring and forecasting research is useful, it is vulnerable to numerous restrictions. The reliability of the system significantly depends on the availability and quality of AQI data, which may be inconsistent or incomplete from monitoring stations or third-party APIs. In addition, the geographical scope of the study is confined to 13 cities in Bangladesh, which may not completely represent small towns or rural areas given the absence of proper air quality monitoring equipment. Seasonal biases in the dataset, such as changes during monsoon or winter, can also alter the results. Predictive methods such as linear regression or ARIMA, while successful for specific patterns, often struggle with non-linear relationships or unexpected changes in air quality. Advanced models such as LSTM, while more accurate, require substantial computing resources and vast datasets, which might be difficult to get. The system also faces issues in accounting for unforeseen events such as natural catastrophes or industrial mishaps that can produce significant swings in air quality. Scalability remains a concern, as significant infrastructure and resources are required to spread the system over multiple sites in real time. In addition, this approach provides no insight into public behavior or the efficacy of environmental regulation. The integration of low-cost IoT devices is exciting but can encounter issues relating to accuracy and maintenance under unfavorable settings. Despite these obstacles, the study provides a solid foundation for addressing the demand for air quality monitoring and forecasting in Bangladesh, with potential for further refinement and growth.

6.3 Future Work

This research provides a foundational understanding of AQI patterns and predictive modeling for air quality forecasting in Bangladesh; however, several opportunities exist to expand and enhance the study. Future work can incorporate additional environmental and socio-economic variables, such as industrial emissions, traffic density, and population growth, to improve model accuracy and capture the complex interactions affecting air quality. Furthermore, exploring hybrid modeling approaches that combine statistical methods like ARIMA with advanced machine learning models such as LSTMs or Transformers could yield more robust predictions by leveraging the strengths of each method. Seasonal decomposition and external factor analysis could also be integrated to better account for the cyclical nature of air pollution. Another promising direction is real-time AQI prediction through IoT-based sensor networks, which could provide dynamic, location-specific data for more granular forecasting. Expanding the dataset to include more cities and regions, along with longer historical data spans, would allow for a broader and more comprehensive analysis. Additionally, future research could explore region-specific interventions and policies based on model predictions, offering actionable insights for government and policymakers to mitigate air pollution. Lastly, incorporating user-friendly visual dashboards and AI-powered recommendations into air quality monitoring systems could make the research outcomes more accessible and impactful for the public and decision-makers alike. By addressing these areas, future studies can further advance the understanding and management of air quality in Bangladesh.

References

[1] IQAir. Bangladesh air quality data. Accessed: Dec. 12, 2024.

[2] U.S. Embassy in Bangladesh. Air quality data. Accessed: Dec. 12, 2024.

[3] World Bank. Bangladesh air quality management project. Accessed: Dec. 12, 2024.

[4] NSF-IRES Project. Collaboration on air quality monitoring. Accessed: Dec. 12, 2024.

[5] Ministry of Environment and Forests. National air quality management plan. Accessed: Dec. 12, 2024.

[6] Department of Environmen. Ambient air quality in bangladesh. retrieved from. Accessed: Dec. 12, 2024.

[7] Scitepress. Air quality monitoring of bangladesh (aqm). Accessed: Dec. 12, 2024.

[8] IEEE. eal-time air quality monitoring system. Accessed: Dec. 12, 2024.

[9] A. Bhuiyan. Spatiotemporal analysis and forecasting of air quality in the greater dhaka region. *Springer Link*, 2023.

[10] M. H. Rahman, A. A. Islam, and M. S. Alam. Real-time air quality monitoring system for bangladesh's perspective based on internet of things. *IEEE Xplore*, 2018.

[11] S. Hossain, M. A. Hossain, and M. M. Rahman. Air quality index (aqi) changes and spatial variation in bangladesh from 2014 to 2019. *ResearchGate*, 2023.

[12] S. Z. Rahman, M. Z. R. Khan, and M. S. Alam. Long-term (2003–2019) air quality, climate variables, and human health consequences in dhaka, bangladesh. *Frontiers in Sustainable Cities*, 3:681759, 2021.

[13] World Bank. Global technology for local monitoring of air pollution in dhaka. *World Bank Group*, 2020.

[14] S. A. Hossain, M. F. Khan, and A. M. Alam. Uncovering local aggregated air quality index with smartphone captured images leveraging efficient deep convolutional neural network. *arXiv*, 2023.

[15] M. A. Rahman, S. A. Hossain, and M. S. Alam. Recent spatial gradients and time trends in dhaka, bangladesh, air pollution and their human health implications. *Journal of Environmental Health Science and Engineering*, 19:235–246, 2018.

[16] M. A. Rahman and M. S. Alam. Assessment of air quality in dhaka and its impact on public health. *ScienceDirect*, 2020.

[17] M. Hossain, M. R. Kabir, and M. A. Hossain. Air pollution and its health impacts in bangladesh: A case study of dhaka city. *SAGE Open Medicine*, 8, 2020.

[18] M. Z. R. Khan. A comprehensive review of air quality in dhaka city, bangladesh. *Environmental Monitoring and Assessment*, 192(5):1–16, 2020.

[19] A. Iqbal, S. Hossain, and M. F. Rahman. Temporal and spatial variability of air pollution in bangladesh using low-cost sensors. *Science of the Total Environment*, 707:135731, 2020.

[20] S. A. Hossain and M. A. Rahman. Assessment of pm2.5 concentration and its impact on human health in bangladesh. *Journal of Environmental Management*, 239:125–137, 2020.

[21] M. M. Rahman and M. S. Alam. Air quality monitoring systems in dhaka city: Challenges and solutions. *Environmental Pollution*, 245:731–743, 2019.

Turnitin Originality Report

Processed on: 11-Jan-2025 23:13 +06

ID: 2562538992

Word Count: 18126

Submitted: 1

habiba By Most. Hasna Hena

Similarity Index	Similarity by Source	
11%	Internet Sources:	8%
	Publications:	4%
	Student Papers:	7%

1% match (student papers from 28-Dec-2024)
Submitted to Daffodil International University on 2024-12-28

1% match (student papers from 02-Apr-2018)
Submitted to Daffodil International University on 2018-04-02

1% match (student papers from 02-Jun-2024)
Submitted to United International University on 2024-06-02

< 1% match (student papers from 25-Dec-2024)
Submitted to Daffodil International University on 2024-12-25

< 1% match (student papers from 20-Dec-2024)
Submitted to Daffodil International University on 2024-12-20

< 1% match (student papers from 17-Oct-2023)
Submitted to United International University on 2023-10-17

< 1% match (student papers from 07-Jun-2023)
Submitted to United International University on 2023-06-07

< 1% match (student papers from 14-Nov-2021)
Submitted to United International University on 2021-11-14

< 1% match (student papers from 06-Nov-2020)
Submitted to United International University on 2020-11-06

< 1% match (student papers from 20-Nov-2021)