**Project Management Proposal**

**Aim & Goal**

The aim of this project is to develop a robust and accurate air quality prediction system capable of forecasting key air pollutants (e.g., PM2.5, PM10, NO₂, SO₂, O₃) with high temporal and spatial resolution. The system will leverage machine learning and deep learning techniques to analyze historical and real-time data, providing forecasts up to 72 hours in advance. The goal is to empower policymakers, urban planners, and the public with actionable insights to mitigate health risks and support environmental planning.

**Team Members**

- Taqi Haider

- Shuai Hu

- TBD

- TBD

**Technical Proposal**

- Sources: Historical air quality data from government monitors, satellite imagery (e.g., NASA MODIS), meteorological data (temperature, humidity, wind speed/direction), traffic data, and industrial emission records.

- Tools: custom develop platform, web scraping, cloud storage

- Preprocessing: Handle missing values (interpolation), normalize features, and align temporal/spatial resolutions.

* + Feature Engineering
  + Temporal features、Spatial features、Lag features
* Model development
* Development
  + Cloud-based API (using Flask/Django, streamlit) for predictions.
  + Dashboard for visualization (Tableau or custom web app with React/D3.js, streamlit).
  + Automated pipeline (Apache Airflow) for data ingestion, preprocessing, and retraining.

**Expected Outcomes**

* + A functional air quality prediction system with ≥85% accuracy (R² score) for 24-hour forecasts.
  + A public dashboard for real-time air quality updates and forecasts.

**Risks & Mitigation**

- Data quality issues: Backup data sources and robust preprocessing.

- Model underperformance: Ensemble methods and hyperparameter tuning.

- Deployment delays: Agile development with iterative testing.

**Timeline**