AirAware: Smart Air Quality Prediction System

Unveiling our progress on AirAware, a cutting-edge system designed to predict air quality with precision and empower users with crucial environmental insights.

This presentation covers the significant strides made in Milestone-3, marking 75% project completion.

Milestone-4: Laying the Foundation

Data Collection & Prep

Acquired Delhi AQ dataset from Kaggle, followed by meticulous cleaning and normalization processes.

Backend & Database

Successfully set up MongoDB for data storage and developed a basic FastAPI backend with essential endpoints.

Dashboard & Docs

Implemented a React dashboard skeleton, including summary cards, data tables, and comprehensive project documentation.

Milestone-1 established the critical infrastructure, providing a robust base for advanced features in Milestone-2.

Milestone-4 Goals: Advancing Prediction & Insights

Prediction Engine Development

- Build a complete Machine Learning pipeline.
- Train and evaluate three distinct models: Linear Regression, Random Forest, and SVM.
- Predict the next PM2.5 value.
- Compare actual vs. predicted values and compute accuracy metrics.
- Develop and integrate dedicated Prediction APIs.

Dashboard & Chatbot Integration

- Integrate prediction capabilities into the existing dashboard.
- Create an interactive error heatmap for visual analysis.
- Develop the AirAware Chatbot as a separate, intelligent module.

ML Pipeline: From Data to Prediction

01

Data Loading & Target Variable

Cleaned dataset loaded. Defined "next PM2.5" as the primary target for prediction.

02

Feature Selection

Identified key features: PM2.5, PM10, CO2, NO2, SO2, temperature, and humidity for model training.

03

Train-Test Split & Model Training

Data partitioned into training and testing sets. Three distinct ML models were trained.

04

Evaluation & Selection

Models rigorously evaluated using R^2, MSE, and RMSE. The best-performing model was selected.

05

Model Persistence

The chosen model and a comparison dataset were saved for future use and integration.

Models Trained & Best Performer

Linear Regression

A foundational model, providing a baseline for predictive performance.

Random Forest Regressor

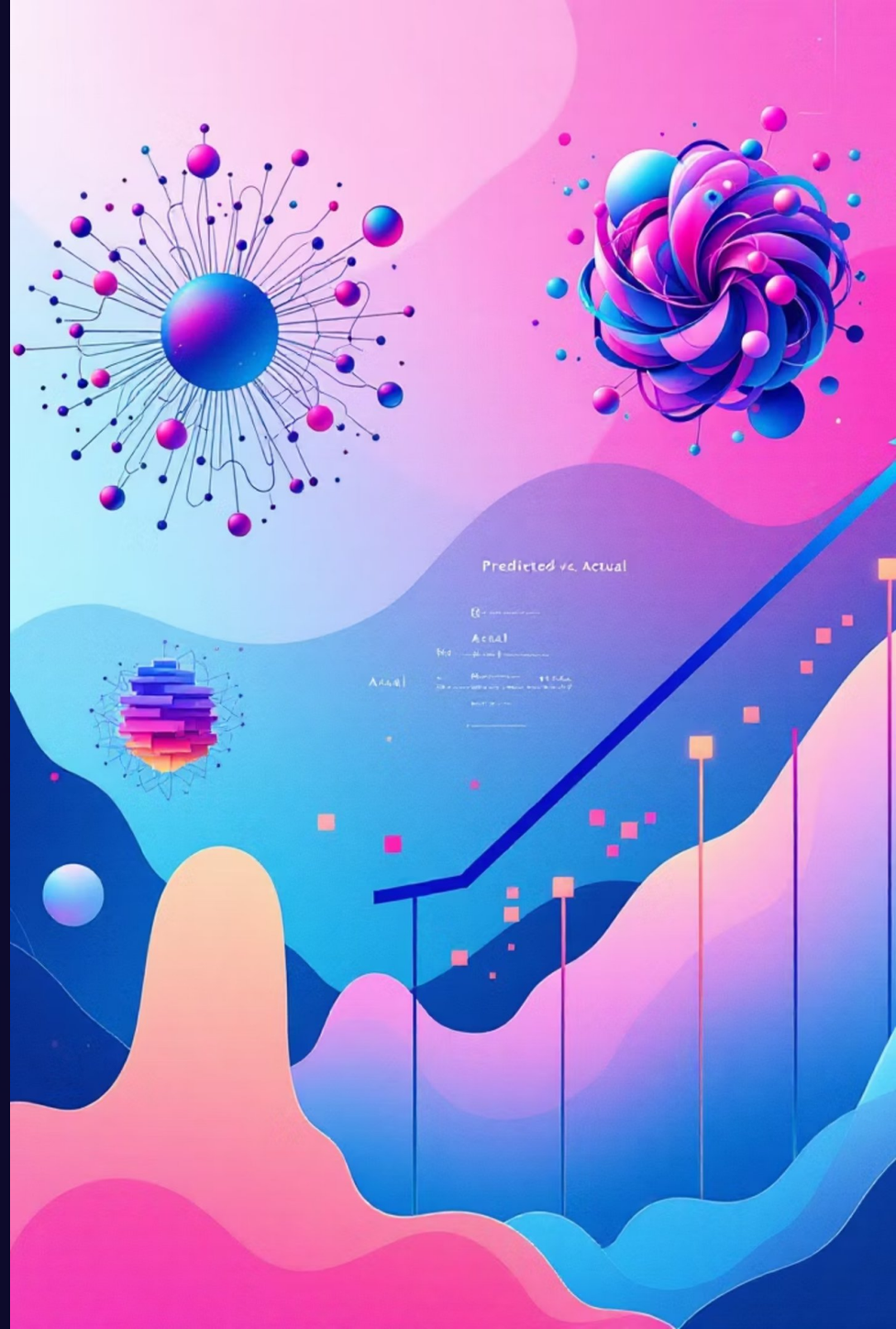
Ensemble learning at its best, exhibiting superior accuracy for PM2.5 prediction.

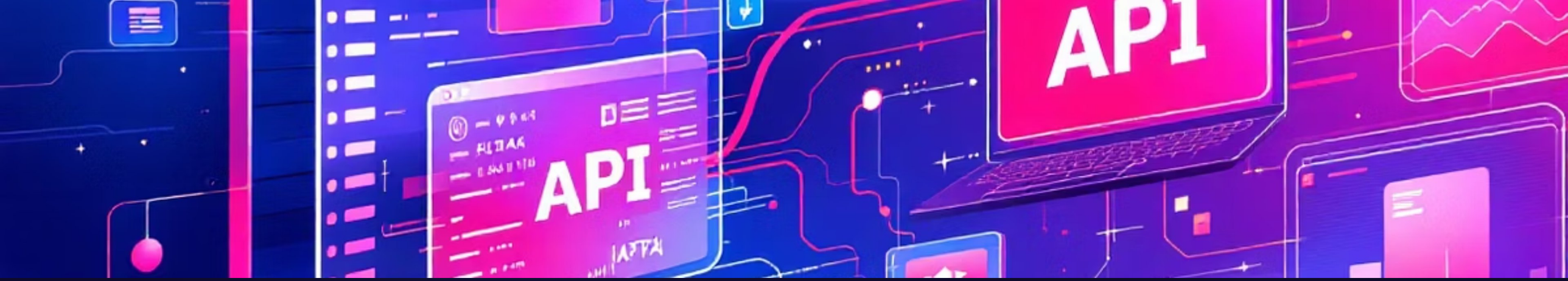
Support Vector Machine (SVR)

Effective in high-dimensional spaces, offering another perspective on predictive capabilities.

After thorough evaluation, the **Random Forest Regressor** emerged as the top performer in accuracy, validated by our demo values.

- Generated Outputs: Predicted values, actual vs. predicted graphs, accuracy tables, and error heatmap data.





Prediction APIs: Fueling Frontend Insights

/predict/latest

Retrieves the next predicted PM2.5 value and its corresponding Air Quality category.

/predict/batch?limit=100

Provides arrays of actual and predicted values for a specified limit, enabling historical comparison.

/predict/accuracy

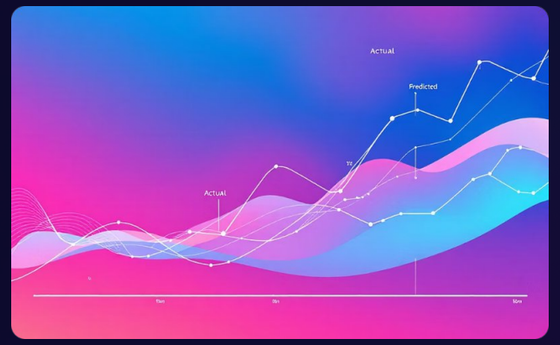
Returns key accuracy metrics: R^2 Score, Mean Squared Error (MSE), and Root Mean Squared Error (RMSE).

/predict/heatmap

Delivers the necessary data matrix to generate a visual error heatmap on the dashboard.

These new FastAPI endpoints are fully integrated and provide real-time data to the AirAware dashboard, enhancing the user experience with dynamic, AI-driven insights.

Dashboard Enhancements: Visualizing AI Insights



Prediction Chart

A dynamic graph showcasing the "Actual vs Predicted PM2.5" values over time for clear comparison.

Prediction Card

A concise display of the "Next Predicted PM2.5 Value" for immediate user awareness.

Accuracy Card

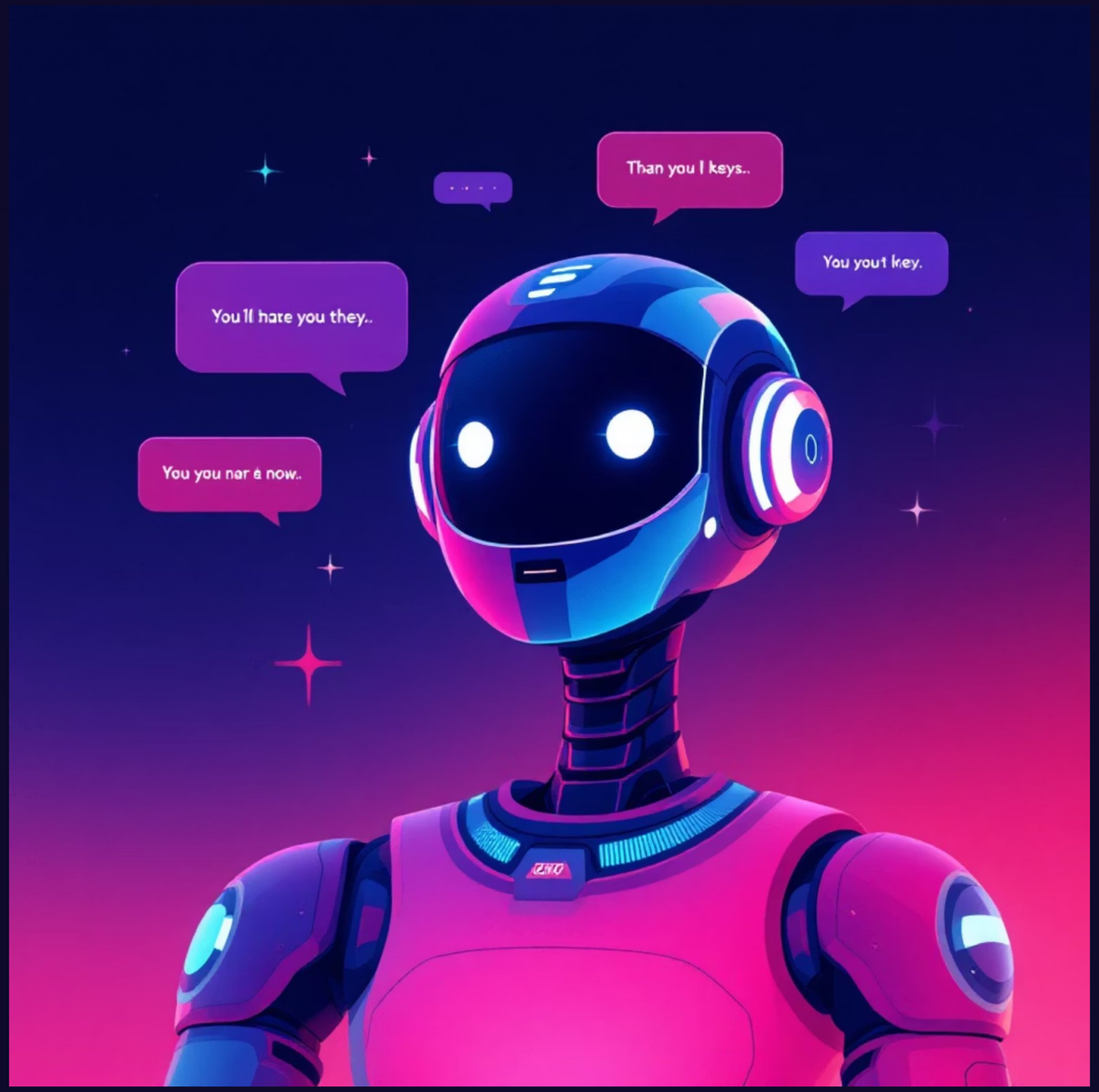
Presents core accuracy metrics (R^2, MSE, RMSE) to convey model performance.

Heatmap Component

Visually represents the error distribution of predictions, highlighting areas of greater variance.

The AirAware dashboard now offers comprehensive, AI-based insights, transforming raw data into actionable information for users.

AirAware Assistant: Your Intelligent Chatbot



1

Pollutant Explanations

Clarifies various pollutants (PM2.5, NO2, CO2) and their impact.

2

AQI Category Breakdown

Helps users understand different Air Quality Index categories and what they mean.

3

Health Suggestions

Offers personalized health advice based on current air quality conditions.

4

Project Q&A

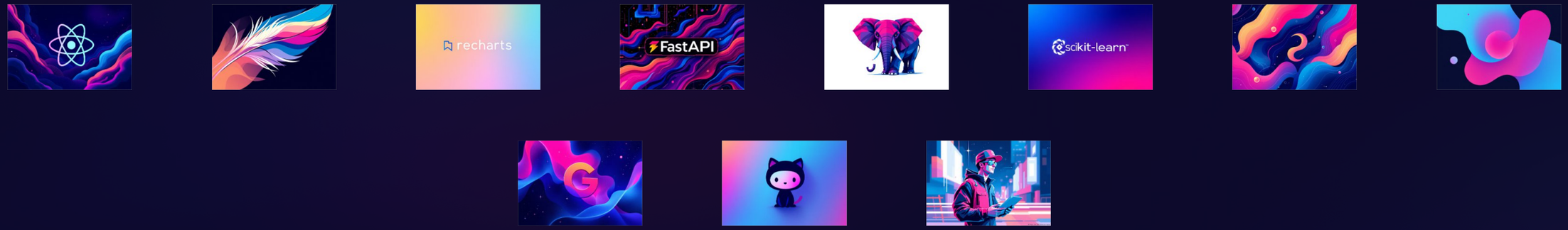
Answers questions about the AirAware project and its functionalities.

5

ML Simplified

Explains complex machine learning concepts in easy-to-understand terms.

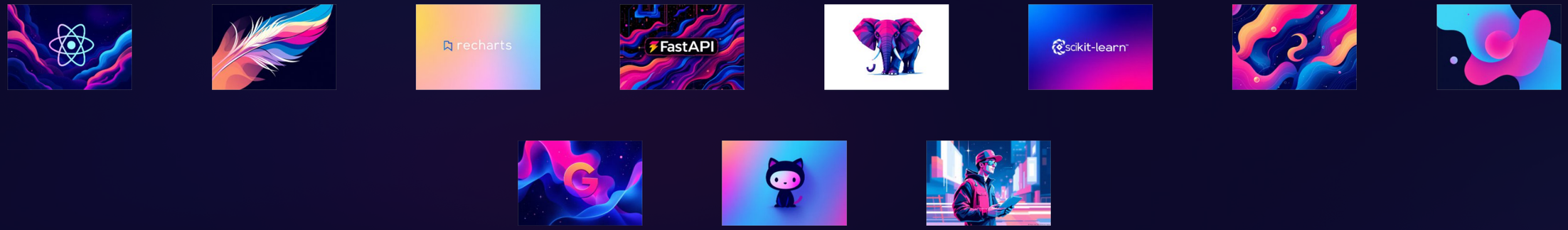
Technologies Powering AirAware



Our robust technology stack ensures a scalable, efficient, and user-friendly AirAware system.

From frontend development to advanced machine learning and natural language processing, each tool plays a critical role in bringing our vision to life.

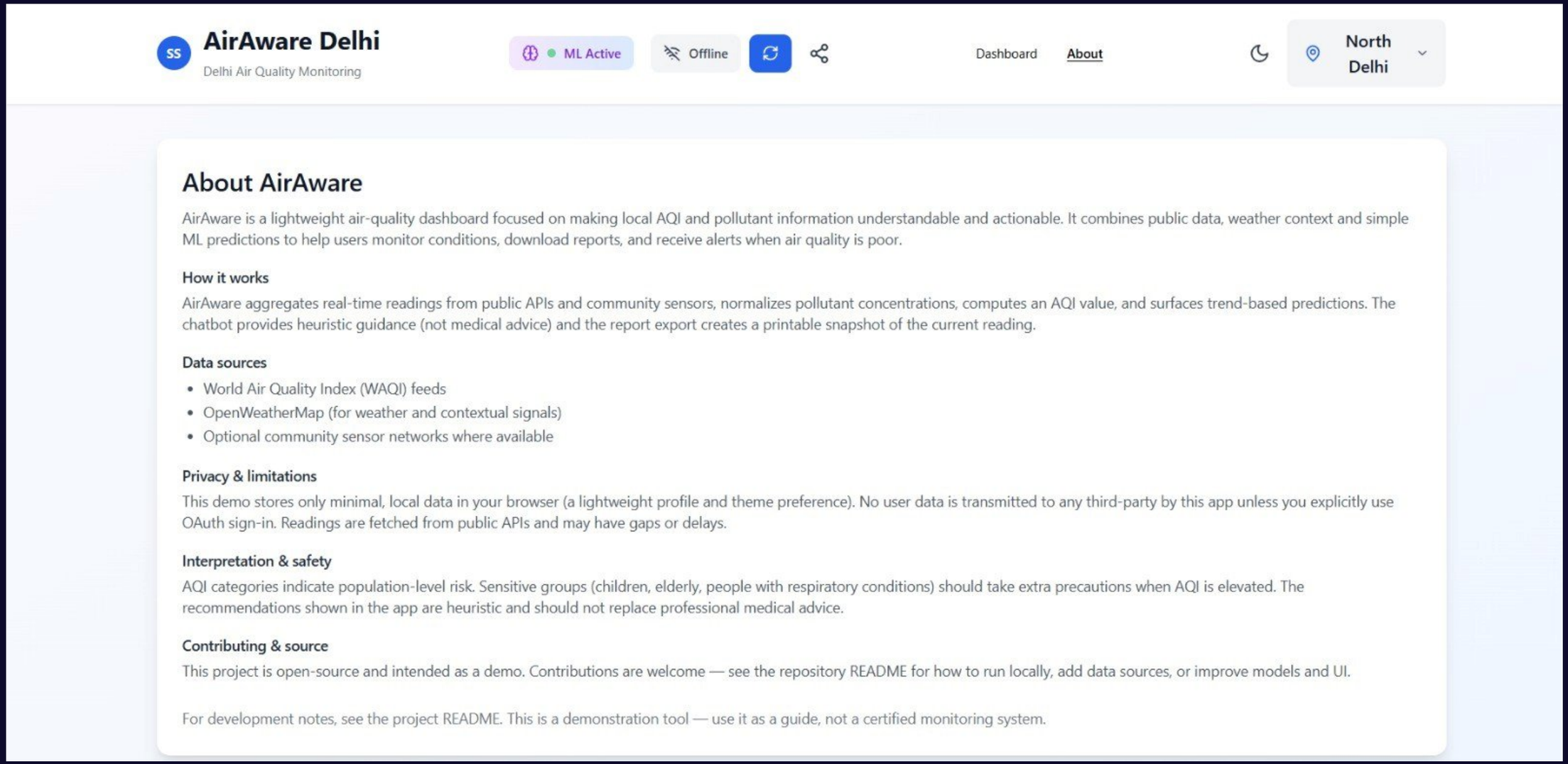
Technologies Powering AirAware



Our robust technology stack ensures a scalable, efficient, and user-friendly AirAware system.

From frontend development to advanced machine learning and natural language processing, each tool plays a critical role in bringing our vision to life.

OUTPUT SCREENS



SS **AirAware Delhi**
Delhi Air Quality Monitoring

ML Active | Offline

Dashboard | About

North Delhi

Current Air Quality

Hazardous

Last updated: 2 minutes ago • North Delhi

430
AQI Index

0 Good | 50 | 100 | 150 | 200 | 300+ Hazardous

What this means
Health warnings of emergency conditions. Sensitive groups should consider reducing prolonged outdoor activities.

Quick Stats

Current AQI	**430**
Category	**Hazardous**
Last Updated	12:08:56 am
7-Day Avg	**430 AQI**

Real-time Updates
Manual refresh only

Weather
Temp: 14.8°C
Wind: 3.1 km/h

Download AQI Report

Pollutant Breakdown

PM2.5	425 µg/m³	100%
PM10	430 µg/m³	100%

Health Recommendations

- Limit outdoor activities
- Everyone should avoid prolonged outdoor exertion
- Keep windows closed and use air purifiers

Pollutant Breakdown

PM2.5	425 µg/m³	↗ 100%
PM10	430 µg/m³	↗ 100%
O3	2.3 ppb	↘ 3.2857142857142856%
NO2	22.4 ppb	↘ 44.8%
SO2	5.1 ppb	↘ 14.571428571428571%
CO	28.7 ppm	↘ 100%

Legend: PM = Particulate Matter, O3 = Ozone, NO2 = Nitrogen Dioxide, SO2 = Sulfur Dioxide, CO = Carbon Monoxide

Health Recommendations

- Limit outdoor activities
- Everyone should avoid prolonged outdoor exertion
- Keep windows closed and use air purifiers

Sensitive Groups Include:

Children · Elderly · Asthma patients · Heart disease · Active outdoors

Activity Guidance:

Walking:	❌ Not recommended
Running:	❌ Not recommended
Cycling:	❌ Avoid
Indoor workouts:	✅ Safe

North Delhi

Industrial and densely populated area with major commercial hubs.

Key Characteristics:

High vehicle density · Industrial emissions · Residential congestion · Construction activity

ML Model Integration Demo

Test the Gradient Boosting AQI Predictor

API Status

Input Pollutants

Randomize

PM2.5 **45.2** µg/m³	PM10 **78.5** µg/m³	NO2 **38.7** µg/m³	SO2 **12.3** µg/m³
CO **1.2** ppm	O3 **42.1** µg/m³	NH3 **180** ppb	

Predict AQI with ML Model

ML Predicted AQI

79

Moderate

Air quality is acceptable. However, there may be a risk for some people

MODEL DETAILS

Algorithm: Gradient Boosting | Test R^2: 0.9978 | Features: 7 pollutants | Training samples: 151

Live ML Integration: This component connects to the Flask API (localhost:5000) running your trained Gradient Boosting model. Predictions are computed in real-time based on pollutant inputs.

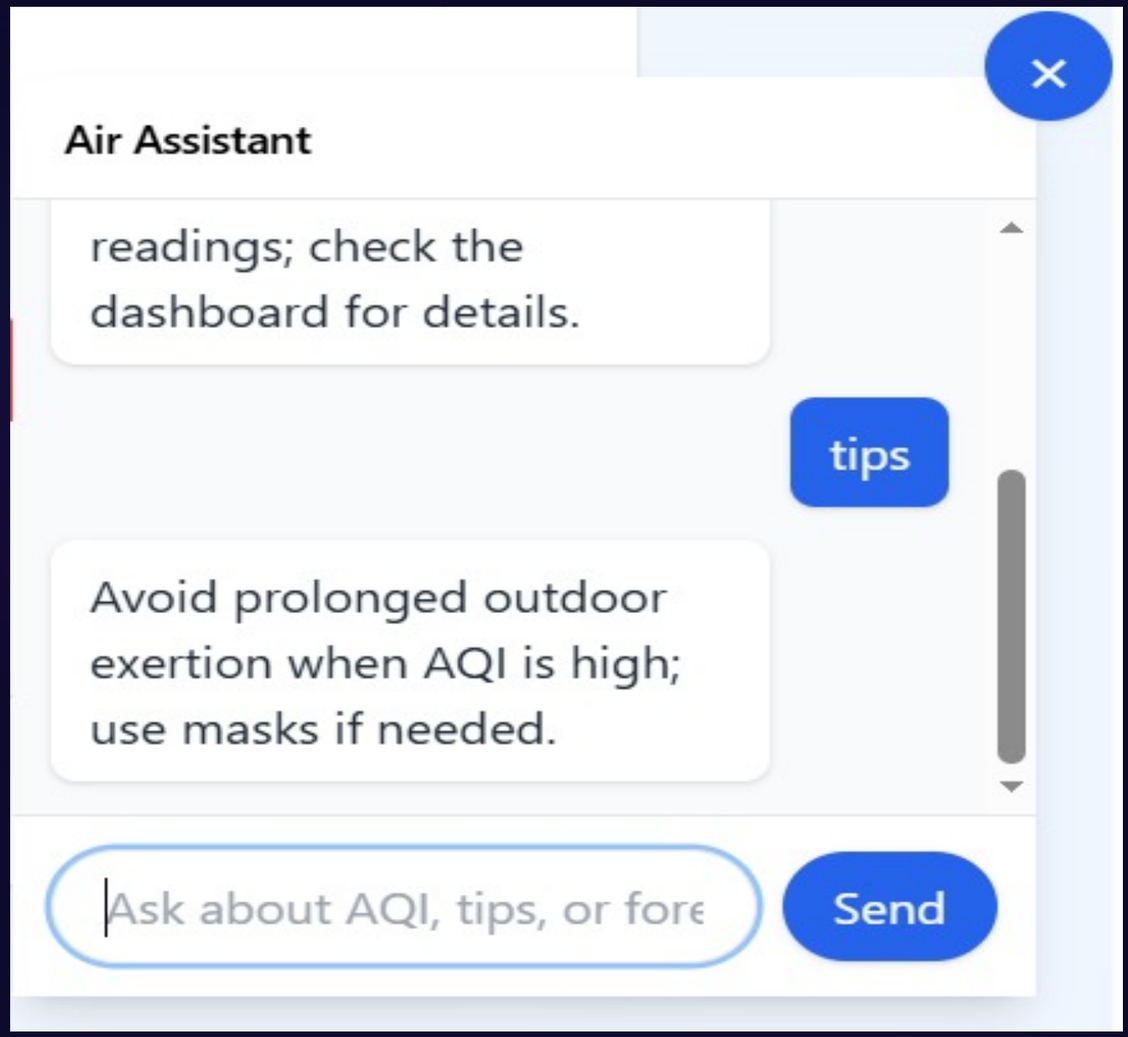
Live ML Integration: This component connects to the Flask API (localhost:5000) running your trained Gradient Boosting model. Predictions are computed in real-time based on pollutant inputs.

7-Day Forecast

Predicted air quality trends

Good (0-50) Moderate (51-100) Unhealthy (101-150) Updated hourly

Data updates on refresh • Predictions powered by machine learning algorithms • Real-time data from WAQI and OpenWeatherMap APIs



AirAware AQI Report

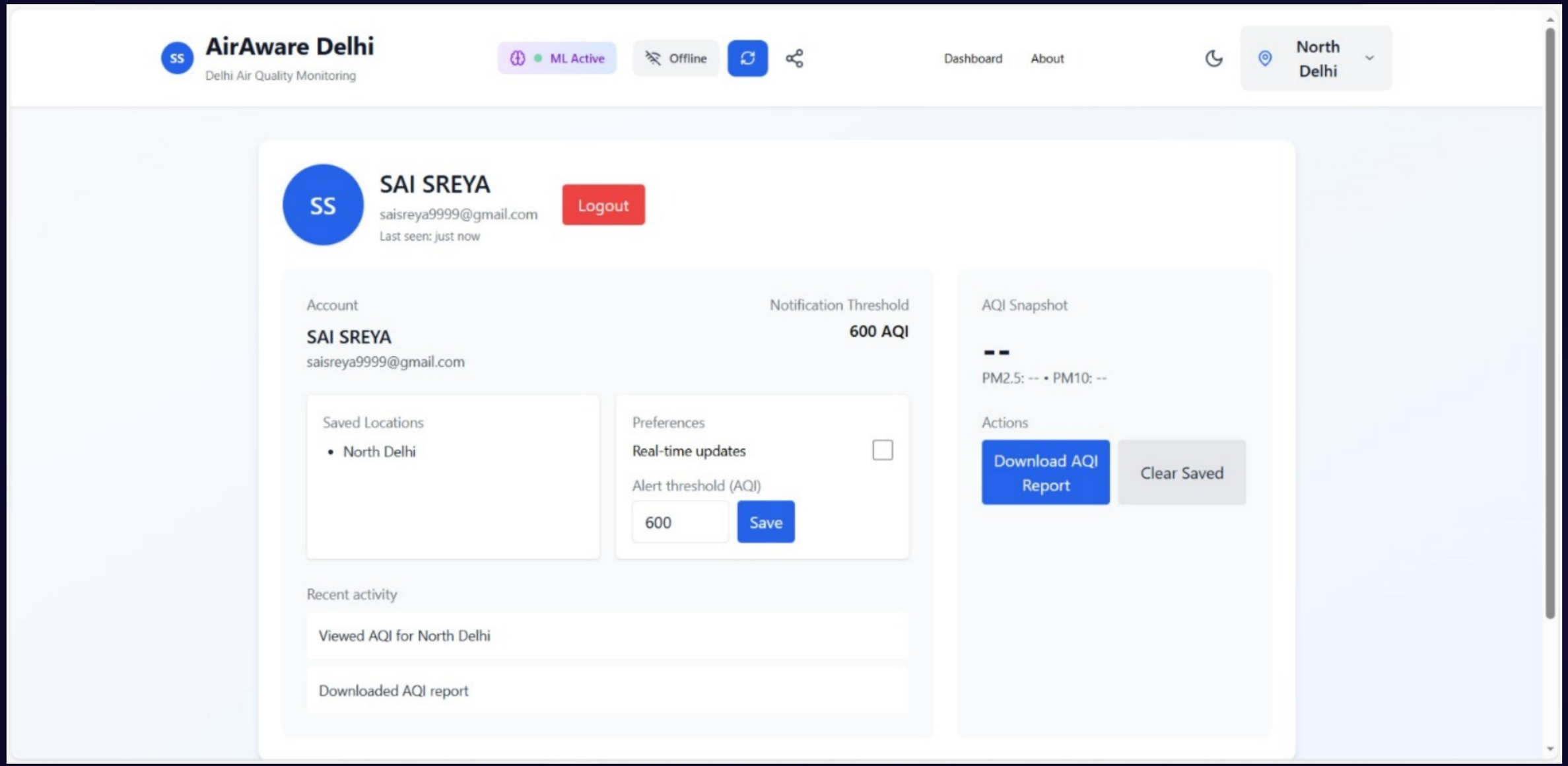
Location: North Delhi

AQI: 430 (Hazardous)

Time: 30/12/2025, 12:08:56 am

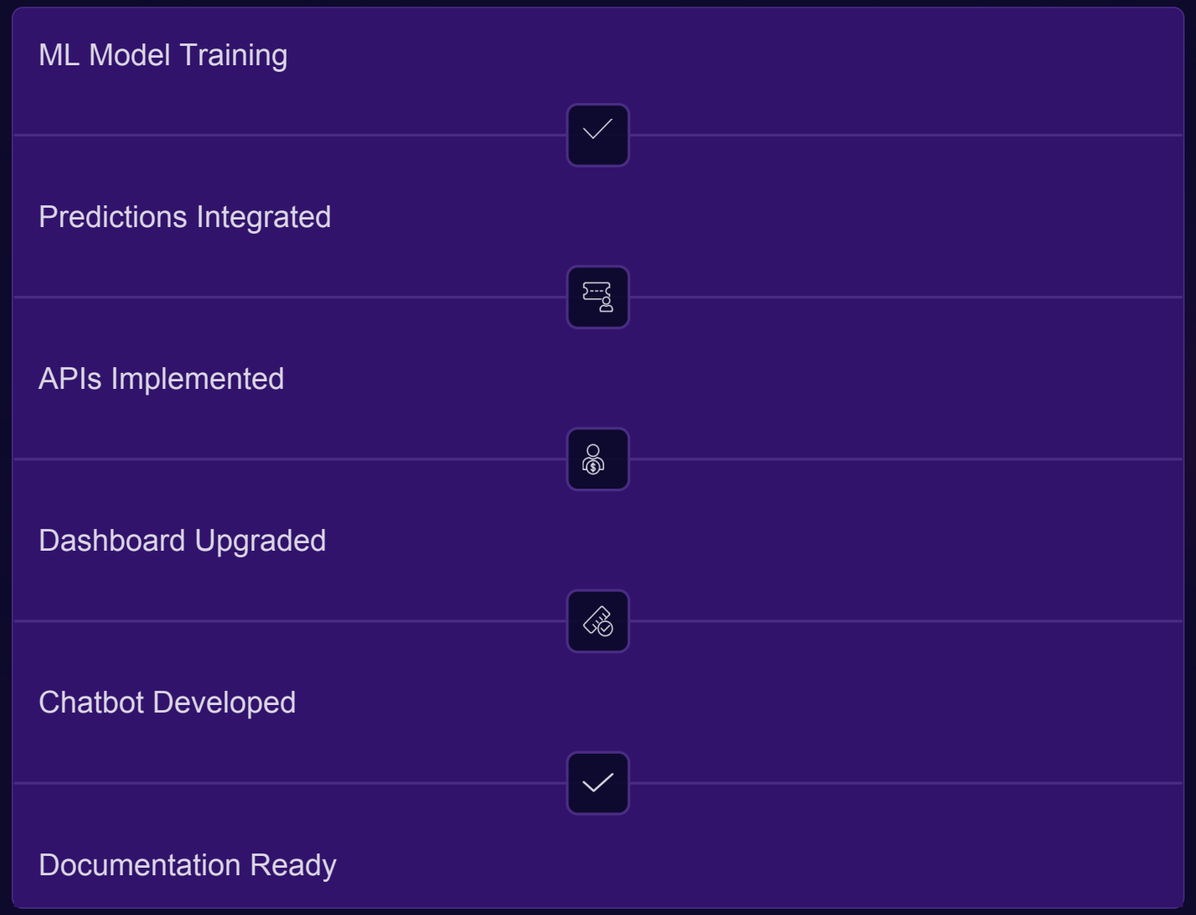
Pollutants

- **PM2.5:** 425 µg/m³
- **PM10:** 430 µg/m³
- **O3:** 2.3 ppb
- **NO2:** 22.4 ppb
- **SO2:** 5.1 ppb
- **CO:** 28.7 ppm



Summary & Next Steps

Milestone-4 Achievements



Milestone-3 (75% Completion)

Multi-Pollutant Prediction

Expanding predictions beyond PM2.5 to include a wider range of pollutants.

Real-time Weather Integration

Incorporating live weather data for enhanced prediction accuracy.

AQI Computation

Developing comprehensive Air Quality Index calculation for holistic understanding.

Alerts Automation

Implementing automated alerts for critical air quality changes.

UI Final Enhancements

Polishing the user interface for an even more intuitive and engaging experience.

Thank you for your attention. We are excited for the next phase of AirAware!

Team 1: Rajalaksmi, Rahul, Sreya, Lokesh, Divija Nandana.