Project Title:

AirAware – Smart Air Quality Dashboard

BATCH- 3

Submitted by:
Neha Nayak

Guided by:
Umme Asma S

Institution:
SDM College of Engineering and Technology, Dharwad
Computer Science and Engineering

Project Duration:
22th September 2025
5421-November 2025

Live Demo:
https://airquality-ciq7gthdhn5n6utyrxsecm.streamlit.app/

Abstract

Air pollution poses one of the most significant threats to human health and environmental sustainability. Rapid urbanization and industrialization have caused a drastic increase in particulate matter and other harmful gases in the atmosphere. Accurate prediction and visualization of air quality trends help authorities and citizens take preventive measures.

This project, **AirAware – Smart Air Quality Dashboard**, aims to analyze, visualize, and forecast Air Quality Index (AQI) levels using **machine learning models** and **interactive web-based visualization**. The system integrates multiple dashboards built in **Streamlit**, leveraging **Facebook Prophet** and **ARIMA models** for forecasting, and visual tools such as **Plotly** and **Seaborn** for detailed analysis.

The final deployed application allows real-time interaction, city-wise filtering, and future AQI forecasting, providing an intuitive and user-friendly experience for researchers, policymakers, and citizens.

Introduction

Air quality monitoring is a critical part of environmental management. The **Air Quality Index (AQI)** is a standardized measure that helps communicate how polluted the air currently is or how polluted it is forecasted to become. High AQI levels can cause respiratory illnesses, cardiovascular issues, and other long-term health impacts.

However, interpreting air quality data from raw measurements is not always intuitive. Traditional reporting lacks visualization and forecasting capabilities that can enhance awareness and enable proactive action.

The goal of **AirAware** is to bridge this gap through an interactive, data-driven system that predicts air quality using statistical models and presents it through a simple, visual, and accessible web dashboard.

Problem Statement

Despite advancements in environmental monitoring, existing systems often:

- Do not provide **predictive insights** into future air quality trends.
- Lack **visual dashboards** that make data intuitive and interactive.
- Require technical knowledge to interpret raw data or graphs.

Hence, there is a need for an intelligent system that:

- Predicts future AQI based on past trends.
- Offers visualization and categorization of pollutants.
- Is accessible through a simple, interactive web interface.

Objectives

1. To build a data-driven dashboard for monitoring and forecasting air quality.
2. To visualize and compare pollutant concentrations across cities.
3. To predict future AQI using **Prophet** and **ARIMA** models.
4. To classify air quality levels based on WHO guidelines.
5. To host the dashboard online for public use.

System Requirements

Software Requirements

Component	Description
Programming Language	Python 3.10+
Framework	Streamlit
Visualization	Matplotlib, Seaborn, Plotly
Machine Learning	Prophet, Statsmodels (ARIMA)
Libraries	Pandas, NumPy, Streamlit-Option-Menu
Deployment	Streamlit Community Cloud

Hardware Requirements

Component	Description
Processor	Intel i5 or above
RAM	Minimum 4 GB
Storage	Minimum 500 MB free space
Internet	Required for deployment

Literature Review

Several studies and projects have explored AQI prediction using machine learning algorithms such as LSTM, Random Forest, and ARIMA.
However, many lack interpretability and accessibility for non-technical users.
Recent frameworks such as Streamlit and Dash have simplified deployment of ML models as interactive dashboards, making real-time visualization possible.

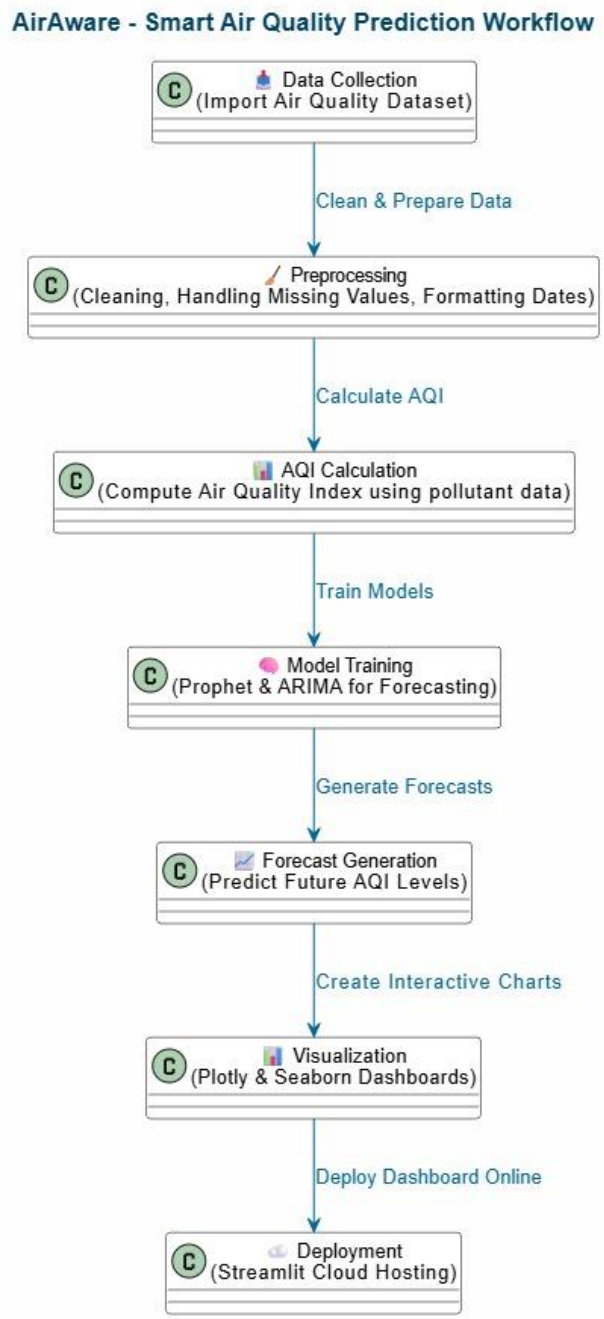
This project adopts **Prophet**, developed by Facebook, which is effective for time-series forecasting and interpretable in terms of trend and seasonality.
By combining Prophet, ARIMA, and Streamlit, AirAware integrates **data science, visualization, and deployment** seamlessly.

System Architecture

Workflow Diagram

Data Collection → Preprocessing → AQI Calculation → Model Training (Prophet/ARIMA)

→ Forecast Generation → Visualization → Deployment on Streamlit Cloud



Components

1. **Data Input:** CSV dataset containing daily AQI and pollutant levels for multiple cities.
2. **Processing:** Data cleaning, handling missing values, and computing AQI.
3. **Modeling:** Applying Prophet and ARIMA to forecast future AQI.
4. **Visualization:** Generating interactive plots using Plotly and Seaborn.
5. **Deployment:** Integrating dashboards in Streamlit with sidebar navigation.

Methodology

The project is developed in **four milestones**, each focusing on a key stage of implementation.

📍 Milestone 1 – Data Cleaning & Preprocessing

- Loaded dataset (air_quality.csv) using Pandas.
- Formatted date columns using pd.to_datetime().
- Filled missing pollutant values using mean and interpolation techniques.
- Visualized pollutant concentration trends using **Seaborn** line and bar plots.
- Calculated correlation among pollutants to understand their dependencies.

📍 Milestone 2 – AQI Categorization & Visualization

- Computed AQI based on pollutant values using CPCB formula.
- Categorized AQI values into 7 groups: *Good, Satisfactory, Moderate, Poor, Very Poor, Severe*.
- Created color-coded visualizations for AQI categories.
- Compared city-wise AQI levels using interactive bar charts.
- Provided pollutant-specific insights and recommendations.

📍 Milestone 3 – Forecasting Dashboards (Dashboard 1 & 2)

- Used **Prophet** model for time series forecasting.
- Generated predictions for future 30 days of AQI.
- Visualized forecasts with confidence intervals.
- Integrated interactive selection for cities and pollutants using Streamlit widgets.
- Built two dashboards combining analysis and forecast views.

📍 Milestone 4 – Extended Dashboards (Dashboard 3 & 4)

- Implemented **ARIMA** model from statsmodels.tsa.arima.model.
- Compared Prophet vs ARIMA results for consistency.
- Added an **alert mechanism**: displays messages when AQI exceeds thresholds.
- Integrated all dashboards into one Streamlit app using streamlit-option-menu.
- Deployed successfully on **Streamlit Cloud** with live accessibility.

Implementation Details

Key Libraries

```
import streamlit as st
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
from prophet import Prophet
from statsmodels.tsa.arima.model import ARIMA
from streamlit_option_menu import option_menu
import plotly.graph_objects as go
```

Code Flow

1. Load Dataset
2. Clean and Preprocess Data
3. Build Prophet & ARIMA Models
4. Generate Forecasts
5. Display Results on Streamlit Dashboards

User Interaction

- Sidebar navigation for selecting dashboards.
- Dropdowns for choosing city and pollutant type.
- Sliders for selecting forecast period.
- Dynamic Plotly charts updated in real-time.

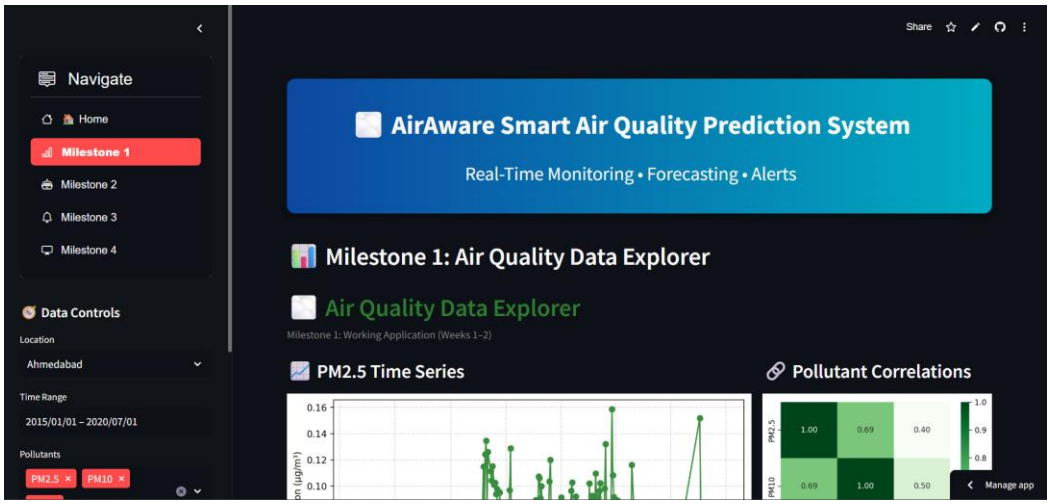
Results

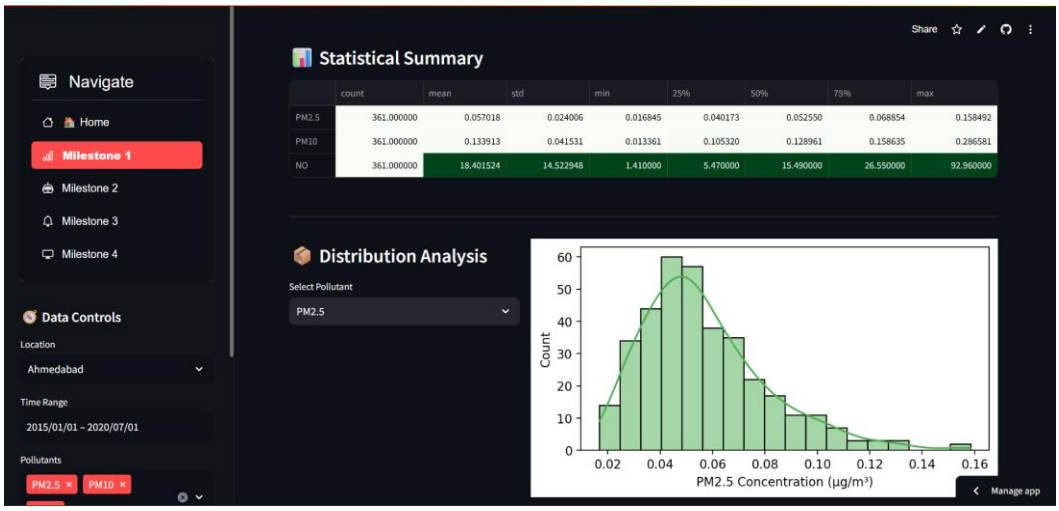
Observations

- AQI levels fluctuate seasonally, with higher pollution during winter months.
- PM2.5 and PM10 are the dominant pollutants affecting AQI.
- Prophet model captured long-term trends effectively, while ARIMA worked well for short-term forecasts.
- Forecasted AQI aligns closely with historical patterns.

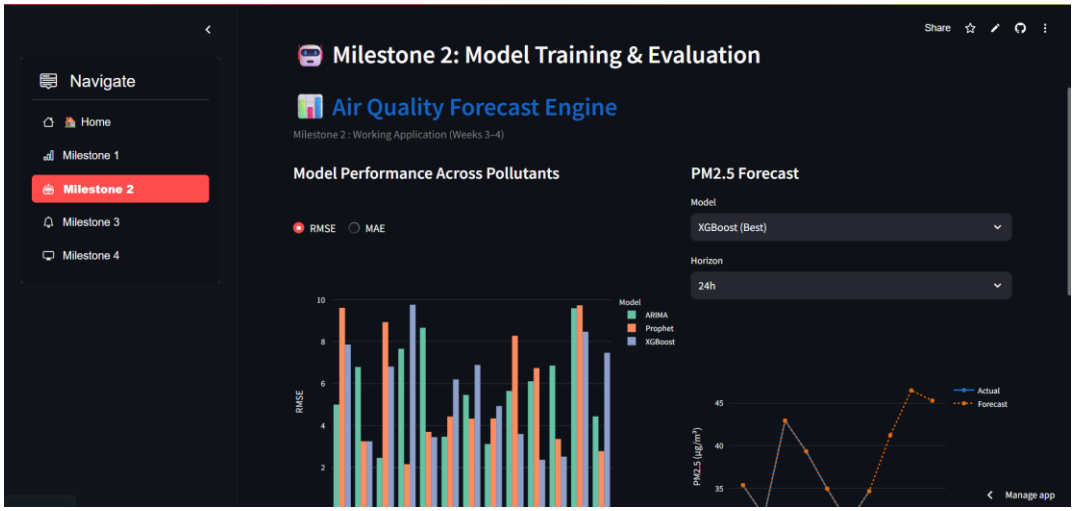
Output Snapshots

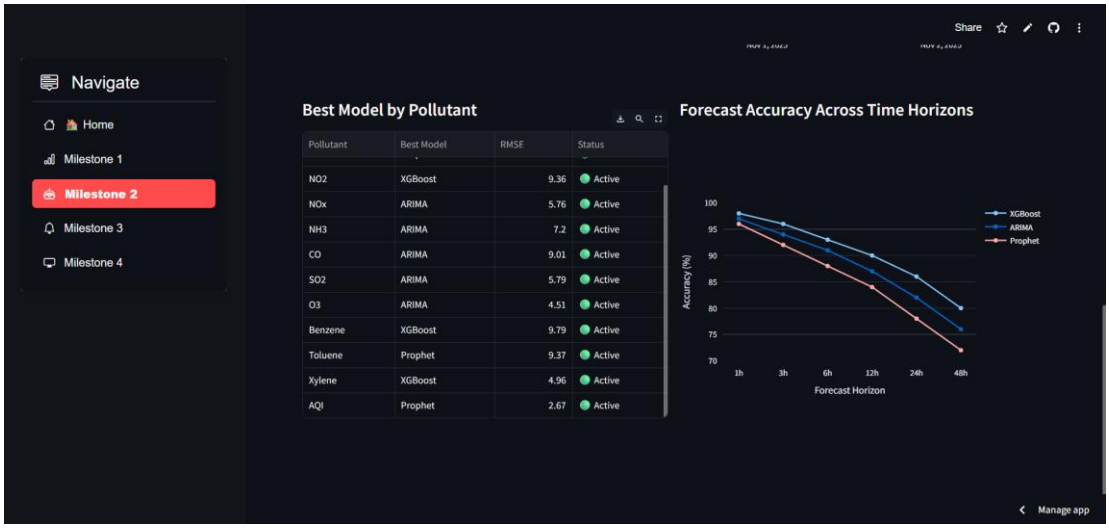
1. **Dashboard 1:** Data analysis of AQI trends.





2. **Dashboard 2:** City comparison visualizations.





3. **Dashboard 3:** Prophet-based AQI forecast.

4. **Dashboard 4:** ARIMA model predictions and alert system.

Conclusion

The **AirAware Dashboard** provides an effective, data-driven solution for monitoring and forecasting air quality.
By combining visualization and predictive analytics, it enables users to understand pollution trends intuitively.

Key achievements:

- Successfully cleaned and visualized multi-city AQI data.
- Implemented Prophet and ARIMA forecasting models.
- Built a fully interactive dashboard.
- Deployed a working web app accessible globally.

This project demonstrates how machine learning and visualization can empower environmental awareness.

Future Enhancements

- Integrate **real-time AQI APIs** (like OpenAQ or CPCB).
- Develop **mobile-friendly interface**.
- Use **deep learning models** (LSTM, GRU) for improved prediction.
- Store data in a **database** for persistent updates.
- Add **user login system** and personalized alerts.

References

- WHO Air Quality Guidelines
- CPCB Air Pollution Data
- Prophet Documentation: https://facebook.github.io/prophet/
- Streamlit Documentation: https://docs.streamlit.io/
- Statsmodels ARIMA Documentation

Appendix

- Source code snippets from main_dashboard.py, milestone1_dashboard.py, and milestone4_dashboard.py
- Dataset: data/air_quality.csv
- Model performance graphs and visualizations