Smart Air Pollution Detection

Lokmanya Tilak college of engineering
Kedar N. Ghostekar

Problem Statement

- Brief Overview:
  Rising Air Pollution Levels: Urbanization and industrial activities have significantly increased air pollution, posing serious health and environmental risks.
  Lack of Real-Time Monitoring: Traditional air quality monitoring systems are costly, limited in coverage, and do not provide real-time data.
  Health Hazards: Pollutants like PM2.5, NO_2, and SO_2 contribute to respiratory diseases, heart conditions, and other long-term health problems.
  Need for an Efficient Solution: There is a growing demand for an affordable, accurate, and accessible air pollution detection system to provide real-time alerts and data-driven insights.

- Key Objectives:
- Identify Major Pollutants – Detect harmful air pollutants like PM2.5, NO_2, and SO_2 affecting public health.
- Enhance Public Awareness – Provide real-time pollution data to individuals and communities for informed decision-making.
- Enable Predictive Analysis – Use collected data to forecast air quality trends and potential pollution spikes.
- Promote Scalability & Accessibility – Develop a cost-effective and scalable solution for widespread adoption.
-

Dataset Overview:

- Dataset Description:

  Location Details: State, city, and type of area (residential, industrial, etc.).
  Pollutant Levels: SO_2, NO_2, RSPM, SPM, and PM2.5 concentrations.
  Collected from air quality monitoring stations across different locations.

- Data Preprocessing:
  Handling missing values by imputing mean values for certain pollutants.
  Encoding categorical variables (state, location, type) for model training.
  Splitting dataset into training and testing sets for better model evaluation.

Methodology

- Approach:
  1. Data Collection
  2. Data Preprocessing
  3. Model Selection & Training
  4. Model Evaluation
  5. Deployment & Insights

- Algorithms Used:
  Random Forest Regressor
  Label Encoding
  Mean Imputation

Conclusion

- Summary:
  The Air Pollution Detector effectively predicts PM2.5 levels using machine learning, helping in real-time air quality monitoring.
  Random Forest Regressor proved to be a reliable model with good accuracy for pollution prediction.
- Results & Impact: The model successfully predicts air pollution trends, helping authorities and individuals take proactive measures.
- Conclusion: This project provides a scalable, data-driven approach to air quality monitoring, with potential improvements like real-time IoT integration.

Future Work:

- Experiment with advanced algorithms like Gradient Boosting, XGBoost, or Deep Learning for better predictions.
- Integrate weather conditions, traffic density, and seasonal variations into the dataset.

GitHub Repository Link of a project

- https://github.com/Ghosting-around/Edunet-Shell

References

- https://www.linkedin.com/in/kedar-ghostekar-2003a5314

Thank You