<div align="center">

Software Requirements Specification
Impact of Air Quality on Vulnerable Groups

</div>

Aim:

This project aims to study the effects of varying Air Quality Index (AQI) levels on specific vulnerable groups, such as individuals with respiratory problems, heart conditions, older adults, and children.

Requirements:

To prepare the Software Requirements Specification document, we need a good understanding of the project and the necessary editor tools like Overleaf for its preparation.

Contents

1 Introduction

Air pollution, characterized by pollutants such as Carbon Monoxide (CO), Nitrogen Dioxide (NO2), Ozone (O3), and Fine Particulate Matter (PM2.5), poses significant health risks to populations worldwide. The Air Quality Index (AQI) serves as a crucial tool for assessing outdoor air quality, providing a comprehensive evaluation of pollutant levels. This project focuses on classifying diverse demographic groups based on their susceptibility to AQI values and predicting the associated health impacts. Specifically targeting respiratory conditions, cardiovascular diseases, and other health risks, the study aims to identify those most vulnerable to air pollution at specific AQI thresholds. By gaining insights into these vulnerabilities, the research intends to inform targeted strategies and policies aimed at safeguarding public health from the detrimental effects of environmental pollution.

1.1 Purpose

The purpose of this project is to analyze how different levels of air pollution, measured by the Air Quality Index (AQI), affect various demographic groups. By focusing on vulnerable populations such as those with respiratory conditions and cardiovascular diseases, the study aims to identify who is most at risk from different AQI thresholds. Ultimately, the research seeks to inform strategies and policies that can effectively protect public health from the harmful effects of environmental pollution.

1.2 Intended Audience

This document is intended for researchers in environmental health, public health officials, policymakers involved in air quality regulation, and stakeholders concerned with the impact of air pollution on vulnerable populations. Additionally, it targets software developers and data analysts who will implement and utilize the systems and tools required for data collection, analysis, and reporting within the project.

1.3 Product Scope

The scope of this project includes

- Identifying vulnerable groups and their susceptibility to different AQI levels.

- Analyzing the health impacts of pollutants like Carbon Monoxide (CO), Nitrogen Dioxide (NO2), Ozone (O3), and Fine Particulate Matter (PM2.5) on these groups.

- Developing recommendations for public health interventions and policies based on the findings.

1.4 Definitions, Acronyms, and Abbreviations

- **AQI**: Air Quality Index

- **CO**: Carbon Monoxide

- **NO2**: Nitrogen Dioxide

- **O3**: Ozone

2   Overall Description

2.1   Product Perspective

The project focuses on conducting research and analysis rather than developing a physical product. It involves gathering and analyzing data to generate insights into the relationship between air pollution levels and health outcomes for vulnerable populations.

2.2   Product Functions

The primary functions of this project as follows:

- **Data Collection**: Gathering real-time or historical data on AQI levels and pollutant concentrations from reliable sources.Ensure that the dataset includes information on AQI levels categorized into ranges (e.g., Good, Moderate, Unhealthy, etc.).

- **Data Processing and Analysis**:

    - Pre-process the data by removing irrelevant columns like "Good" and "Bad" indicators.
    - Add a new column that categorizes AQI ranges into specific health impact categories.
    - Implement algorithms or statistical methods to correlate AQI levels and pollutant concentrations with health outcomes specific to vulnerable groups.

- **Reporting**: Generating comprehensive reports and visualizations summarizing the findings of the research, including recommendations for public health interventions and policies based on the observed data.

2.3   User Characteristics

The primary users of the project include researchers, public health officials, policymakers, and stakeholders involved in environmental health and public policy. These users will utilize the research findings to make informed decisions regarding air quality management and protection of vulnerable populations.

2.4   Constraints

- Availability of accurate and reliable data on AQI levels and pollutant concentrations.

- Compliance with ethical standards and regulations governing research involving human subjects is essential.

2.5   Assumptions and Dependencies

- **Assumption**: The availability of sufficient data sources for AQI and pollutant levels.

- **Dependency**: Cooperation and collaboration with relevant institutions and organizations for data sharing and research collaboration.

3 Requirement Specification

3.1 Functional Requirements

3.1.1 Data Collection

:

- The system shall collect real-time or historical data on AQI levels from reliable sources.

- The system shall gather data on pollutant concentrations (CO, NO2, O3, PM2.5) corresponding to AQI levels.

3.1.2 Data Analysis

:

- The system shall analyze the collected data to identify correlations between AQI levels, pollutant concentrations, and health outcomes for vulnerable groups.

- The system shall generate statistical reports and visualizations to present the findings of the analysis.

3.1.3 Reporting

- The system shall produce comprehensive reports summarizing the impact of air quality on vulnerable populations.

- Reports shall include recommendations for public health interventions and policies based on the research findings.

3.2 Non-functional Requirements

- **Performance:** The system shall be capable of processing and analyzing large datasets efficiently.

- **Reliability:** The system shall operate reliably with minimal downtime during data collection and analysis phases.

- **Security:** Data transmission and storage shall adhere to best practices for security and privacy protection.

- **Usability:** The user interface shall be intuitive and user-friendly for researchers and stakeholders.

4 Hardware and Software Requirements

4.1 Hardware Requirements

- Processor: AMD® Ryzen 3 3250u with radeon graphics

- RAM: 5.7GB

- Storage: Minimum 500MB

- Display: Screen resolution of at least 1366x768(16:9)2

4.2 Software Requirements

4.2.1 Operating Systems

- Linux (Ubuntu 20.04 version)

4.2.2 Programming Language

- Python: Version 3.7 or higher

4.2.3 Dependencies

- Numpy: For numerical operations and array handling

- pandas: Data Manipulation

- sklearn: Used for Implementing Classification,Regression,clustering and Dimensional Reduction

4.2.4 Development Environment

- IDE: Jupyter Notebook

4.2.5 Data visualization

- Matplotlib

- seaborn

4.2.6 Additional Software Requirements

- Ensure all required Python libraries are installed and updated to their latest versions.

5 Testing Requirements

5.1 Unit Testing

- Data Cleaning

- Data Transformation

- Feature Creation

5.2 Integration Testing

- end-to-end data flow

- Datbase integration

5.3 Machine Learning Model Integration

- Data input for models

- Model predictions

5.4 Model Testing

- **Regression Metrics:**

  - Mean Absolute Error
  - Root Mean Squared Error

- **Classification Metrics:**

  - Accuracy
  - Precision
  - Recall
  - F1-Score

5.5 Model Validation Techniques

- Train-Test Split

5.6 Performance Testing

- **Load Testing:**

  - Data processing load

6   Conclusion

Our project used machine learning to study how air quality affects vulnerable groups like the elderly and children. By analyzing data from global air pollution records, we found that higher air pollution levels correlate with increased health problems in these groups. Our findings underscore the importance of addressing air quality issues to protect vulnerable populations, informing better public health policies and interventions. Ongoing data collection and model updates will be crucial for improving our understanding and response to these health challenges in the future.

References

[1] Sheng He,Mohammad Ashraf Ottom, Fayha Alshibliand Mohammed M. Abu Shquier *Khalid M.O.Nahar: AIR QUALITY INDEX USING MACHINE LEARNING* ,Department of Computer Sciences, Yarmouk University, Irbid, Jordankhalids@yu.edu.jo,Department of Information Systems, Yarmouk University, Irbid, Jordanottom.ma@yu.edu.jo,Department of Land-Water and Environment, Jordan University, f.shibli@ju.edu.jo,Department of Computer Sciences, Jerash University, shquier@jpu.edu.jo

[2] Khalid Nazim Abdul Sattar,Saudi Arabia,Siddesh G M *Sridhar,Latha: AIR QUALITY INDEX USING MACHINE LEARNING* In:Department of Electronics and Communication Engineering Nitte Meenakshi Institute of TechnologyBengaluru, India In: Dept. of ECE Nitte Meenakshi Institute ofTechnology, Bengaluru, India In:Department of Electronics and Communication Engineering GSSS Institute of Engineering and Technology for Women,Mysuru, India In:Department of Department of CSI College of Science, Majmaah University Al Majmaah 11952, Saudi Arabia. In:Department of Electronics and Communication Engineering GSSS Institute of Engineering and Technology for Women,Mysuru, India

[3] `https://www.kaggle.com/datasets/hasibalmuzdadid/global-air-pollution-dataset`