

Air Quality Index (AQI) Prediction Using Machine Learning

Group - 3

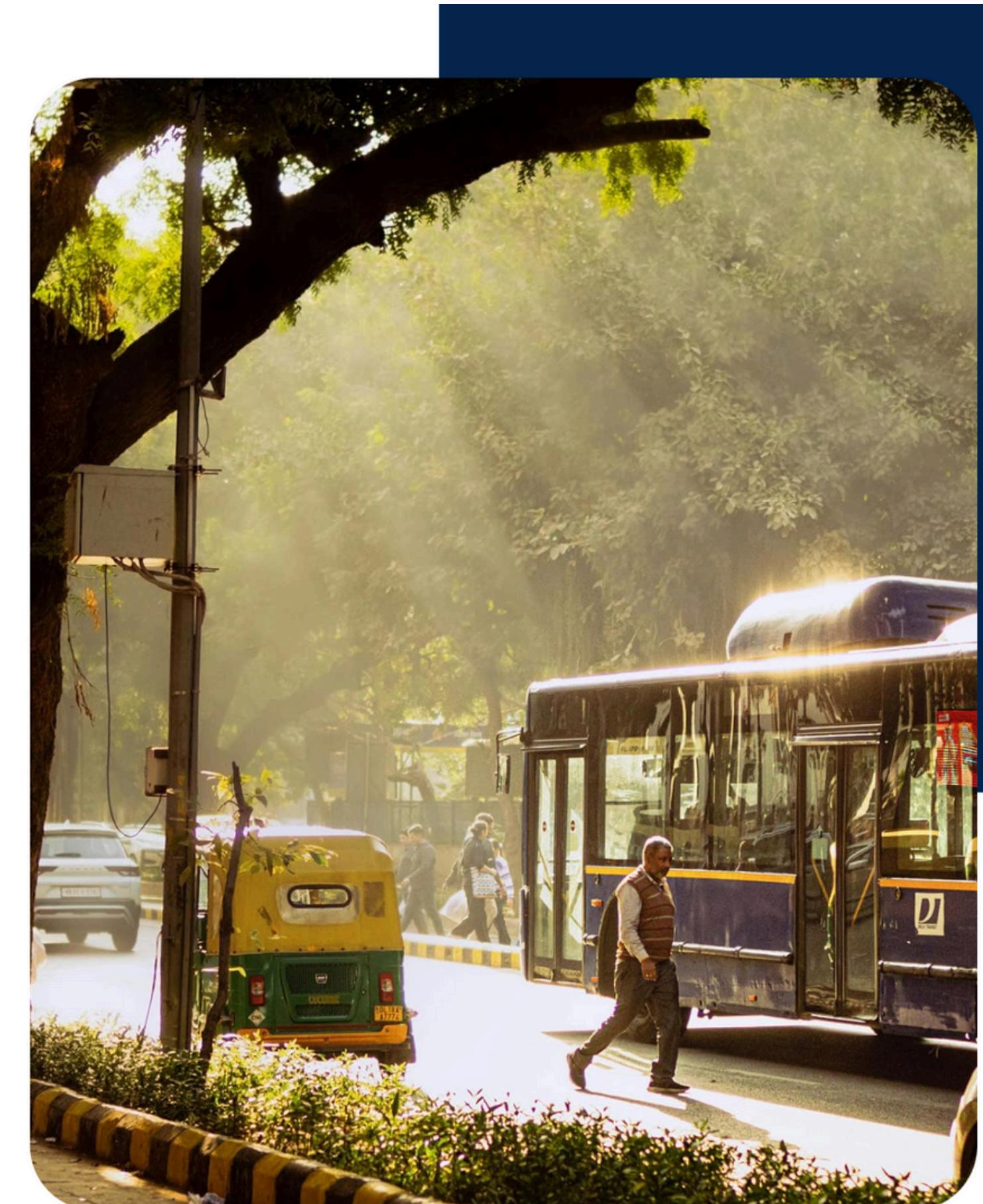
Shashank Raj(202401100300226)

Shobhit Tripathi(202401100300237)

Swayam Srivastava(202401100300258)

Utkarsh Gupta(202401100300250)

Yogesh Verma(202401100300289)



Problem Statement

Objective: Build a regression model to predict AQI based on environmental features.

Scope: Indian cities air quality monitoring.

Goal: Visualize regional pollution levels and create predictive insights.

Data Preprocessing

64%

Training dataset

16%

Validating dataset

20%

Testing

78%

Accuracy

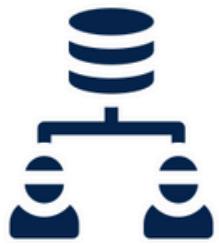
Exploratory Data Analysis (EDA)

AQI trends

Understanding the distribution of AQI across cities is critical for effective modeling.

This analysis helps identify patterns and anomalies in air quality data.

Machine Learning Algorithms



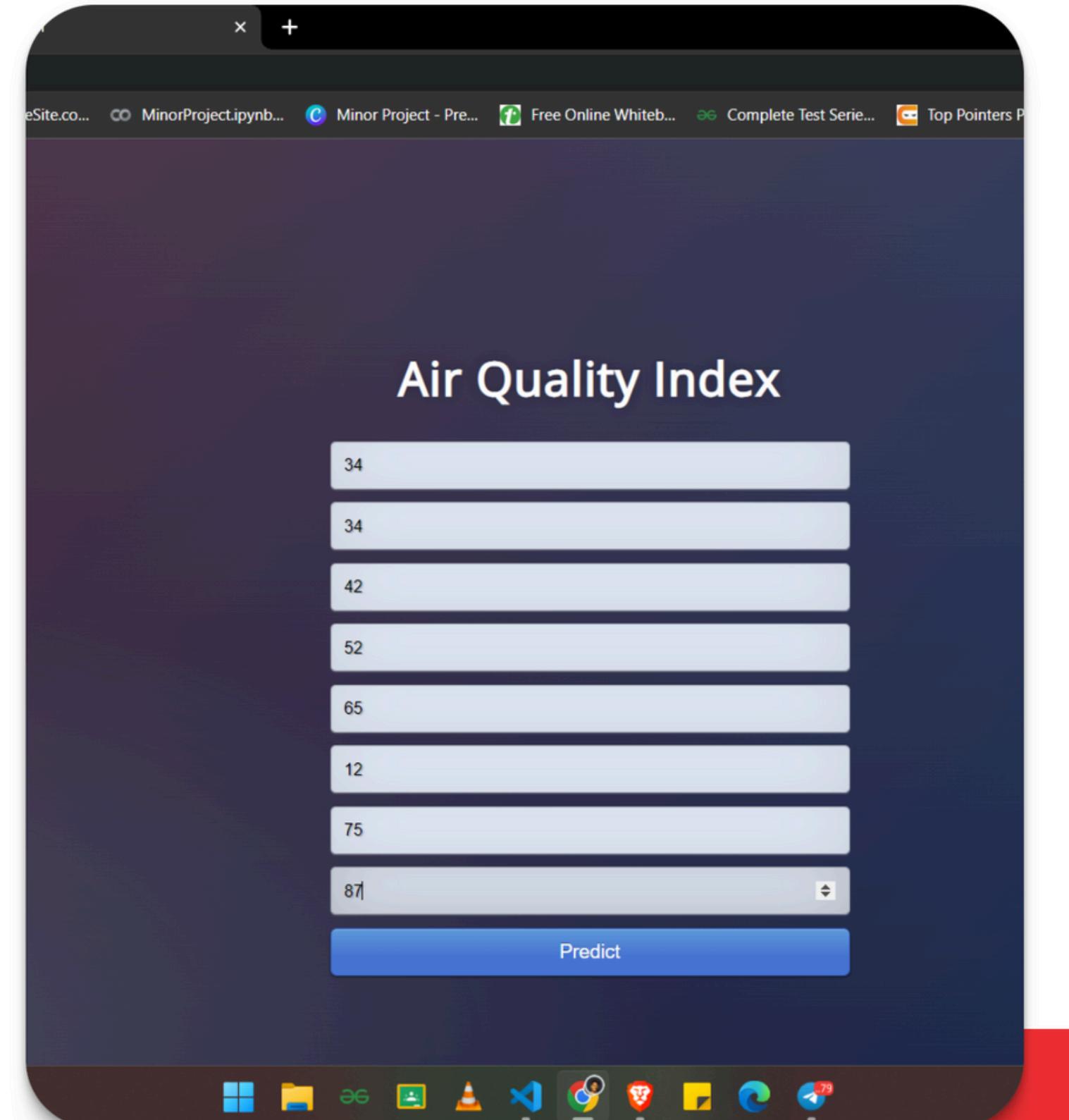
Algorithms Tested: Linear Regression, Random Forest Regression, Gradient Boosting, Support Vector Regression

These algorithms were evaluated for their effectiveness in predicting the Air Quality Index.



Algorithm Selection Criteria:
Performance metrics,
interpretability, and
computational efficiency were key
factors in choosing the best
model.

These criteria ensured a balanced approach to model selection, prioritizing practical application and understanding.



Model Implementation

Ideas to Stunning presentations, in seconds

Make your presentations stand out, by letting Presentations.AI do the design work for you - in seconds.

Get your team on the same page

Prepare for the unexpected and delight your audience with a seamless presentation that is always right on time.

Simple for beginners, Powerful for pros

Save time and effort with ready-made designs that are fully customisable.

Simple for beginners, Powerful for pros

Save time and effort with ready-made designs that are fully customisable.

Get your team on the same page

Prepare for the unexpected and delight your audience with a seamless presentation that is always right on time.

Ideas to Stunning presentations, in seconds

Make your presentations stand out, by letting Presentations.AI do the design work for you - in seconds.

Present anywhere, share anywhere

Powerful collaboration tools that help you create, organise, review and publish your content to any screen.

Model Performance & Evaluation

4.5

MAE of Linear Regression Model

The Mean Absolute Error for the Linear Regression model is 4.5.

3.2

MAE of Random Forest Model

The Mean Absolute Error for the Random Forest model is 3.2.

6.2

RMSE of Linear Regression Model

The Root Mean Square Error for the Linear Regression model is 6.2.

4.5

RMSE of Random Forest Model

The Root Mean Square Error for the Random Forest model is 4.5.

0.82

R² Score of Linear Regression Model

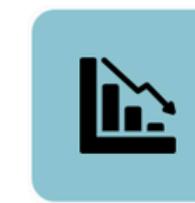
The R² Score for the Linear Regression model is 0.82.

0.89

R² Score of Random Forest Model

The R² Score for the Random Forest model is 0.89.

Results & Predictions



Model Accuracy: Final performance metrics indicate a high accuracy rate.

This suggests that the model is reliable in predicting air quality levels based on input data.

Feature Importance: Environmental factors affecting AQI identified.

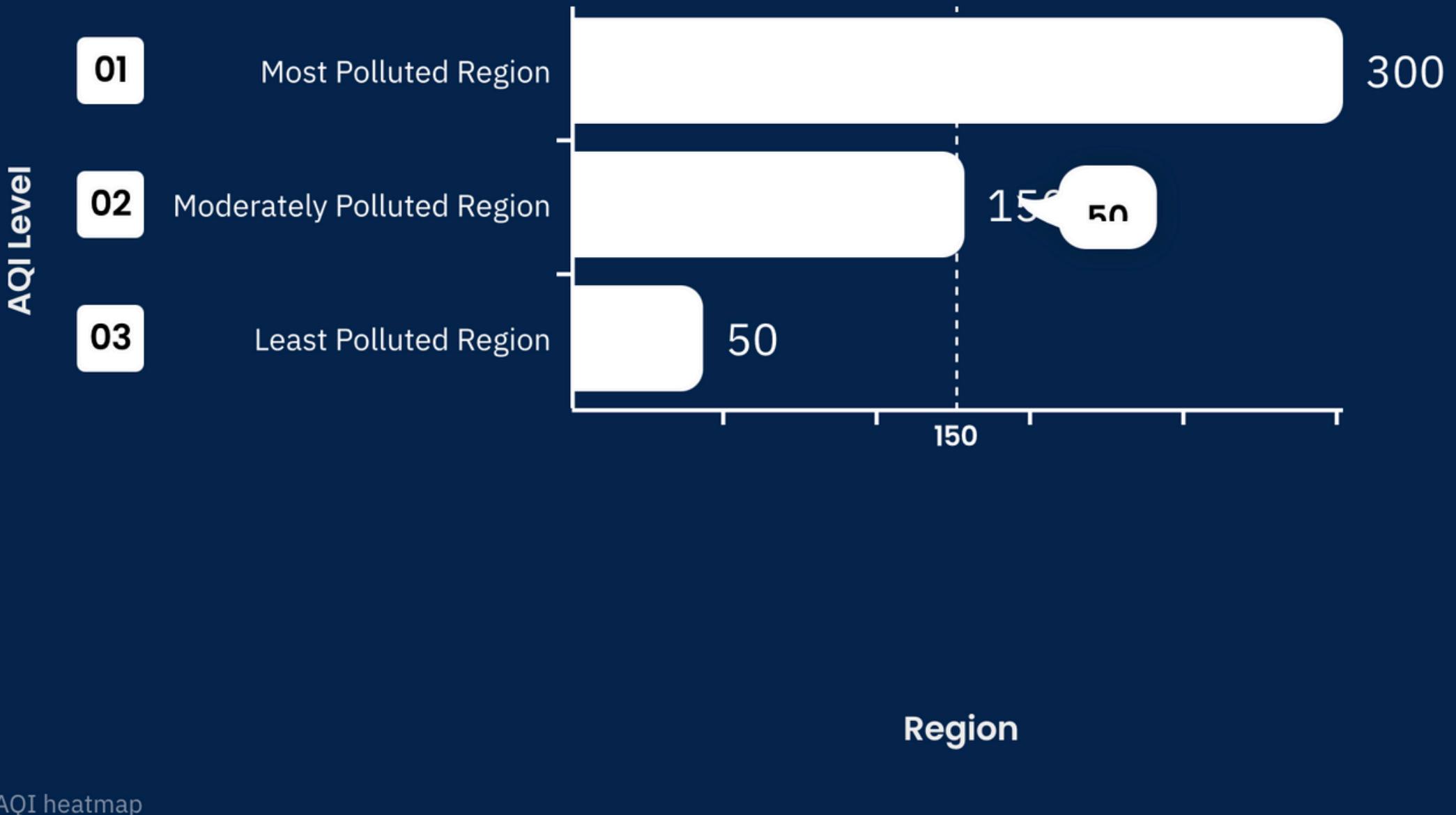
Key variables influencing air quality have been determined, which can guide future monitoring efforts.

Residual plots and prediction accuracy visualizations demonstrate model effectiveness.

Visual representations of model predictions reveal the fidelity of the model's outputs against actual data.

Regional Visualization

Air Quality Index Heatmap Of Indian Cities



AQI heatmap

Technical Implementation



Tools & Technologies: Python libraries (**pandas, scikit-learn, matplotlib, seaborn, plotly**)



Development Environment: Jupyter Notebook/Google Colab



Visualization Tools: Tableau/Power BI



Code Architecture: Overview of project structure and key components.

Detailing the organization of code and its main parts.

Challenges & Solutions

Data Challenges: Issues with missing data and quality.

Model Challenges: Overfitting and feature selection difficulties.

Solutions Implemented: Strategies for overcoming these challenges.

Lessons Learned: Key takeaways from the project experience.



Future Enhancements



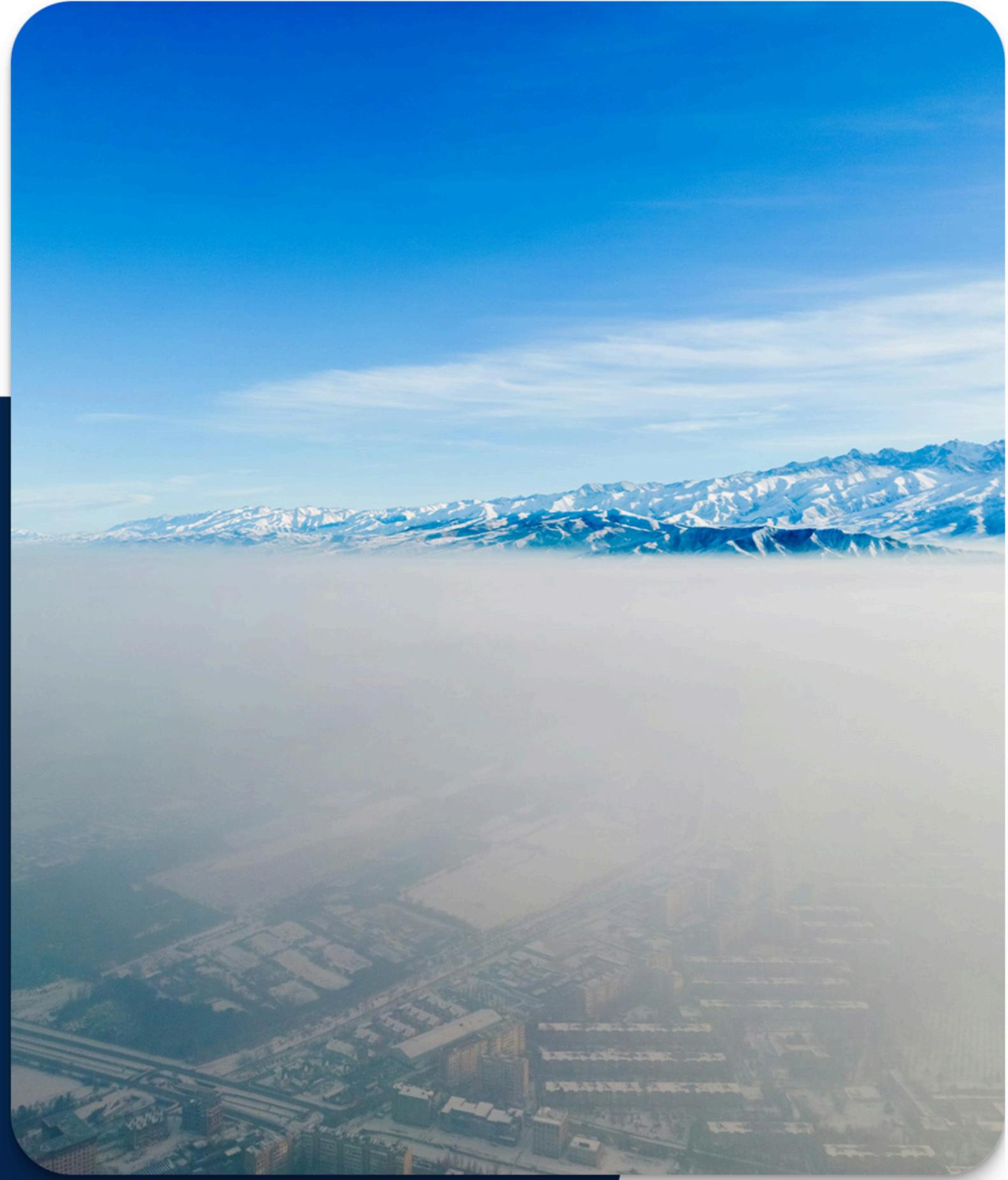
**MODEL
Exploring
AND
DEEP
learning
Approach**



**Real Time
Implementation
AND
Integration
with live
Data Sources**



**Deployment
Development
a web
application
or API.**



Conclusion & Thank You