Name: Muhammad Ahmed Javed **Name: Misbah Munir**

Roll#: L21-7692 **Roll#: L21-6237**

Section: BDS-5D

Project Title: Air Quality Monitoring and Prediction in Pakistani Cities

Background/Problem Statement:

"Air pollution is a critical environmental issue that poses significant risks to public health and the overall well-being of communities. Major cities in Pakistan are grappling with increasing levels of air pollution, adversely affecting the quality of life and posing health hazards to residents". In response to this pressing concern, our project aims to create an integrated system for monitoring, analyzing, and predicting air quality in these urban areas.

Project Description:

We are developing a comprehensive system for monitoring and predicting air quality in major cities across Pakistan. The project involves collecting historical air quality data, analyzing the impact of various factors (e.g., weather conditions), and building a predictive model for future air quality. Additionally, visualize the data to provide insights and warnings to the public.

Dataset Sources:

1. **VisualCrossingWebServices/API**
2. **WAQI**
3. **OpenWeatherMap API**

Tools:

1. **Data Scraping and Wrangling:**
   - We will use Python with libraries like selenium, BeatifulSoup, pandas, and scrapy for scraping and cleaning data (removing outliers, handling missing values etc.)
2. **Data Analysis and Visualization:**
   - We will utilize Jupyter Notebooks with Pandas, Matplotlib, and Seaborn for exploratory data analysis.
   - Plotly for interactive visualizations to showcase air quality trends over time and much more.
3. **Geospatial Visualization:**
   - We will use Folium, GeoPandas, and Plotly to create a geospatial visualization of air quality across different cities
4. **Models:**
   - Scikit-Learn for machine learning tasks, feature extraction (including correlation analysis and univariate analysis), and model evaluation. We will evaluate the model based on accuracy, recall, and precision.

- We have planned to implement *TensorFlow* and *Keras* for building and training deep learning models

Data Sources for City Images:

- **Google Maps API**
- **Government Websites**

Working Process:

We are diving our tasks into 3 categories:

a) **Task#1:**

Data collection and pre-processing.

b) **Task#2:**

Exploratory Data Analysis includes geospatial analysis as well.

c) **Task#3:**

Model Development

Note:

There can be minor changes according to our needs or requirements. We may select other data sources as per need but using deep learning models, and data collection with scraping or using API will not be changed. In the **Background** heading the sentences in quotation marks are our problem statement.