

November 15, 2025

AIR QUALITY FORECASTING PROJECT





Project details

Project Title: Air Quality Forecasting Project

Purpose:

This project addresses the significant health and environmental risks posed by air pollution in urban areas. By developing a time series forecasting system, it aims to predict Air Quality Index (AQI) and pollutant levels, aiding authorities and the public in making timely and informed decisions to promote healthier city environments.

Tools and Technologies: Streamlit, Python

Mentor: Ms. Umme Asma S

Batch Number: 3

Start Date: 22 September 2025

Intern Name: Payal Rawat

Problem Statement

Air pollution is a growing concern in urban environments, directly impacting human health and environmental quality. Accurate forecasting of air quality can help authorities take preventive measures, issue warnings, and inform the public.

Sources of air pollution



Industrialization
traffic congestion
Fossil fuels
Deforestation

Key Pollutants



Particulate matter from combustion and dust, affecting lungs.



PM10

Particulate matter affecting lung function significantly.



Gases from vehicles and industry, respiratory irritants.



SO₂

Gases from industry causing respiratory irritants.



Carbon monoxide from incomplete combustion, harmful if inhaled.



O₃

Ground-level ozone impacting respiratory health directly.

Risk of Air Pollution

Health problem

- Respiratory Diseases
- Cardiovascular Issues
- Other Ailments

Environmental Impacts

- Crop Damage
- Climate Change
- Ecosystem Degradation

Air quality forecasting project

Purpose and focus of project

This project aims to build a time series forecasting system that predicts air quality index (AQI) and pollutant level using historical pollution data .The goal is to empower agencies, city planner, and the public with timely information for healthier decision-making.

Project Objective

- A time series forecasting model to predict AQI and key pollutants (PM2.5, PM10, NO2, etc.).
- Visual dashboard showing historical trends and future air quality predictions.
- Alerts or warnings when predicted AQI crosses safety thresholds.
- Ability to analyze pollutant contributions and seasonal trends.
- Admin functionality to upload new datasets and retrain models for updated forecasting.

Technology_Stack

- Python for core development
- Pandas/NumPy for data manipulation
- Scikit-learn for machine learning models
- TensorFlow for deep learning(LSTM) models
- ARIMA for statistical time-series forecasting
- Prophet for trend-based time-series forecasting
- XGBoost-for boosting -based regression prediction
- Streamlit for web application
- Plotly/Matplotlib for visualization

Milestone 1: Data preprocessing and EDA

Data Sources & Agencies

CPCB & SAFAR

Official government monitoring stations provide reliable data.

OpenAQ & WAQI

Crowdsourced platforms enhance coverage and engagement.

EDA (Exploratory Data Analysis)

Purpose: Understand data, detect patterns, missing values & outliers, plan features

Visualizations:

Time Series Plots – Show pollutant trends, daily cycles, support forecasting

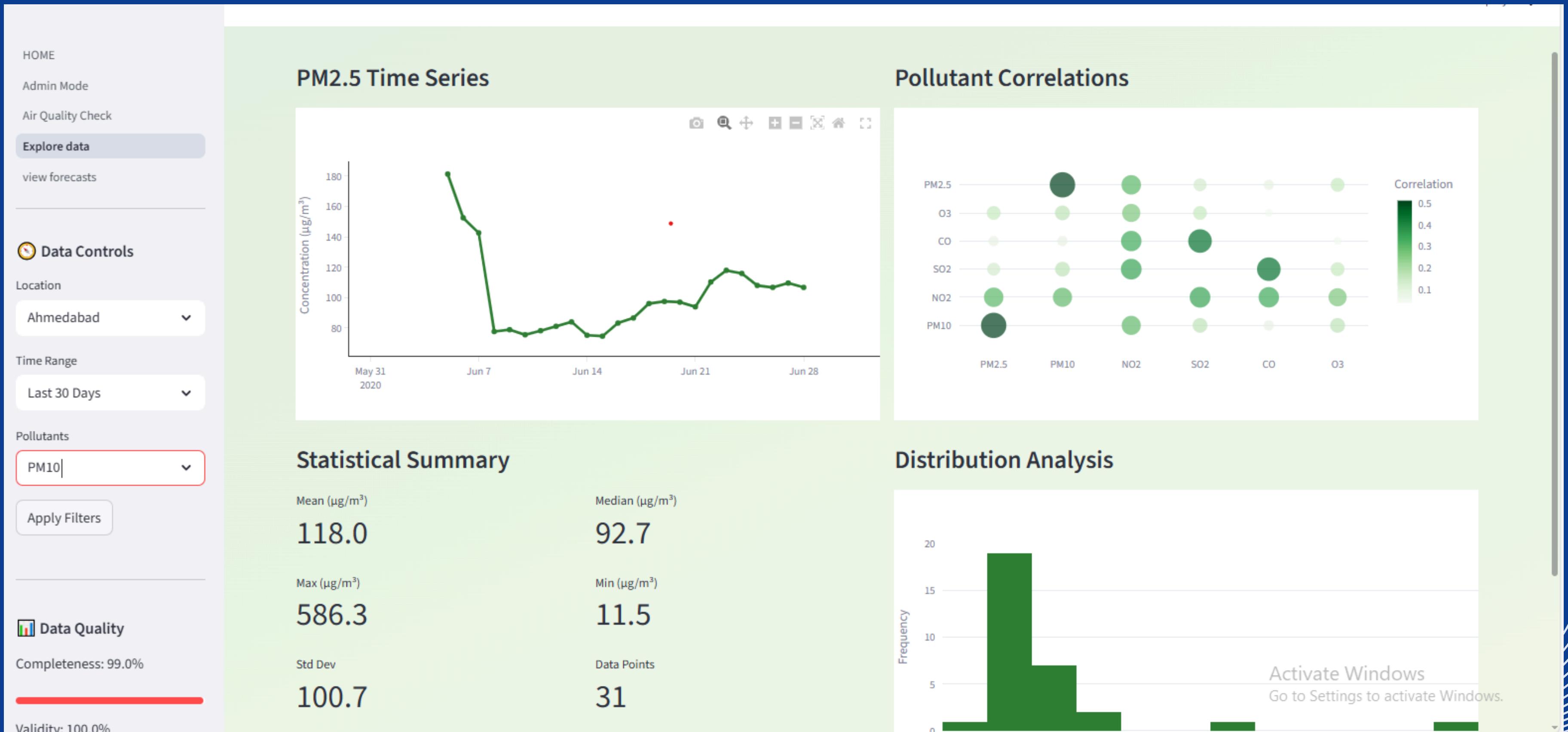
Correlation Analysis – Identify related pollutants, improve model features

Statistical Summary – Mean, median, max/min; detect extremes

Distribution Analysis – Frequency of values, detect outliers, check skewness

Dashboard 1

Air quality data explorer



Milestone 2

Model training and Evaluation



ARIMA

ARIMA is a statistical time-series model that forecasts future values based only on past values. It combines:

AR (Auto Regression) – depends on previous values

I (Integrated) – makes data stationary

MA (Moving Average) – uses past errors



LSTM

A deep learning model for sequential data analysis.

- ✓ Handles sequential time-series data like pollutant levels over hours/days
- ✓ Learns complex patterns & trends automatically
- ✓ Works well when data is large and pollution depends on previous values Useful for multi-step forecasting (next hour/day/week pollution levels)



XGBoost

A gradient boosting algorithm for high performance.

- ✓ Handles tabular data (pollutant values + weather features)
- ✓ Works well with small & medium datasets
- ✓ Captures non-linear relationships
- ✓ Faster and more accurate than normal regression models
- ✓ Good for feature importance — tells which factor affects pollution most



Prophet

Prophet is an open-source time-series forecasting model developed by Facebook. It automatically detects:

- ✓ Trend
- ✓ Seasonality (daily, weekly, yearly)
- ✓ Holidays/festivals impact

Visualization & Insights

Model Comparison Chart – Compare RMSE/MAE across pollutants

Forecast Graph – Actual vs predicted trends with confidence intervals

Best Model Table – Shows optimal model & performance per pollutant

Forecast Accuracy Chart – Accuracy vs forecast horizon

The screenshot shows a sidebar with the following navigation links:

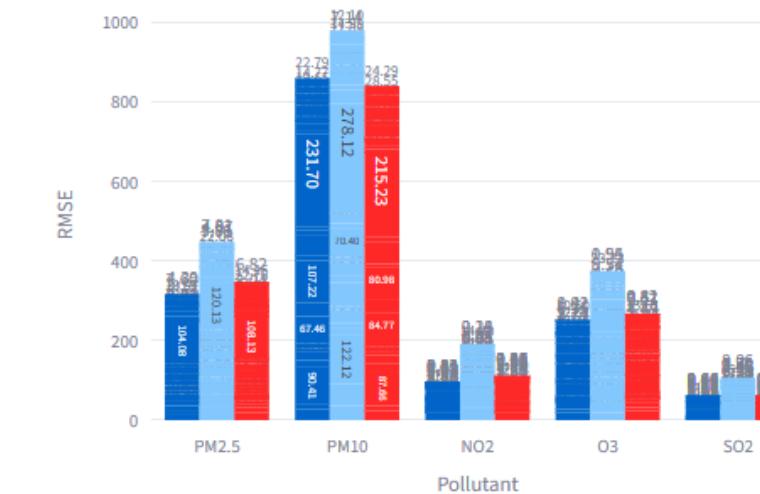
- HOME
- Admin Mode
- Air Quality Check
- Explore data
- view forecasts**

Air Quality Forecast Engine

Model Performance

RMSE MAE

Model Performance (RMSE)



Best Model by Pollutant

	City	Pollutant	Best_Model	RMSE
0	Ahmedabad	PM2.5	ARIMA	16.7354
1	Ahmedabad	PM10	XGBoost	22.8277
2	Ahmedabad	NO2	Prophet	17.0911
3	Ahmedabad	O3	ARIMA	18.2229
4	Ahmedabad	SO2	ARIMA	1.8653
5	Chennai	PM2.5	Prophet	5.3932
6	Chennai	PM10	XGBoost	12.6622
7	Chennai	NO2	ARIMA	1.7933
8	Chennai	O3	XGBoost	8.8046
9	Chennai	SO2	XGBoost	1.1915

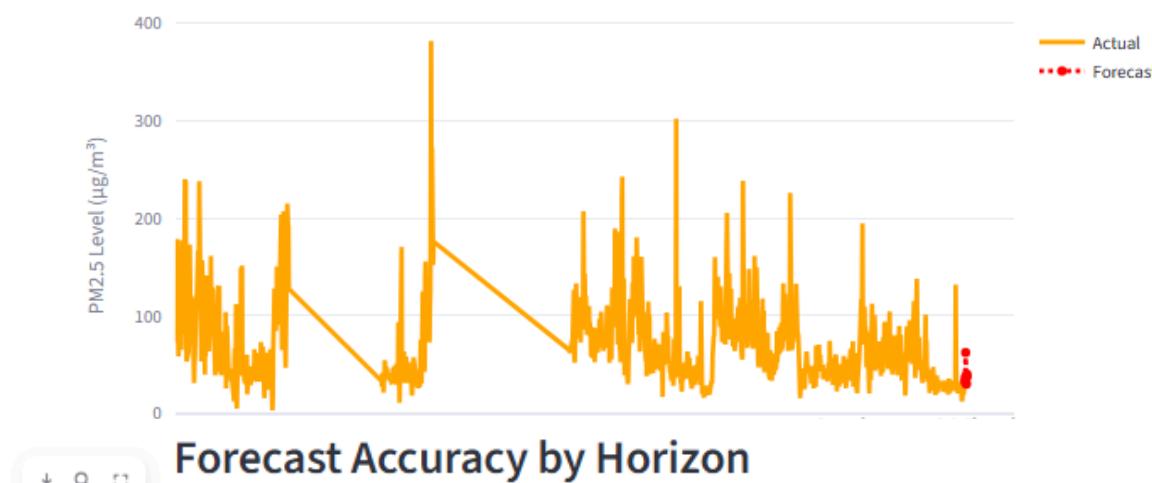
PM2.5 Forecast

City

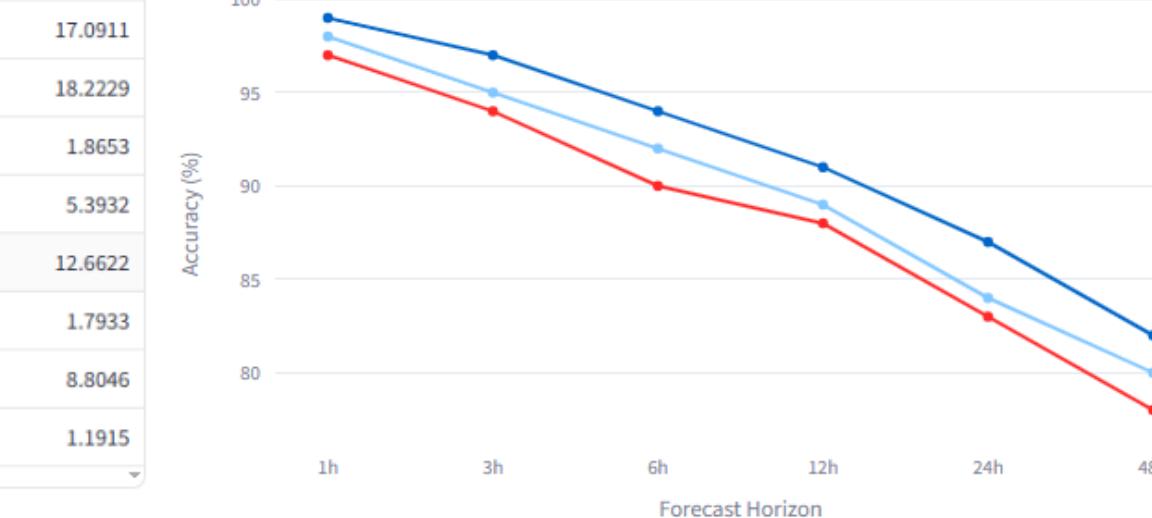
Ahmedabad

Pollutant

PM2.5



Forecast Accuracy by Horizon



Milestone 3

Alert logic and visualization

what is AQI:-AQI is a standardized index that converts pollutant levels into a single value to represent overall air quality.



Visualization



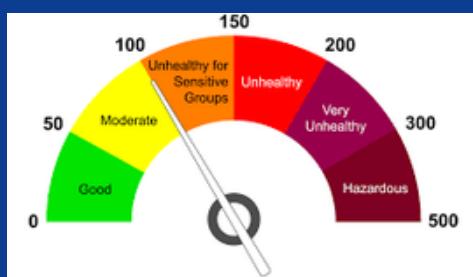
Real-time AQI Gauge

Display current air quality status dynamically.



7-day Forecast Cards

Show upcoming air quality trends visually.



Pollutant Concentration Charts

Analyze detailed pollutant levels effectively.

HOME
Admin Mode
Air Quality Check
Explore data
view forecasts

City: Ahmedabad
Forecast range: 2020-06-22 → 2020-07-01
Pollutant for trend: PM10

Air Quality Alert System

Milestone 3: Working Application — 7-day forecasts, AQI, and pollutant trends

Current Air Quality



7-Day Forecast



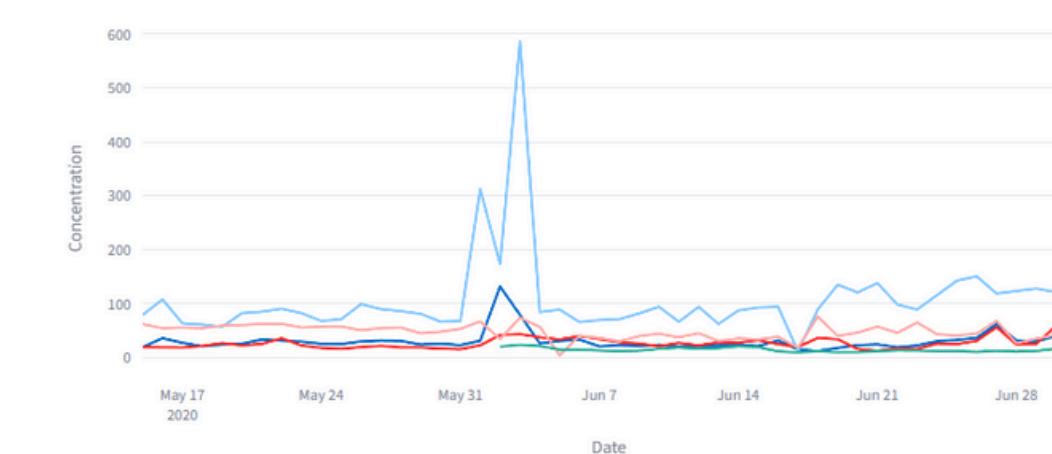
HOME
Admin Mode
Air Quality Check
Explore data
view forecasts

City: Ahmedabad
Forecast range: 2020-06-22 → 2020-07-01
Pollutant for trend: PM10

Pollutant Analysis & Alerts

Pollutant Concentrations (48 hrs)

Recent Pollutant Concentrations



Active Alerts (Next 7 Days)

- Unhealthy for Sensitive — Mon, 22 Jun
Predicted AQI: 121
- Unhealthy for Sensitive — Tue, 23 Jun
Predicted AQI: 121
- Unhealthy for Sensitive — Thu, 25 Jun
Predicted AQI: 121
- Unhealthy for Sensitive — Fri, 26 Jun
Predicted AQI: 121
- Unhealthy for Sensitive — Sat, 27 Jun
Predicted AQI: 121
- Unhealthy for Sensitive — Mon, 29 Jun
Predicted AQI: 121
- Unhealthy for Sensitive — Tue, 30 Jun
Predicted AQI: 121

Milestone 4

streamlit web dashboard

1. Interactive UI to select station, timeframe, and forecast variable.

- Display AQI gauge, line plots, and alerts.

- Admin interface to upload new data and retrain models.

The screenshot shows a Streamlit dashboard titled "Air Quality Prediction System". The top navigation bar includes links for "HOME", "Admin Mode" (which is selected), "Air Quality Check", "Explore data", and "view forecasts". A sidebar menu offers "User Panel" and "Data Upload & Training" options. The main content area displays real-time air quality data for the "Jorapokhar" station on "2025/11/17". The data includes PM2.5 (14.00), PM10 (58.00), CO (48.00), NO₂ (71.00), SO₂ (64.00), O₃ (24.00), Temperature (°C) (-10.00), and Humidity (%) (0.00). Below the data is a "Predict AQI" button, which has triggered a green success message: "Predicted AQI: 88.7" and "Category: Moderate". At the bottom, there's an "Upload Data & Retrain Models (Open Access)" section with a file upload input and a "Retrain All Models" button. A watermark for "Activate Windows" is visible in the background.

wf Air Quality Monitoring Dashboard

Explore, Analyze, and Forecast Air Quality Data

Explore Data

View city-wise pollutant levels, summary, and statistics.

[Go → Explore Data](#)

View Forecasts

Check time-series predictions for PM2.5 and other pollutants.

[Go → View Forecasts](#)

Air Quality Check

AQI insights and pollution category classification.

[Go → Check AQI](#)

Retrain Model

check AQI,upload files.

[Go → Retrain Model](#)

Developed for Air Quality Analysis Project

Activate Windows
Go to Settings to activate Windows.

Thank You

