High Level Design (HLD)

Air Quality Index Prediction

Revision Number: 1.0
Last date of Revision: 02/09/2021

Milind Sai
Kartik Bhargav
Nooruddin Shaikh
Saurabh Jejurkar

Document Version Control

Date Issued	Version	Description	Author
15-08-2021	1.0	Introduction and General Description	Nooruddin Shaikh
25-08-2021	1.0	Design Details	Saurabh Jejurkar
31-08-2021	1.0	Performance	Milind Sai
2-09-2021	1.0	Conclusion and Reference	Kartik Bhargav

Contents

Document Version Control ……2
Abstract……4
1 Introduction…… 5
1.1 Why this High-level Design Document?……5
1.2 Scope……5
1.3 Definations……5
2 General Description……6
2.1 Product Perspective…… 6
2.2 Problem Statement……,..6
2.3 Proposed Methodology……6
2.4 Further Improvements……6
2.5 Data Required……6
2.6 Tools used……...6
2.7 Constraints……..7
2.8 Assumptions……7
3 Design……7
3.1 Process Workflow……7
3.1.1 Model Training and Evaluation……8
3.1.2 Deployment Process……9
3.2 Error Handling……..9
4 Performance……9
4.1 Reusability……9
4.2 Application of compatibility……9
4.3 Resaurce Utilization……9
4.4 Deployment……..10
5 Conclusion……..10
6 Reference……10

Abstract

Pollution is major problem in 21st century. There is a lot of work going on to reduce the pollution. Air pollution is important issue in all over the world. In India it become worst in winter season in cities like Delhi and Gurugram. Air quality for a human being and all the living species is a vital part of life. Variation of air quality creates a high impact on human life; continuous monitoring and evaluation and investigation of AIR quality are critical. By using various factors air quality is calculated i.e. Air quality Index (AQI) mathematically. But it's also possible to calculate AQI by using machine learning technique.

1 Introduction

1.1 Why this High-Level Design Document?

The purpose of this High-Level Design (HLD) Document is to add the necessary detail to the current project description to represent a suitable model for coding. This document is also intended to help detect contradictions prior to coding, and can be used as a reference manual for how the modules interact at a high level.

The HLD will:

- Present all of the design aspects and define them in detail
- Describe the user interface being implemented
- Describe the hardware and software interfaces
- Describe the performance requirements
- Include design features and the architecture of the project
- List and describe the non-functional attributes like:
    - Security
    - Reliability
    - Maintainability
    - Portability
    - Reusability
    - Application compatibility
    - Resource utilization
    - Serviceability

1.2 Scope

The HLD documentation presents the structure of the system, such as the database architecture, application architecture (layers), application flow (Navigation), and technology architecture. The HLD uses non-technical to mildly-technical terms which should be understandable to the administrators of the system.

1.3 Definitions

Term	*Description*
AQI	Air Quality Index
PM2.5	Particulate Matter of size 2.5
NO2	Nitrogen Dioxide
CO	Carbon Mono oxide
SO2	Sulphur Dioxide
O3	Ozone

2 General Description

2.1 Product Perspective

Air quality prediction webapp is machine learning based regression model which helps to determine air quality index of area. It can help to instruct the people about air quality so that they take precaution.

2.2 Problem Statement

To create the machine learning based solution to predict air quality index based on the parameters.

2.3 Proposed Solution

Develop the web application to predict the air quality index and quality of air, which help to alert the citizen in particular area.

2.4 Further Improvements

This application can be used with IOT devices.

2.5 Data Required

Data is completely depending upon our problem statement.

- For training the model we need the data that consist of PM2.5, NO2, CO, SO2, O3 and AQI value.

2.6 Tools used

Python programming language and frameworks such as Numpy, Pandas, Scikit-learn, Matplotlib, Seaborn are used to build the whole model

- PyCharm and Visual Studio Code is used as IDE.
- For visualization of the plots, Matplotlib and Seaborn are used.
- Heroku is used for deployment of the model.
- Front end development is done using Streamlit
- GitHub is used as version control system.

2.7 Constraints

The AQI prediction website should be user friendly.

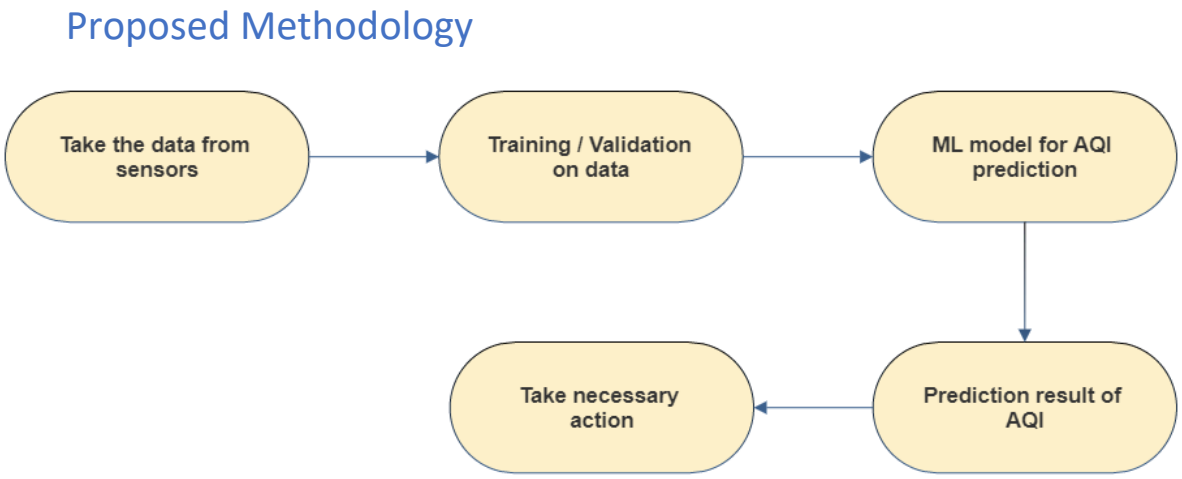
2.8 Assumptions

The main objective of the project is to ingest live data from IoT devices and AQI monitoring parameters for current air based on the training data used by using Machine Learning. It is also assumed that all aspects of this project have the ability to work together in the way the designer is expecting.

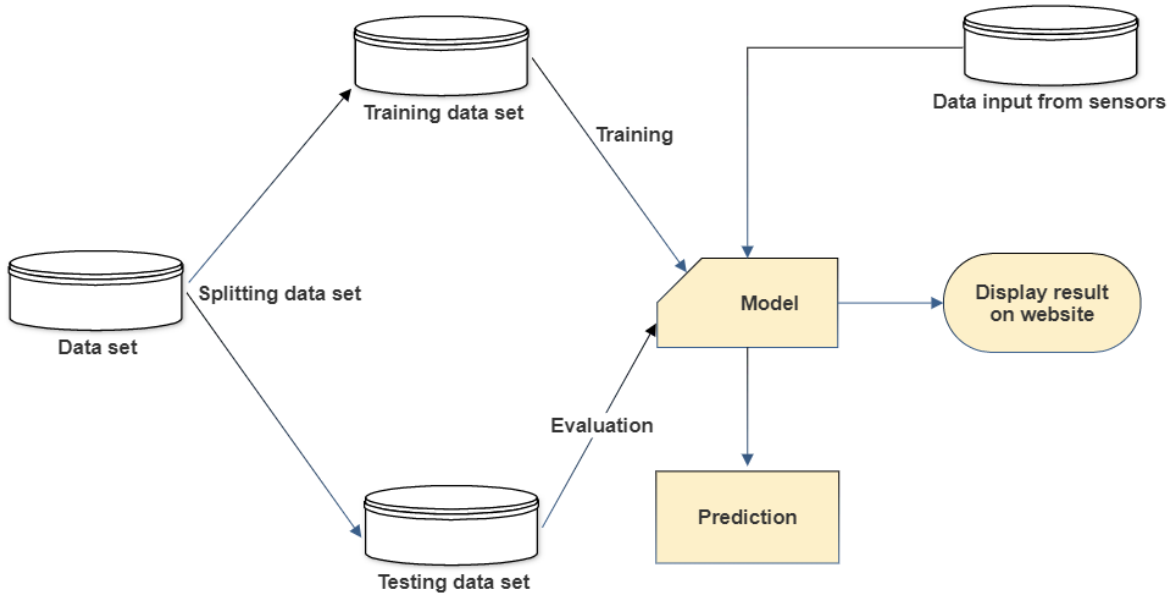
3 Design Details

3.1 Process Workflow

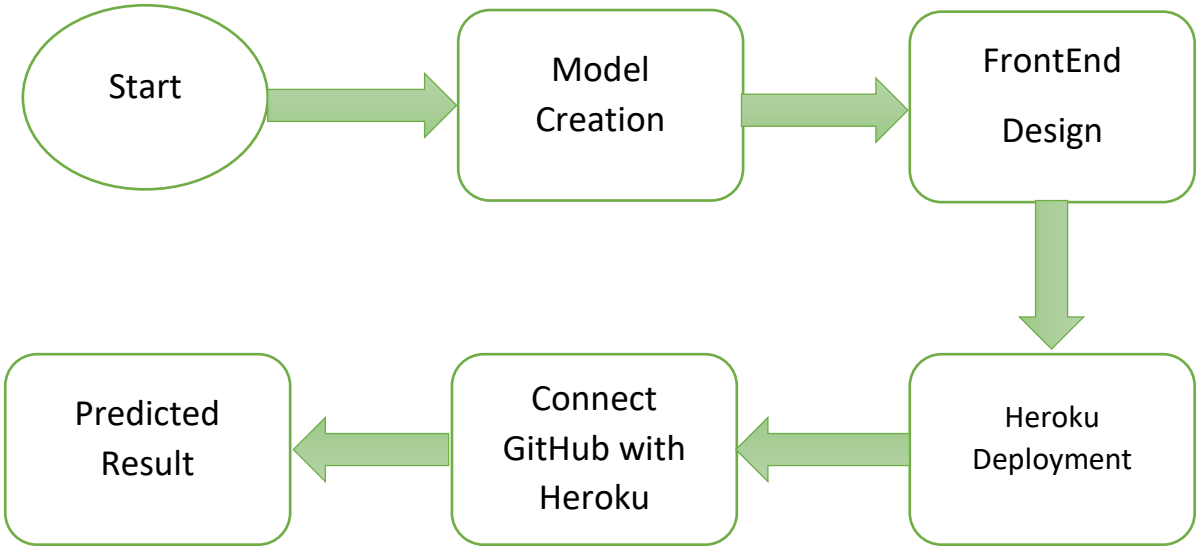
For identifying the different types of anomalies, we will use a machine learning model. Below is the process flow diagram.



3.1.1 Model Training and Evaluation



3.1.2 Deployment Process



3.2 Error Handling

Should errors be encountered, an explanation will be displayed as to what went wrong? An error will be defined as anything that falls outside the normal and intended usage.

4 Performance

AQI Prediction should be as accurate as possible, so that it will not mislead the user. Best possible model will be used to predict AQI.

4.1 Reusability

The code written used should have the ability to be reused with no problems.

4.2 Application compatibility

Python is use for model creation, so it is compatible for any platform.

4.3 Resource utilization

When any task is performed, it will likely use all the processing power available until that function is finished.

4.4 Deployment

5 Conclusions

This project proposes the machine learning model for AQI prediction. This model can be used for alerting the citizens when there is air quality high. It will be helpful to reduce disease and problems occur by air pollution. It also helps the government to take an action in that area against air pollution.

6 Reference

1. https://docs.streamlit.io/en/stable/
2. https://scikit-learn.org/stable/user_guide.html