International Journal of Advance and Innovative Research
Volume 13, Issue 1 (IV): January - March 2026
ISSN 2394 - 7780

JVM's Mehta Degree College, Sector 19, Airoli
NAAC Re-accredited "A+" Grade
IQAC in association with Western Regional Centre, ICSSR Organized one day National Conference on "Integrating Multidisciplinary Approaches to Build a Resilient and Sustainable Future", held on 10th January 2026

AIR POLLUTION PREDICTION SYSTEM USING AI-BASED TECHNIQUES

Dr. Sunitha Joshi
Assistant Professor, Department of Information Technology, JVM Mehta Degree College

ABSTRACT
Air pollution is a major environmental and public health concern, requiring accurate and real-time monitoring for effective control and mitigation. Traditional air quality monitoring systems rely on expensive and geographically limited sensor networks, which fail to provide large-scale real-time data. This project explores the application of Machine Learning (ML) techniques for air pollution detection and prediction, leveraging historical pollution data, meteorological parameters, and IoT-based sensor inputs. Various ML algorithms, including Support Vector Machines (SVMs), Decision Trees, Random Forest, and Deep Learning models, are employed to analyze pollutant concentration levels and predict future trends. The proposed system enhances air quality monitoring by offering cost-effective, scalable, and real-time predictive insights to assist policymakers, environmental agencies, and the general public in taking timely preventive measures. By integrating ML with real-time data collection, this approach significantly improves the accuracy and efficiency of air pollution detection, contributing to better environmental management and public health protection.

Keywords: *Air Pollution, Machine Learning, Air Quality Index, Prediction Models, IoT Sensors, Random Forest, Environmental Analytics*

1. INTRODUCTION
Air pollution has become one of the most critical environmental challenges worldwide, particularly in urban regions where vehicular emissions, industrial activities, and population density are high. According to global health studies, prolonged exposure to polluted air leads to respiratory diseases, cardiovascular disorders, and premature mortality. The Air Quality Index (AQI) serves as a standardized metric to communicate pollution severity and associated health risks to the public.

Conventional air quality monitoring systems depend on fixed sensor stations that are costly to install and maintain, resulting in limited spatial coverage and delayed reporting. These limitations reduce the effectiveness of pollution mitigation and public awareness. Recent advancements in artificial intelligence and machine learning have enabled data-driven approaches that can model complex, non-linear relationships among pollutants and environmental factors.

This research presents an AI-based air pollution prediction system that integrates historical data, real-time inputs, and machine learning algorithms to provide accurate AQI predictions and health advisories. The system supports real-time monitoring, hourly forecasting, and personalized health impact assessment through an intuitive web interface. By offering scalable and predictive air quality insights, the proposed solution contributes to proactive environmental management and informed public decision-making.

2. LITERATURE SURVEY

2.1 Background
Air quality monitoring has evolved from manual sampling techniques to automated sensor networks and data-driven prediction models. The introduction of the Air Quality Index standardized pollution reporting and enhanced public communication. With the growth of computational power and data availability, machine learning techniques have increasingly been adopted to improve prediction accuracy and real-time responsiveness.

2.2 Related Work

Rahman (2024) proposed *AirNet*, a machine learning-based air quality forecasting system using web-based interfaces. The study demonstrated that Random Forest and Decision Tree models achieved high accuracy when trained on large-scale global datasets, highlighting the effectiveness of ensemble learning techniques in pollution prediction.

Shriram (2021) examined the relationship between AQI and public health, proposing a system that integrates real-time AQI data with health impact analysis and route optimization to reduce exposure. The research emphasized the importance of personalized pollution-aware decision systems.

Yasmin (2023) introduced *AQIPred*, a hybrid MLP–LSTM model for time-specific AQI forecasting. The model achieved high precision and demonstrated the superiority of hybrid deep learning architectures for temporal air quality prediction.

Kumar (2022) conducted a case study on Indian cities using multiple machine learning models to predict air pollution trends. The study found that probabilistic and ensemble models outperformed traditional statistical approaches, particularly when handling diverse urban datasets.

These studies collectively validate the application of machine learning in air pollution prediction and motivate the development of a real-time, user-centric AQI forecasting system.

3. PROBLEM STATEMENT

Air pollution significantly impacts human health and environmental sustainability, yet existing monitoring systems suffer from delayed reporting, limited prediction horizons, and lack of personalized guidance. Most systems provide generalized AQI values without actionable health recommendations or real-time alerts tailored to individual users.

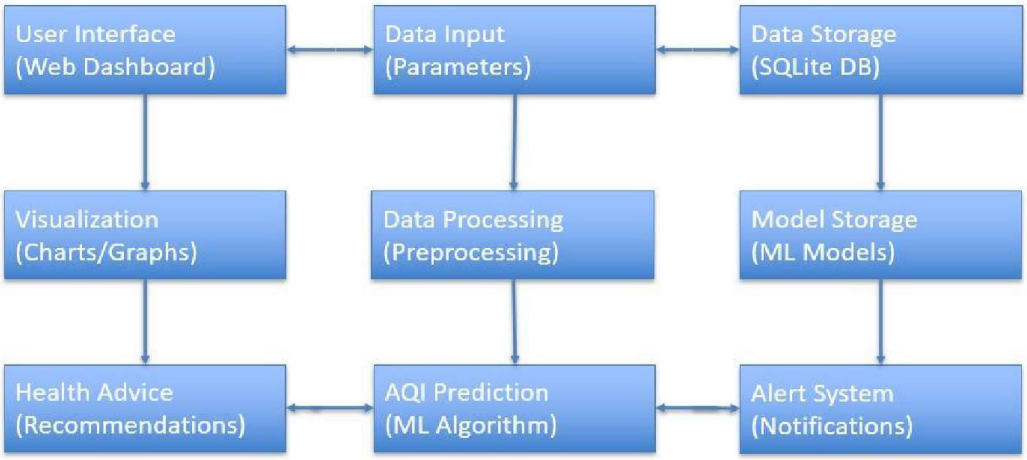
The problem addressed in this research is the need for a **real-time, accurate, and personalized air pollution prediction system** that leverages machine learning to deliver proactive alerts, health impact analysis, and future AQI forecasts.

4. PROPOSED SYSTEM AND IMPLEMENTATION

4.1 System Overview

The proposed system integrates real-time data acquisition, machine learning-based AQI prediction, visualization modules, and a health impact calculator. It processes pollutant parameters such as PM2.5, PM10, NO_2, SO_2, CO, and O_3 to generate current and forecasted AQI values.

4.2 Block Diagram of Proposed System



The architecture consists of:

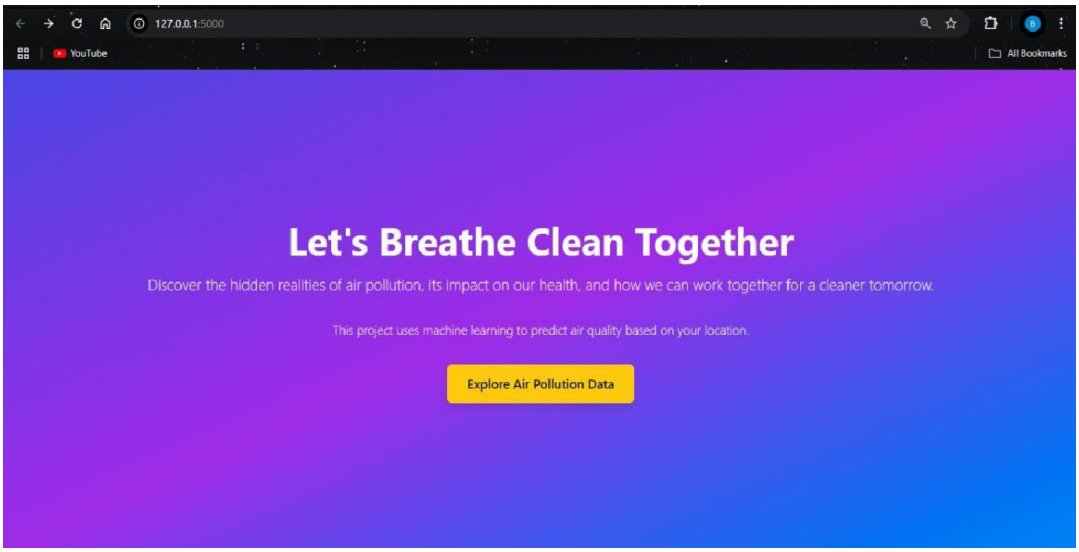
- User Interface (Web Dashboard)
- Data Input via APIs and sensors
- Data Storage
- Data Preprocessing Module

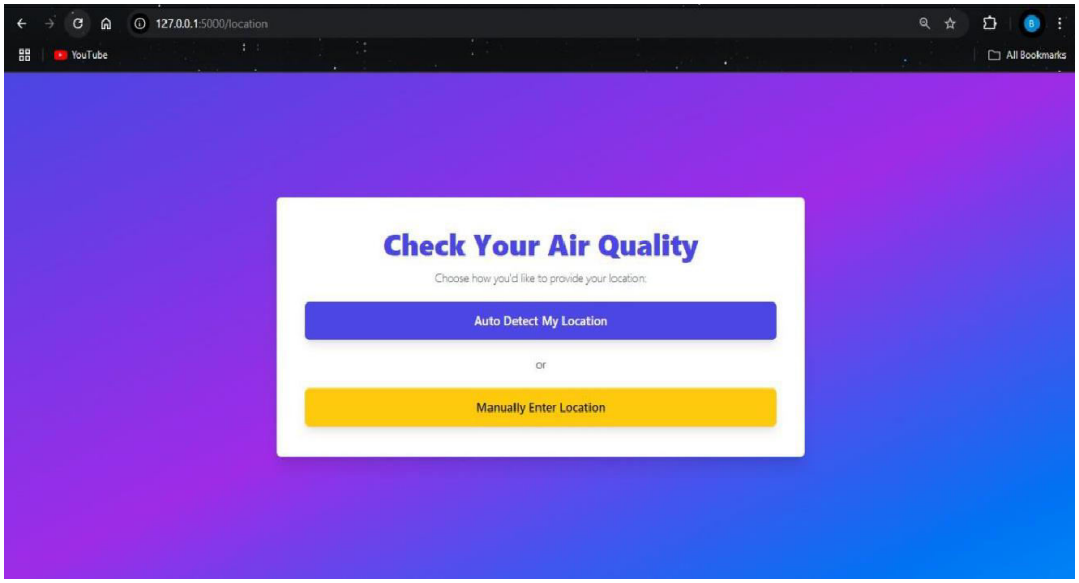
- Machine Learning Model
- AQI Prediction Engine
- Health Recommendation and Alert System

4.3 IMPLEMENTATION

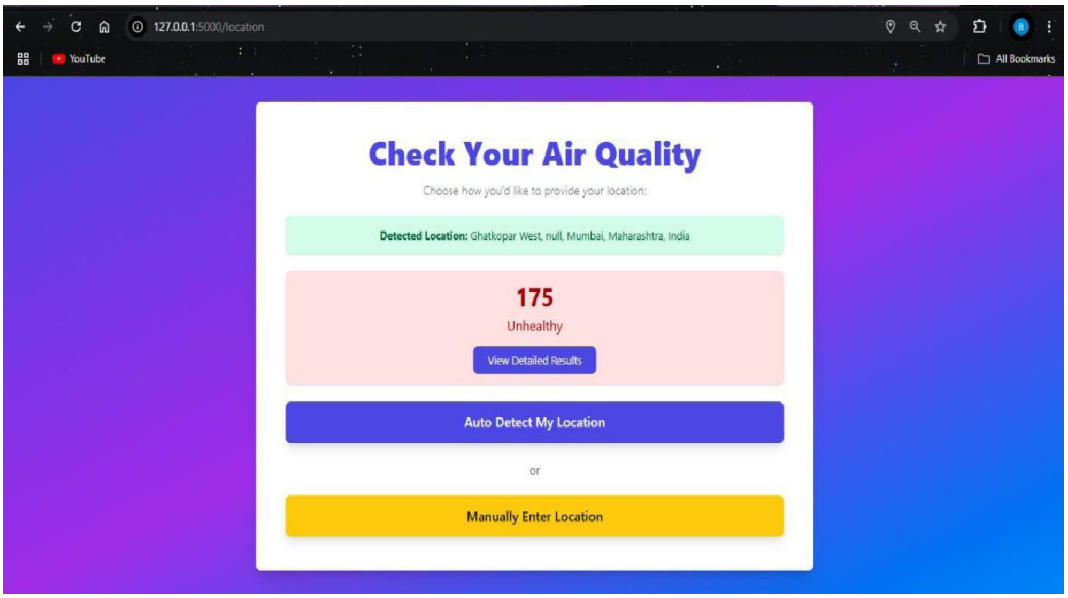
The system is implemented using Python-based technologies, with Flask used for web deployment and Scikit-learn for machine learning. A Random Forest Regressor model is trained on historical air quality datasets to predict AQI values. The application provides:

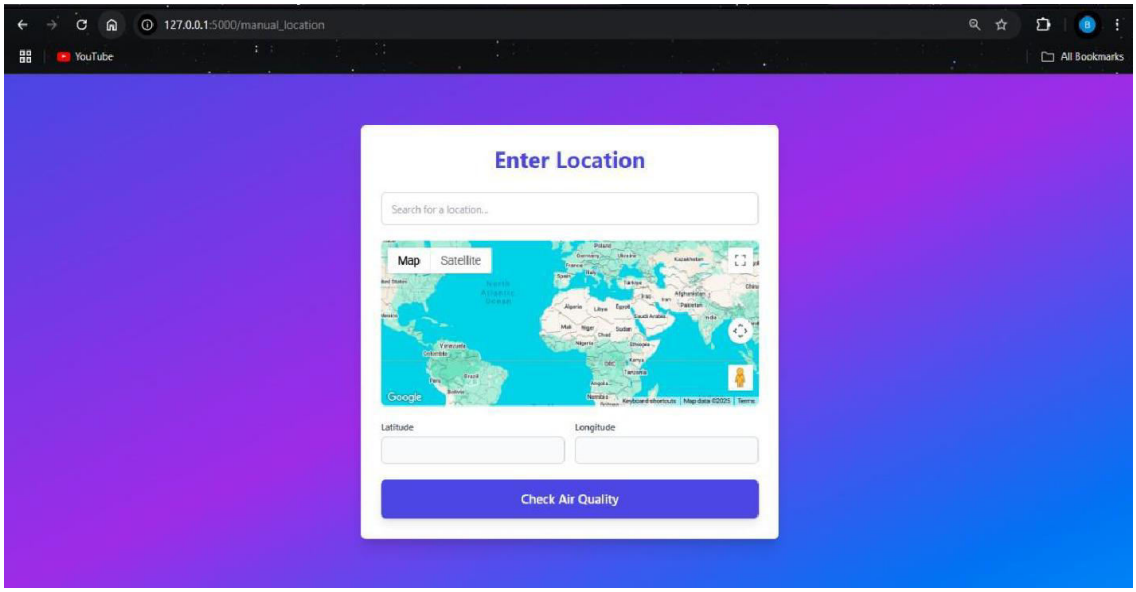
- Automatic and manual location-based AQI detection
- Real-time AQI display with severity classification
- Health impact calculator based on age and health conditions
- Email alerts for hazardous pollution levels

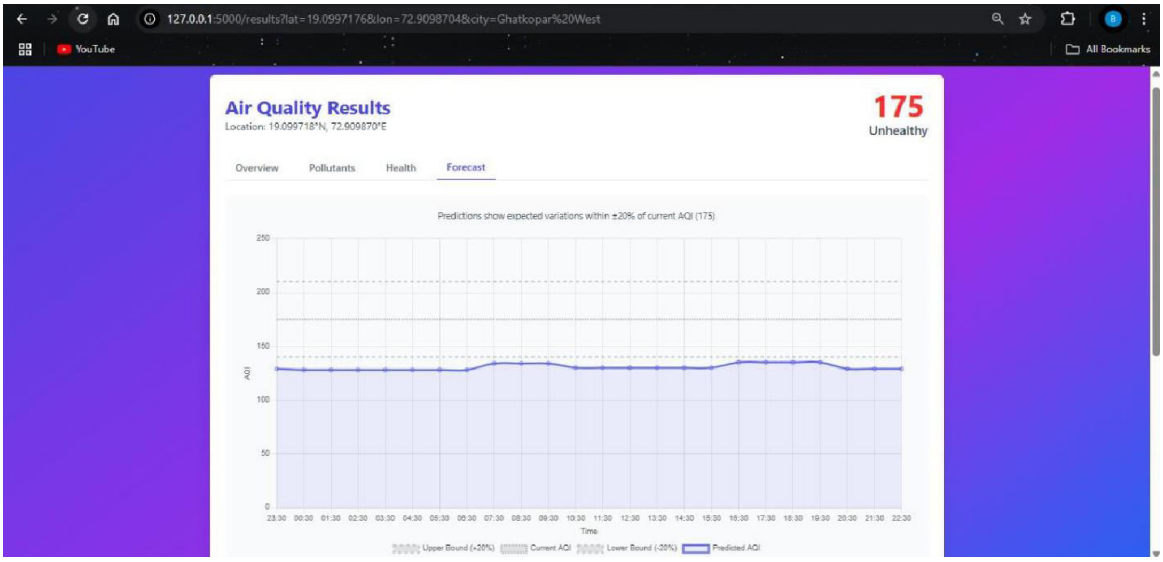




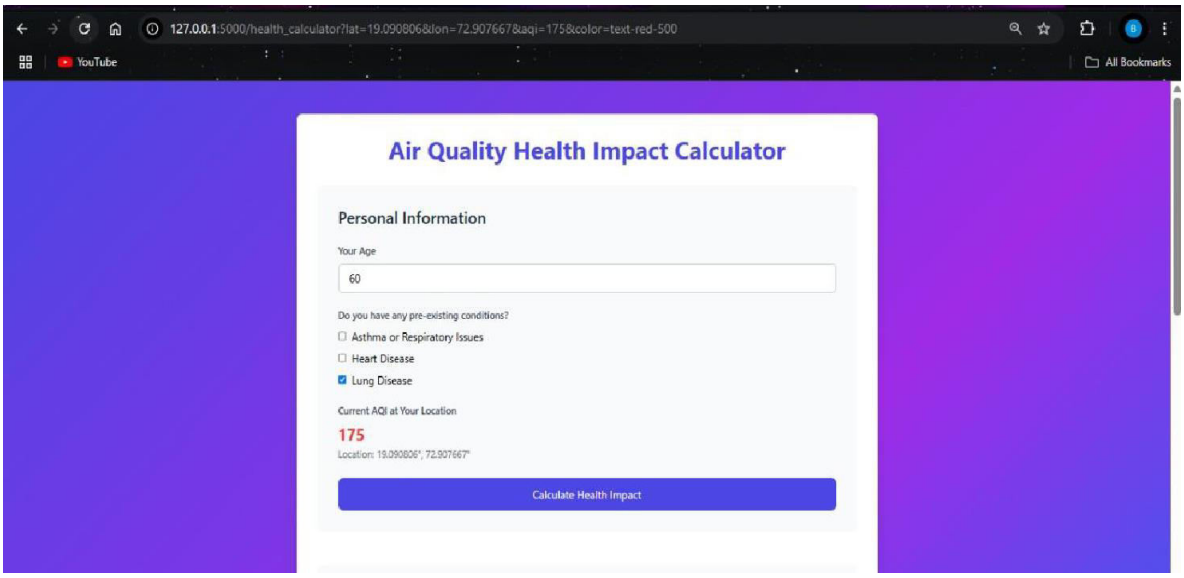
International Journal of Advance and Innovative Research
Volume 13, Issue 1 (IV): January - March 2026
ISSN 2394 - 7780
JVM's Mehta Degree College, Sector 19, Airoli
NAAC Re-accredited "A+" Grade
IQAC in association with Western Regional Centre, ICSSR Organized one day National Conference on "Integrating Multidisciplinary Approaches to Build a Resilient and Sustainable Future", held on 10th January 2026

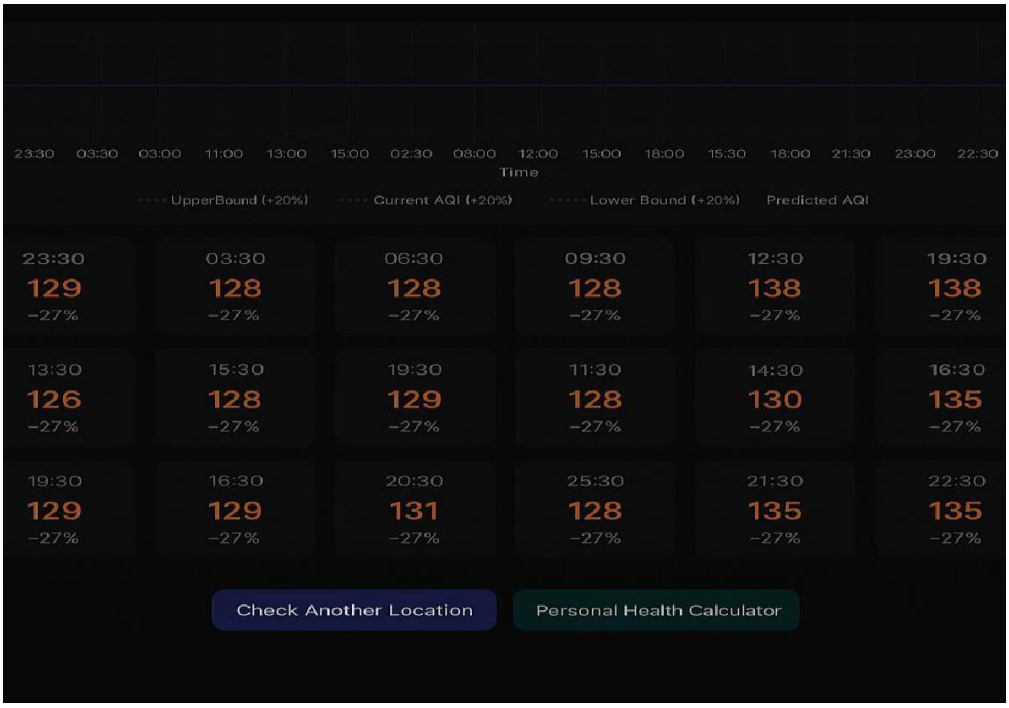






International Journal of Advance and Innovative Research
Volume 13, Issue 1 (IV): January - March 2026
ISSN 2394 - 7780
JVM's Mehta Degree College, Sector 19, Airoli
NAAC Re-accredited "A+" Grade
IQAC in association with Western Regional Centre, ICSSR Organized one day National Conference on "Integrating Multidisciplinary Approaches to Build a Resilient and Sustainable Future", held on 10th January 2026





5. RESULTS AND DISCUSSION

The system demonstrates high prediction accuracy and real-time responsiveness. AQI predictions show strong alignment with observed pollution trends, enabling early detection of unhealthy air conditions.

Performance Highlights:

- Model Accuracy: ~95% using Random Forest
- Response Time: ~2 seconds
- Data Sources: Open APIs and IoT sensors
- Forecasting Capability: Hourly AQI prediction

The health impact calculator effectively translates AQI values into actionable guidance, enhancing usability for sensitive populations.

6. CONCLUSION

The Air Pollution Prediction System successfully overcomes the limitations of traditional air quality monitoring by integrating machine learning-based prediction, real-time AQI reporting, and personalized health advisories. The system provides accurate forecasts, timely alerts, and user-friendly visualization, enabling proactive decision-making and improved public awareness. Its scalable architecture allows future enhancements such as mobile application integration and advanced weather-based forecasting, making it a valuable tool in combating air pollution and protecting public health.

REFERENCES

[1] Md. Mahbubur Rahman (2024), AirNet: Predictive Machine Learning Model for Air Quality Forecasting Using Web Interface.

[2] Pranav Shriram (2021), A Study and Analysis of Air Quality Index and Related Health Impact on Public Health.

[3] Farhana Yasmin (2023), AQIPred: A Hybrid Model for High Precision Time-Specific Forecasting of Air Quality Index with Cluster Analysis.

[4] K. Kumar (2022), Air Pollution Prediction with Machine Learning: A Case Study of Indian Cities.