SMART INDIA HACKATHON 2024

TITLE PAGE

- **Problem Statement ID:** 1734
- **Problem Statement Title:** Downscaling of Satellite based air quality map using AI/ML
- **Theme:** Clean & Green Technology
- **PS Category:** Software
- **Team ID:**
- **Team Name:** Tech Titans

Tech Titans

IDEA TITLE

SMART INDIA HACKATHON 2024

Proposed Solution:

1. Data Collection & Preprocessing:

- Acquiring NO2 data from satellites (e.g., Sentinel-5P) and handle gaps with imputation techniques.

2. Model Development:

- Training a Machine Learning model (e.g., Random Forest) to downscale satellite data to fine resolution.

3. Validation:

- Apply cross-validation and test with independent data to ensure accuracy.

4. Deployment:

- Deploy as a web tool for real-time NO2 monitoring with a user-friendly interface for researchers and government bodies.

How it Addresses the Problem: Provides detailed local NO2 maps, improving air quality assessments.

Innovation & Uniqueness: Combines satellite and ground data with advanced ML for accurate, real-time air quality mapping and no comprehensive solution exists like this.

Fig 1. Snap of dataset

Data Acquisition

•**API**: Accessed via Sentinel Hub using specific coordinates and date range.

•**Token**: Used for API authentication.

Data Processing

•**Script**: Extracted NO2 levels in floating-point format.

•**Request**: Retrieved NO2 data for the region.

Visualization

•**Output Grid**: Displays NO2 levels per pixel, with a color bar indicating concentration levels (dark = low, light = high).

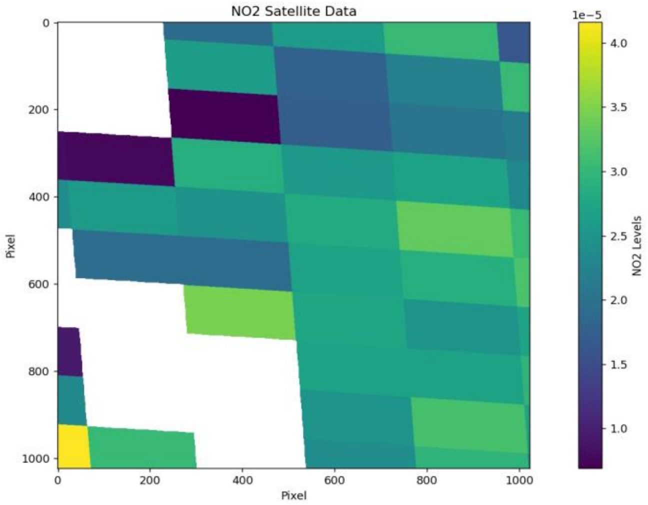


Fig 2. Outcome

Tech Titans

TECHNICAL APPROACH

SMART INDIA HACKATHON 2024

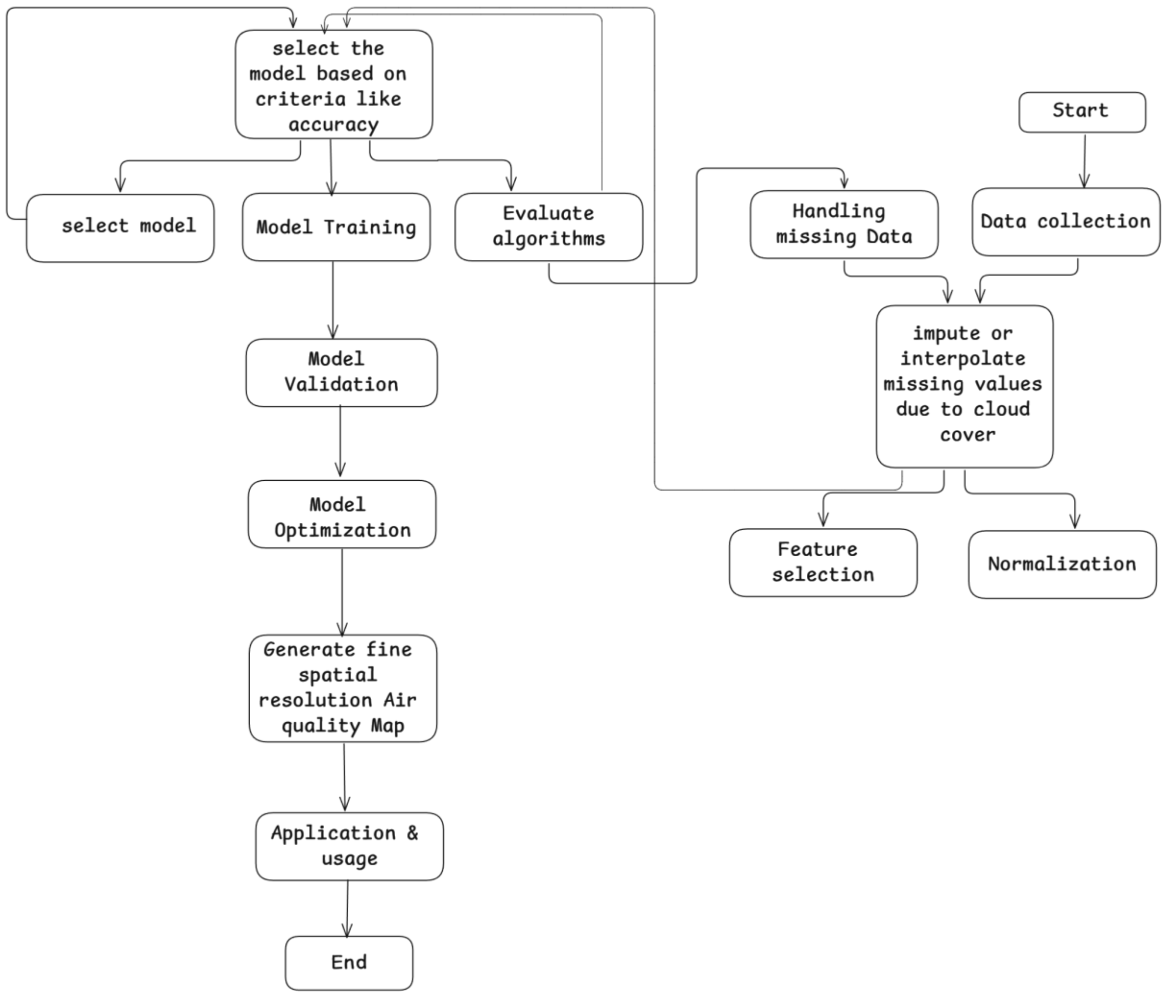
❖ Technologies to be Used

Programming Languages: Python
Database: MongoDB
Frameworks & Libraries: TensorFlow, PyTorch, Scikit-Learn, GeoPandas, Matplotlib, Pandas

❖ Methodology & Process for Implementation

Data Collection: Satellite data ,data from air quality monitoring stations
Data Preprocessing: Handling gaps in data, Extracting relevant environmental features
Algorithm Selection: Random Forest, CNN, GBM
Training: data to train the model
Validation: Independent dataset validation, cross-validation

Block Diagram



Tech Titans

FEASIBILITY AND VIABILITY

SMART INDIA HACKATHON 2024

Feasibility

- **Data Availability**: Abundant satellite and ground data, though preprocessing is needed.
- **Technical Maturity**: Reliable AI/ML algorithms and robust libraries are available.
- **Computational Resources**: Requires high-performance computing or cloud solutions.

Challenges & Risks

- **Data Gaps**: Missing or noisy data could lead to inaccurate maps.
- **Model Complexity**: Balancing model sophistication with risk of overfitting or underfitting.
- **Scalability**: Managing large datasets might cause performance slowdowns.
- **Validation**: Ensuring the model works well across different regions.
- **Interpretability**: Complex models might be difficult for stakeholders to understand.

Strategies

- **Data Gaps**: Implement cloud masking, interpolation, and data augmentation.
- **Model Complexity**: Start with simpler models, use cross-validation, and combine model for better accuracy.
- **Scalability**: Leverage cloud computing and parallel processing to handle big data efficiently.
- **Validation**: Use diverse and independent datasets to ensure the model generalizes well.

Tech Titans

IMPACT AND BENEFITS

SMART INDIA HACKATHON 2024

Target Audience:

•**Researchers:** Access to high-resolution data improves the quality and scope of environmental research.

•**Government Bodies:** Enhanced capabilities for real-time air quality monitoring and rapid response to pollution events.

•**Public:** Better access to localized air quality information empowers individuals to make informed health and lifestyle decisions.

Benefits:

•**Health Improvements:** Facilitates early warning systems for air pollution, reducing health risks and chronic diseases.

•**Community Engagement:** Fosters community involvement in environmental issues through accessible data and visualization tools.

•**Targeted Pollution Control:** Enables precise targeting of pollution sources, leading to more effective cleanup and mitigation efforts.

•**Enhanced Environmental Policies:** Provides data-driven evidence to support stringent air quality regulations and compliance with environmental standards.

•**Sustainable Development:** Supports green initiatives and sustainable practices by offering insights into pollution trends and impacts.

Tech Titans

RESEARCH AND REFERENCES

SMART INDIA HACKATHON 2024

RESEARCH

- "NASA Worldview," NASA Earth Observing System Data and Information System. [Online]. Available: Link.
- "IHME Cloud Storage," Institute for Health Metrics and Evaluation. [Online]. Available: Link.

REFERENCES

- M. Kim, D. Brunner, and G. Kuhlmann, "Importance of satellite observations for high-resolution mapping of near-surface NO2 by machine learning," Remote Sensing of Environment, vol. 264, p. 112573, 2021, doi: 10.1016/j.rse.2021.112573.
- C. Huang, K. Sun, J. Hu, T. Xue, H. Xu, and M. Wang, "Estimating 2013–2019 NO2 exposure with high spatiotemporal resolution in China using an ensemble model," *Environmental Pollution*, vol. 292, Part A, p. 118285, 2022, doi: 10.1016/j.envpol.2021.118285.
- K. L. Chan, E. Khorsandi, S. Liu, F. Baier, and P. Valks, "Estimation of Surface NO2 Concentrations over Germany from TROPOMI Satellite Observations Using a Machine Learning Method," *Remote Sens.*, vol. 13, no. 5, p. 969, Mar. 2021, doi: 10.3390/rs13050969.
- R. H. Narashid and W. M. N. W. Mohd, "Air Quality Monitoring Using Remote Sensing and GIS Technologies," presented at the 2010 CSSR, Universiti Teknologi MARA, 2010.
- P.S. Kanaroglou, N.A. Soulakellis., N.I. Sifakis.,. Improvement of satellite derived pollution maps with the use of a geostatistical interpolation method. International Journal of Geographical Systems. (4); pp. 2002, 193-208.