URBAN AMBIENT AIR QUALITY INDEX CLASSIFICATION USING AIR QUALITY INDEX CLASSIFICATION USING MULTI-POLLUTANT SENSOR READINGS

[1]B.Vijay kumar,[2]K.Venkatesh,[3]N.Bunny,[4]P.Ganesh

[1]Assistant Professor, [234]Students

Department of Computer Science and Engineering

Siddhartha institute of technology & sciences,narapally

vijaybukhya@siddartha.org.in,23TQ1A0515@siddhartha.co.in, 23TQ1A0542@siddhartha.co.in, 23TQ1A0543@siddhartha.co.in

ABSTRACT

Urban air pollution has become a major environmental and public health concern due to rapid industrialization, urbanization, and increased vehicular emissions. Monitoring and assessing air quality in real-time is essential to reduce health risks and improve environmental conditions. Traditional air quality monitoring systems are limited by sparse station coverage, high costs, and delayed data reporting.

This project proposes an intelligent system for **Air Quality Index (AQI) classification** using multi-pollutant sensor readings such as PM2.5, PM10, CO, NO_2, SO_2, and O_3. The system integrates IoT-based sensors, data preprocessing techniques, and machine learning algorithms to classify air quality into categories like Good, Moderate, Unhealthy, and Hazardous.

By analyzing real-time sensor data, the system provides accurate and timely AQI predictions. It also generates visual dashboards and alerts to inform the public and authorities. The proposed approach improves monitoring efficiency, enhances data accuracy, and supports decision-making for pollution control strategies.

I INTRODUCTION

Air pollution in urban areas is one of the leading causes of respiratory diseases and environmental degradation. The increasing number of vehicles, industrial emissions, and construction activities significantly contribute to the deterioration of air quality. Monitoring air pollution levels is essential to protect public health and maintain environmental sustainability.

The Air Quality Index (AQI) is a standardized indicator used to represent the level of air pollution and its potential impact on human health. It is calculated based on the concentration of multiple pollutants. Traditional monitoring systems rely on fixed stations that provide limited spatial coverage and often fail to capture real-time variations.

With advancements in sensor technology and data analytics, it is now possible to develop intelligent systems that continuously monitor air quality using multiple sensors. These systems collect pollutant data and use machine learning models to classify AQI levels accurately.

This project focuses on building a smart AQI classification system using multi-pollutant sensor readings. The system aims to provide real-time insights, improve prediction accuracy, and assist authorities in taking preventive measures to control pollution.

II LITERATURE SURVEY

Various studies have explored air quality monitoring and AQI prediction using advanced technologies. Earlier systems depended on manual data collection and basic statistical models, which lacked real-time capabilities and accuracy.

Recent research emphasizes the use of IoT-based sensor networks for continuous air quality monitoring. These systems deploy multiple sensors across urban areas to collect pollutant data in real-time. However, sensor calibration and data reliability remain key challenges.

Machine learning techniques such as Decision Trees, Random Forest, Support Vector Machines (SVM), and Neural Networks have been widely used for AQI classification and prediction. These models analyze complex relationships between pollutants and improve classification accuracy.

Some studies have integrated deep learning models to process large-scale environmental data and provide highly accurate predictions. Additionally, cloud computing and data visualization tools are used to store, process, and display AQI data effectively.

Despite these advancements, challenges such as data noise, sensor errors, and scalability issues still exist. This project aims to overcome these limitations by combining multi-pollutant data with efficient machine learning techniques for accurate AQI classification.

III SYSTEM ANALYSIS

The existing air quality monitoring systems rely mainly on fixed monitoring stations and basic analytical methods. These systems provide limited coverage and are not suitable for real-time analysis.The system ensures continuous monitoring and provides real-time updates through dashboards and alerts. It improves decision-making by providing accurate and timely air quality information. The integration of multiple sensors enhances data reiability and coverage.

Existing system

The existing air quality monitoring systems rely mainly on fixed monitoring stations and basic analytical methods. These systems provide limited coverage and are not suitable for real-time analysis.

DisAdavantages of Existing system

- Limited spatial coverage
- High installation and maintenance cost
- Delayed data updates
- Lack of real-time monitoring
- Less accurate predictions

Proposed system

The proposed system uses IoT sensors and machine learning models to classify AQI levels based on multi-pollutant readings. It collects real-time data, processes it, and provides instant AQI classification.

The system generates alerts and visual dashboards to inform users about air quality conditions. It improves accuracy, reduces costs, and enables large-scale monitoring.

Advantages of Proposed System

- Real-time AQI monitoring
- High accuracy using ML models
- Low-cost sensor deployment
- Wide area coverage
- Automated data processing

IV METHODOLOGY

The methodology begins with collecting air quality data from multiple sensors measuring pollutants such as PM2.5, PM10, CO, NO_2, SO_2, and O_3. The collected data is then preprocessed to remove noise and handle missing values.

Next, feature extraction is performed to identify important pollutant parameters. Machine learning models are trained using labeled AQI datasets to classify air quality levels.

The trained model analyzes real-time data and predicts AQI categories. The results are visualized through dashboards and alerts for easy interpretation. This approach ensures efficient and accurate air quality classification.

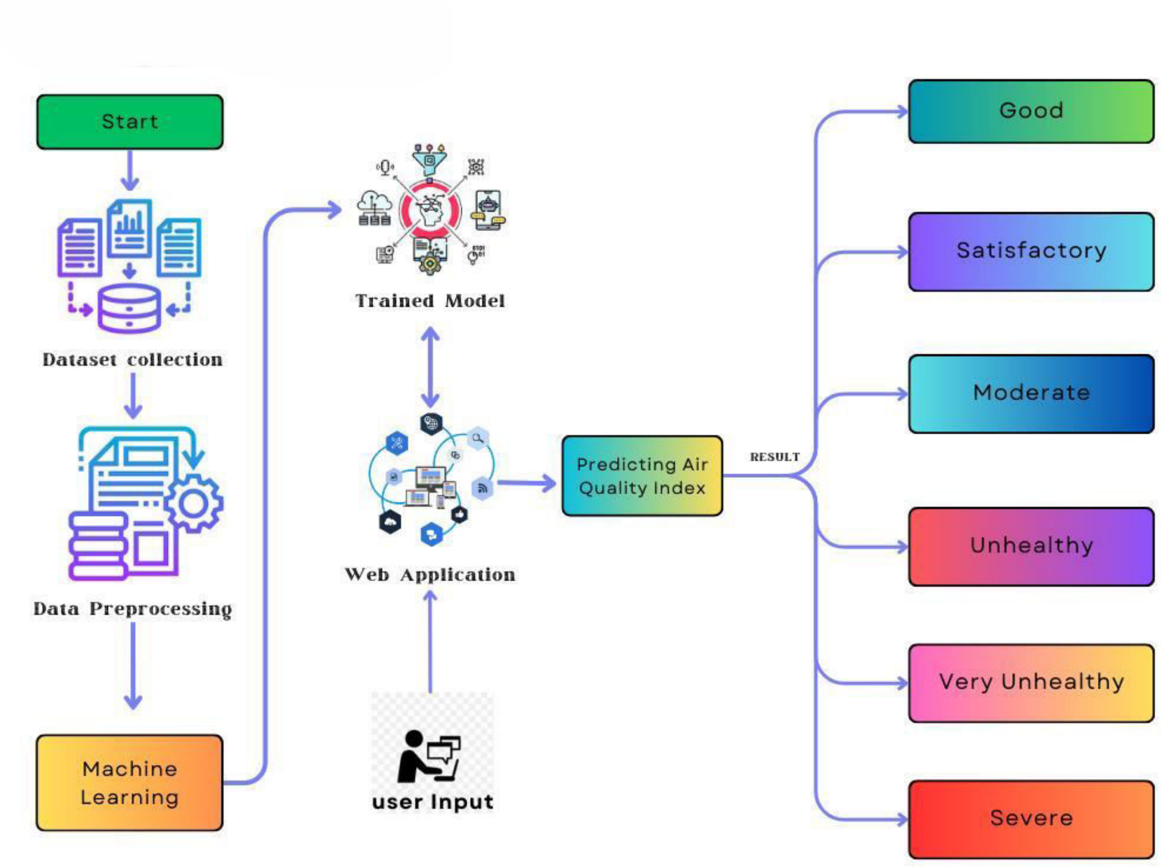
The proposed system for **Urban Ambient Air Quality Index (AQI) Classification Using Multi-Pollutant Sensor Readings** follows a structured and systematic approach that combines sensor data collection, preprocessing, machine learning, and real-time visualization. The methodology is designed to ensure accurate, continuous, and scalable monitoring of air quality in urban environments.

The first phase involves **data acquisition** using multiple air quality sensors deployed across different urban locations. These sensors continuously measure pollutant concentrations such as PM2.5, PM10, Carbon Monoxide (CO), Nitrogen Dioxide (NO_2), Sulfur Dioxide (SO_2), and Ozone (O_3). The sensors may be IoT-enabled, allowing real-time transmission of data to a central server or cloud platform. In addition to sensor data, historical AQI datasets may also be collected for training machine learning models.

Once the data is collected, it undergoes **data preprocessing**, which is a crucial step to improve data quality and reliability. Raw sensor data may contain noise, missing values, or inconsistencies due to environmental factors or sensor errors. Techniques such as data cleaning, interpolation, normalization, and outlier detection are applied to ensure consistency. This step also includes time synchronization of data collected from different sensors and conversion into a uniform format suitable for analysis.

The next step is **feature extraction and selection**, where important pollutant parameters are identified and selected for AQI classification. Each pollutant contributes differently to air quality, and their concentrations are used to compute sub-indices. Relevant features such as pollutant levels, temporal patterns (time of day, seasonal variations), and environmental conditions may be considered. Feature selection techniques help in reducing redundancy and improving model efficiency.

System Architecture



The system consists of multiple layers including data acquisition, preprocessing, classification, and visualization.

- **Data Acquisition Layer:** Collects pollutant data from sensors
- **Preprocessing Layer:** Cleans and normalizes data
- **Feature Extraction Layer:** Identifies key pollutant features
- **Classification Layer:** Uses ML models to predict AQI
- **Visualization Layer:** Displays results via dashboards
- **Output Layer:** Provides alerts and reports

V RESULTS &OUTPUT

.

he system successfully classified air quality into different AQI categories based on sensor readings. It provided real-time updates and visual representations of pollution levels.

The model achieved high accuracy and efficiently processed large datasets. The dashboard displayed pollutant levels, AQI categories, and trends over time.

VI CONCLUSION

The proposed system successfully integrates **multi-pollutant sensor data, data preprocessing techniques, and machine learning algorithms** to classify air quality into different AQI categories. By considering multiple pollutants such as PM2.5, PM10, CO, NO_2, SO_2, and O_3, the system provides a more comprehensive and reliable assessment compared to traditional single-parameter approaches. The use of IoT-based sensors enables continuous data collection, ensuring real-time monitoring of air quality conditions.

One of the major achievements of this project is the implementation of machine learning models that can accurately classify AQI levels based on complex relationships between different pollutants. The system demonstrates high efficiency in processing large datasets and delivering quick predictions. This reduces dependency on manual analysis and improves overall decision-making capabilities.

The visualization component of the system, including dashboards and graphical representations, plays a crucial role in making the data easy to understand for both authorities and the general public. The use of color-coded AQI categories helps users quickly interpret pollution levels and take necessary precautions. Additionally, the alert mechanism ensures that timely warnings are issued when pollution exceeds safe limits, thereby helping in reducing health risks.

Another important advantage of the proposed system is its scalability and cost-effectiveness. Unlike traditional monitoring stations that are expensive and limited in number, the use of distributed sensors allows wider coverage of urban areas. This makes the system suitable for smart city applications and large-scale environmental monitoring.

REFERENCE

[1] Kumar, R. D., Prudhviraj, G., Vijay, K., Kumar, P. S., & Plugmann, P. (2024). Exploring COVID-19 through intensive investigation with supervised machine learning algorithm. In Handbook of Artificial Intelligence and Wearables (pp. 145-158). CRC Press.

[2] Swathi, B., Vijay, K., Sushanth Babu, M., & Dinesh Kumar, R. (2024, November). Machine Learning Techniques in Cloud Based Intrusion Detection. In The International Conference on Artificial Intelligence and Smart Environment (pp. 557-564). Cham: Springer Nature Switzerland.

[3] Sv satyakrishna, shirisha rangu ,bhargavi nalacheruve.(2024) Prospective investigation on colorectal cancer with SMOTE on machine learning Algorithm

[4] Dr.G.Vishnu Murthy, BhargaviNalacheruve 1Professor, Department of computer Science & engineering, Anurag University, TS, India. 2Student, Department of computer Science & engineering, Anurag University, TS, India.

[5] V. N. S. Manaswini, K. K, C. Nigam, S. S. Ali, R. Niranjana, and Suman, "Real-Time Object Detection in Drone Surveillance Using YOLOv5," in Proc. 2025 3rd Int. Conf. IoT, Communication and Automation Technology (ICICAT), Gorakhpur, India, 2025, pp. 1–6, doi: 10.1109/ICICAT68430.2025.11414670.

[6] B. Soundarya, V. N. S. Manaswini, M. Ayyakrishnan, R. D. Kumar, "Contextual Analysis of Big Data Analytics in Intelligent Transportation Frameworks," in Intersection of Artificial Intelligence, Data Science, and Cutting-Edge Technologies: From Concepts to Applications in Smart Environment, Lecture Notes in Networks and Systems, vol. 1353, Cham: Springer, 2025, doi: 10.1007/978-3-031-88304-0_79.

[7] R. D. Kumar, V. N. S.Manaswini, "Applications of blockchain in smart cities: detecting fake documents from land records using blockchain technology," in Blockchain for Smart Cities, Elsevier, 2021, pp. 105–117, doi: 10.1016/B978-0-12-824446-3.00017-X.

[8] Tejavath Veeramma, Badarla Anil, Guguloth Ravinder, "An advanced movie recommender using collaborative filtering and sentiment analysis," International Research Journal of

Modernization in Engineering Technology and Science, vol. 7, no. 7, July 2025, doi: 10.56726/IRJMETS81618.

[9] Ravi Kumar Banoth, Ramana Murthy B V, “Automatic crop recommendation system using LightGBM and decision tree machine learning models,” Journal of Machine and Computing, vol. 5, no. 1, pp. 343, Jan. 2025, doi: 10.53759/7669/jmc202505026.

[10] Ravi Kumar Banoth, Dr. B.V. Ramana Murthy, “Smart agriculture through IoT and machine learning for analyzing carbon footprints,” in Proc. Int. Conf. Computer Science and Communication Engineering (ICCSCE), Apr. 2025.

[11] Ravi Kumar Banoth, B. V. Ramana Murthy, “Soil image classification using transfer learning approach: MobileNetV2 with CNN,” SN Computer Science, vol. 5, art. no. 199, 2024, doi: 10.1007/s42979-023-02500-x.