Infosys Springboard Virtual Internship 6.0
Internship Completion Report

Team Details

(Do not mention any personally identifiable information like email ID, institute details, mobile phone number, etc.)

Batch Number: AirAware: Smart Air Quality Prediction System
Start Date: 10 November 2025

Team Members:

1. Sahil Choudhary
2. Raunit Singh
3. Sneha Reddy
4. Likhita Karri
5. Fariha Naureen

Internship Duration: 40 Days (8 Weeks)

1. Project Title

AirAware: Smart Air Quality Prediction System

---

2. Project Objective

AirAware is an AI-driven air quality monitoring and prediction platform aimed at addressing the growing environmental and public health challenges caused by air pollution.

The primary objectives of this project are to:

- Monitor critical air quality parameters such as AQI, PM2.5, PM10, CO, NO_2, and SO_2 across multiple cities.
- Analyze historical and real-time air pollution data to identify long-term trends and seasonal patterns.
- Predict future air quality levels using machine learning models to enable early warnings.
- Visualize air pollution data through intuitive dashboards and graphical representations.
- Enhance public awareness and decision-making by classifying air quality into health-based categories.
- Support preventive actions for individuals, institutions, and policymakers through data-driven insights.

Overall, the project aims to convert complex environmental data into actionable intelligence using artificial intelligence and data analytics.

---

3. Project Description (Detailed)

AirAware is a smart, web-based air quality prediction system designed to analyze environmental pollution data and provide both current and future air quality insights.

The system works through the following pipeline:

1. Collection of air quality data from reliable datasets and APIs covering multiple pollutants and cities.
2. Preprocessing of raw data, including handling missing values, normalization, and noise reduction.
3. Exploratory data analysis to understand pollution patterns and correlations between pollutants.
4. Feature engineering to identify the most influential parameters affecting AQI.
5. Training of supervised machine learning models using historical pollutant concentration data.
6. Prediction of AQI values for upcoming time periods based on learned patterns.
7. Visualization of current, historical, and predicted air quality using charts, graphs, and

dashboards.

8. Classification of AQI levels into standard health categories such as Good, Moderate, Poor, and Hazardous.

The system serves as both a monitoring and forecasting tool, enabling users to proactively respond to deteriorating air quality conditions.

---

4. Weekly Activity Plan and Completion

Week	Planned Activities	Completed Activities
Week 1	Requirement gathering, environmental study, problem analysis	Clearly defined problem statement, objectives, and system scope
Week 2	Dataset identification and feasibility analysis	Selected validated air quality datasets and APIs
Week 3	Data preprocessing and cleaning	Cleaned datasets, handled missing values, and standardized formats
Week 4	Feature selection and exploratory data analysis	Identified key pollutant features influencing AQI
Week 5	Machine learning model design and training	Implemented AQI prediction model
Week 6	Model evaluation and optimization	Improved prediction accuracy through tuning
Week 7	Frontend dashboard and visualization integration	Developed visual dashboards and graphs
Week 8	Documentation, testing, and final review	Finalized project and prepared reports

5a. Key Milestones

Milestone	Description	Date Achieved
Project Kickoff	Defined project goals, scope, datasets, and responsibilities	10 Nov 2025
Data Readiness	Completed data preprocessing and validation	20 Nov 2025
Model Development	Built and tested AQI prediction model	1 Dec 2025
System Integration	Integrated ML model with visualization layer	15 Dec 2025
Final Submission	Completed documentation and evaluation	05 Jan 2025

5b. Project Execution Details

The execution of the AirAware project followed a systematic and iterative development approach.

Initially, extensive research was conducted on air pollution indicators, AQI standards, and environmental monitoring systems. Relevant datasets were collected and examined to ensure accuracy and completeness.

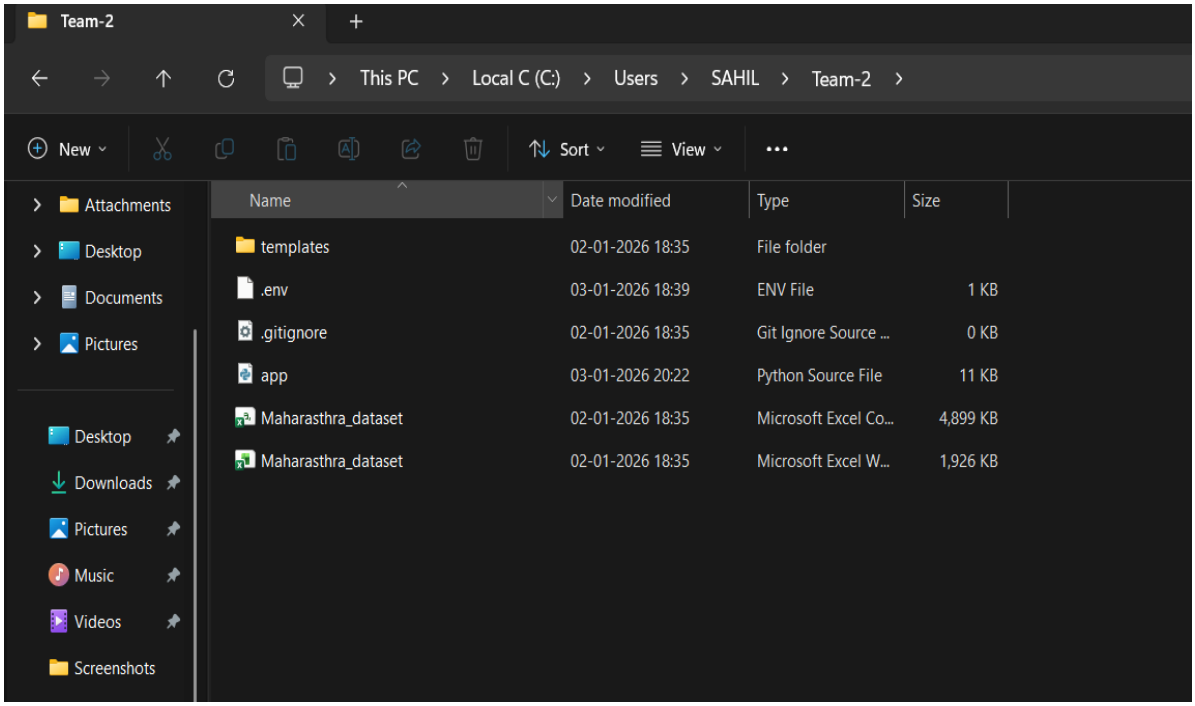
Data preprocessing techniques such as missing value imputation, normalization, and feature scaling were applied to improve data quality. Exploratory data analysis helped identify pollutant correlations and seasonal trends.

Machine learning models were trained using historical pollutant data, enabling the system to predict future AQI levels. The trained model was validated and optimized to ensure reliable and meaningful predictions.

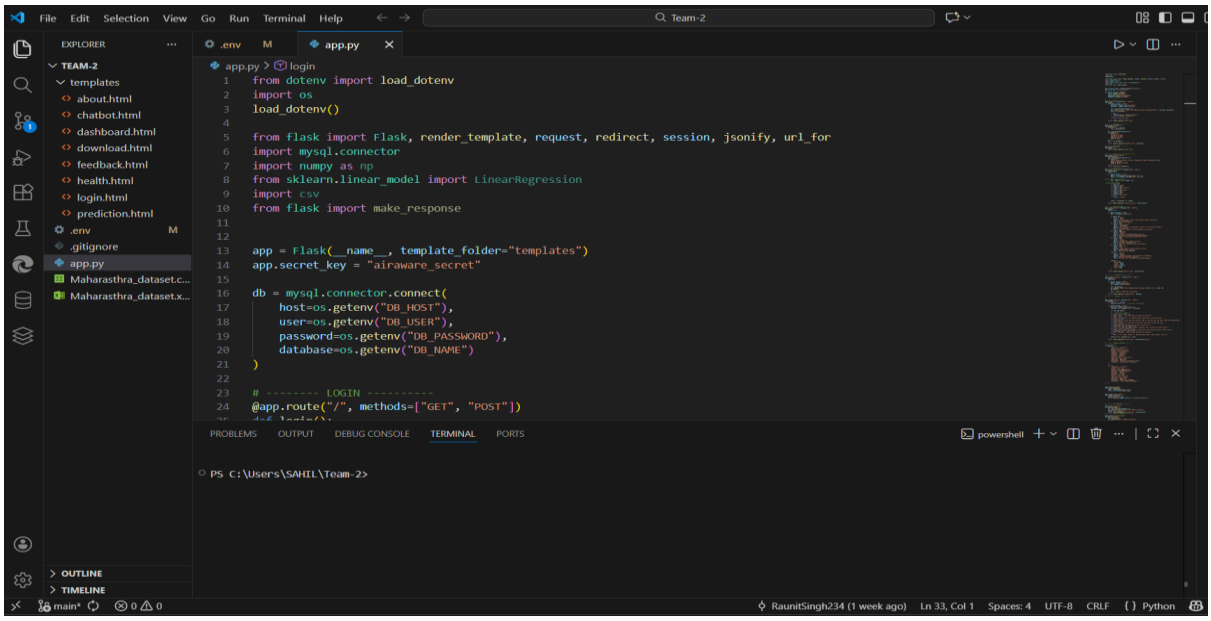
Simultaneously, a user-friendly web interface was developed to visualize air quality metrics through interactive charts and dashboards. Continuous testing and debugging ensured smooth integration and stable system performance.

6. Snapshots / Screenshots

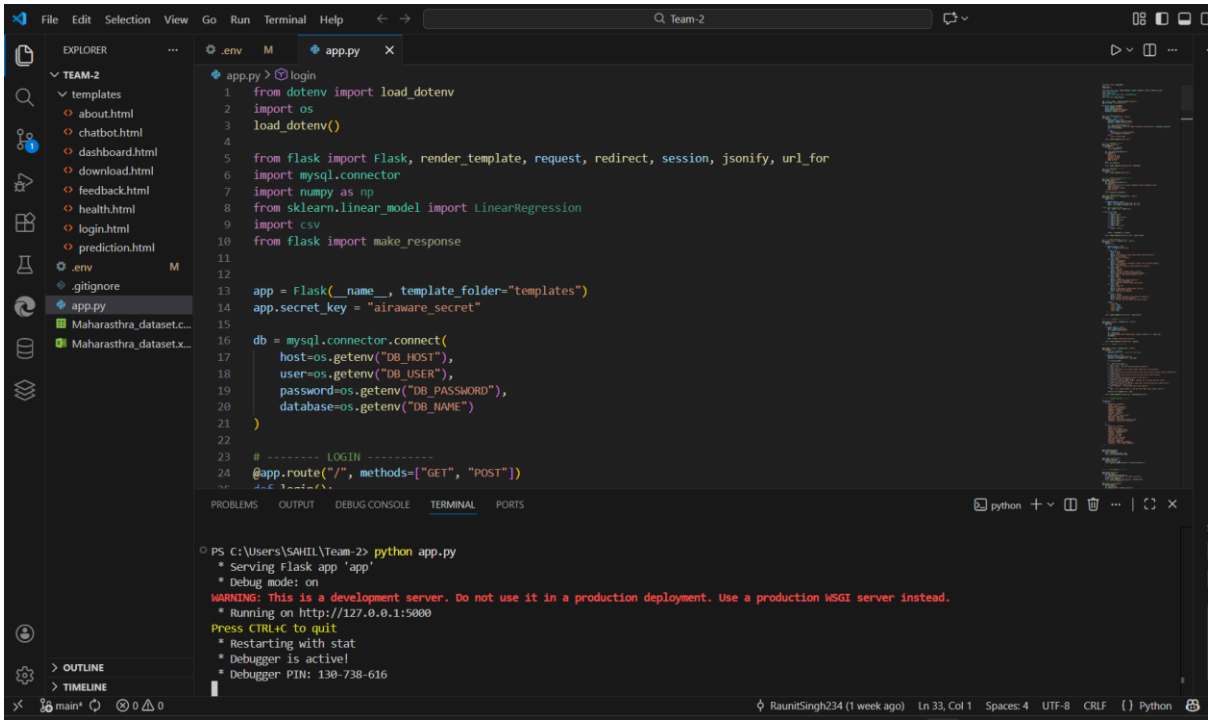
Step 1: Opening the project folder containing all the main files such as app.py, templates,



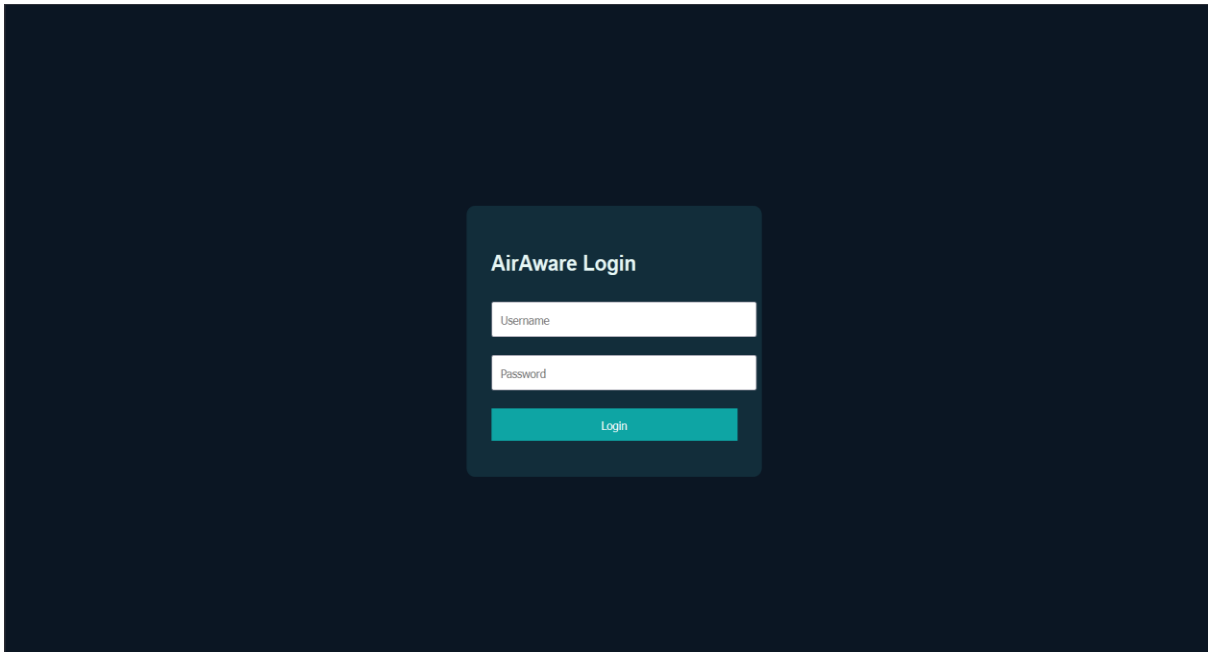
Step 2: Opening the app.py file in VS Code to view the main Flask application code.



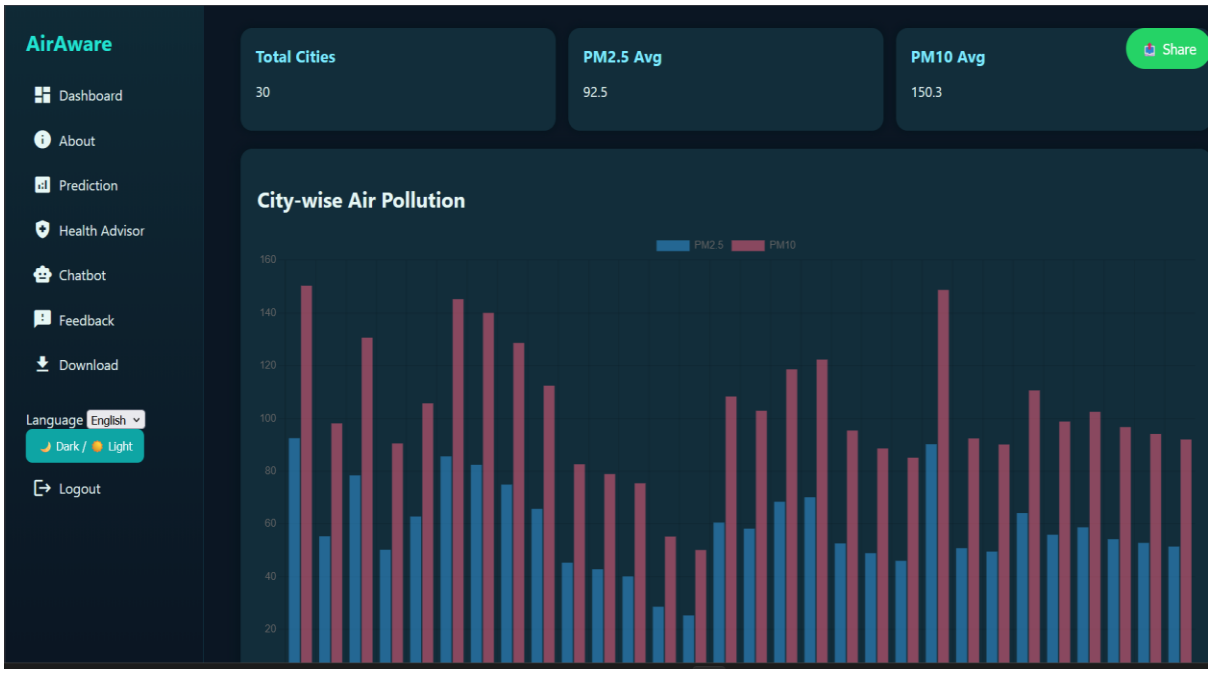
Step 3: Running the Flask app using the command python app.py in the terminal.



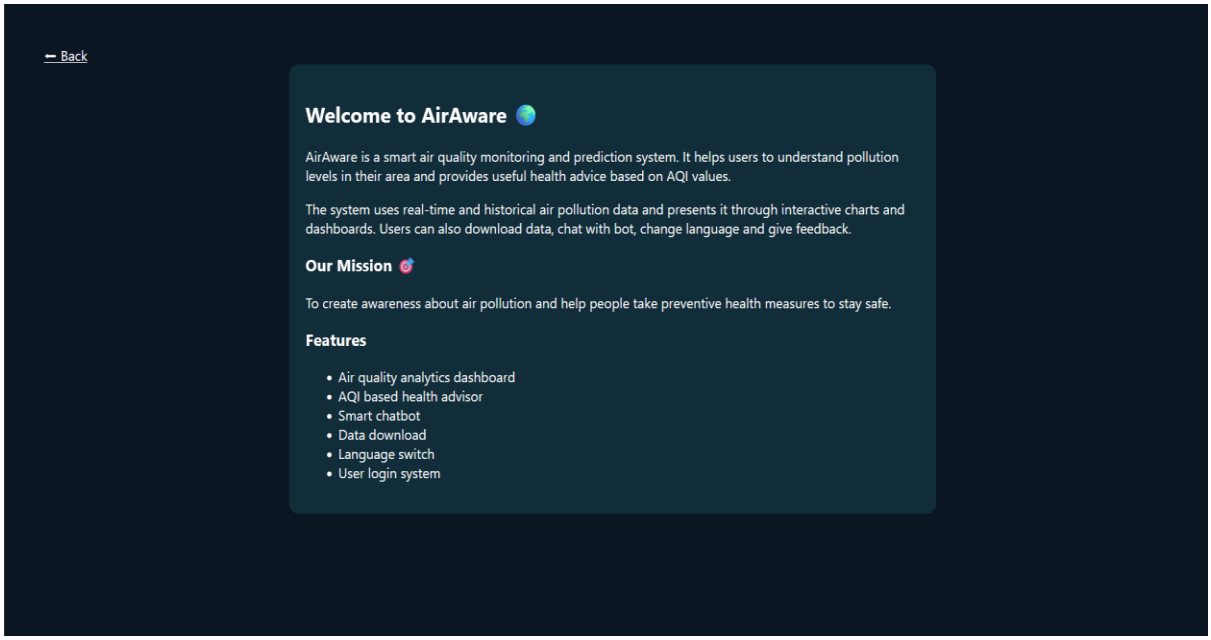
Step 4: Login Page



Step 5: Dashboard: AirAware

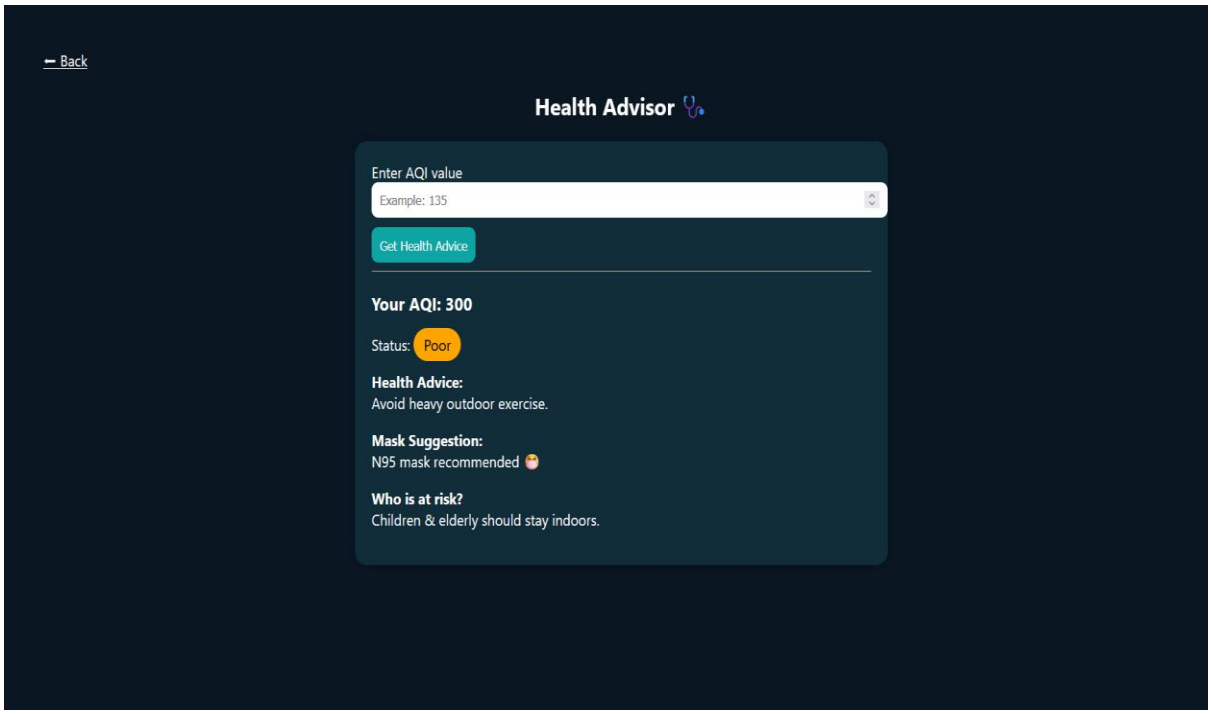


Step 6: AirAware about page

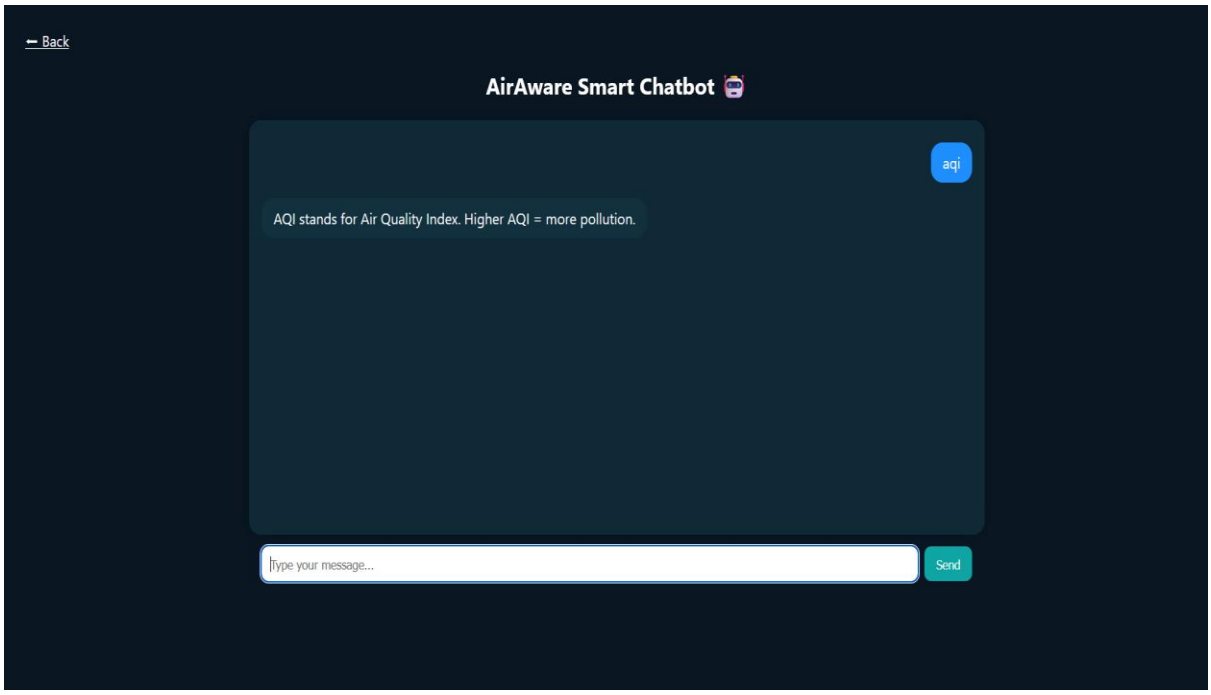


Step 7: Prediction

Step 8: Model gives Health Advisory according the AQI



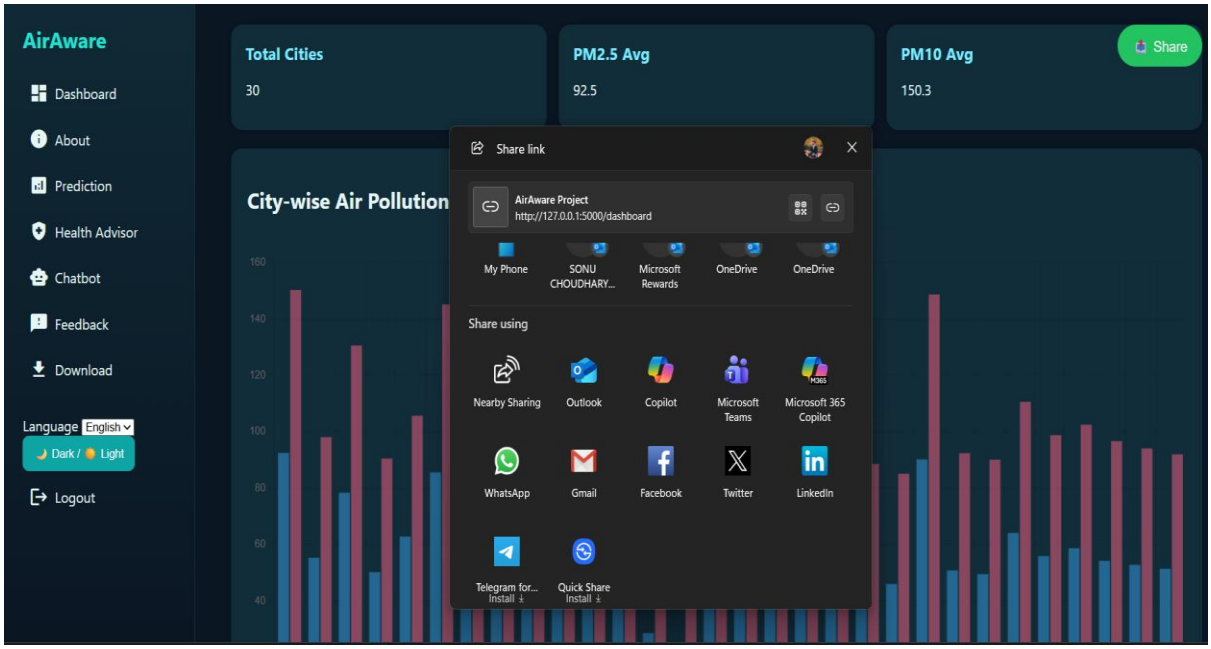
Step 9: Chatbot



Step 10: Feedback

Step 11: User can switch to Dark/Light mode

Step 12: User can also share the data via different platforms



7. Challenges Faced

During the development of AirAware, several challenges were encountered:

1. Variability in air quality data formats across different sources.
2. Missing or incomplete pollution data affecting model accuracy.
3. Identifying the most impactful features influencing AQI.
4. Balancing model complexity with prediction performance.
5. Visualizing multidimensional environmental data effectively.
6. Managing large datasets efficiently.
7. Aligning predictions with official AQI classification standards.

These challenges were resolved through iterative refinement, testing, and optimization.

---

8. Learnings & Skills Acquired

The project resulted in comprehensive technical and professional learning:

1. In-depth understanding of air quality metrics and environmental standards.
2. Practical experience in data preprocessing and exploratory data analysis.
3. Hands-on exposure to machine learning-based predictive modelling.
4. Visualization techniques for environmental data interpretation.
5. System integration between ML models and frontend dashboards.
6. Debugging, optimization, and performance tuning.
7. Technical documentation and structured reporting skills.

The project significantly strengthened problem-solving abilities and AI application skills.

---

9. Testimonials from Team

Team Member 1: "This project enhanced my understanding of AI-driven environmental analysis and strengthened my skills in machine learning, data visualization, and system design."

Team Member 2: "AirAware provided valuable hands-on experience with real-world datasets and helped me understand predictive modelling for environmental monitoring."

Team Member 3: "This project improved my collaborative skills and deepened my knowledge of data preprocessing and analytical visualization."

Team Member 4: "Working on AirAware boosted my confidence in applying machine learning to solve real-world environmental challenges."

Team Member 5: "This internship strengthened my understanding of end-to-end project development and enhanced my teamwork and problem-solving abilities."

---

10. Conclusion

AirAware: Smart Air Quality Prediction System successfully demonstrates the effective application of artificial intelligence and data analytics in addressing environmental challenges. The project delivers meaningful insights into air pollution trends, predicts future AQI levels, and supports proactive decision-making. It highlights strong teamwork, technical expertise, and the practical use of AI for social and environmental impact.

---

11. Acknowledgements

We sincerely thank **Infosys Springboard** for providing an excellent virtual internship platform and industry-oriented learning resources.

We extend our heartfelt gratitude to our mentor **Mrs. Shakthi Gopalakrishnan** for her continuous guidance, valuable feedback, and encouragement throughout the internship.

We also thank our peers and team members for their collaboration and support, which contributed significantly to the successful completion of the AirAware project.