BEYOND THE AQI

Live Data and AQI prediction of Indian cities

DISASTER MANAGEMENT

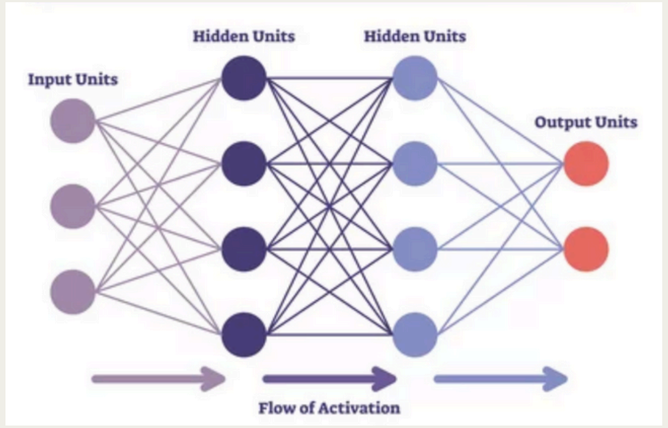
Harshit Kashyap – 2022201050
Ujjwal Prakash – 2022202009
Priyank Mahour – 2022201047
Garima Tripathi - 2021101123

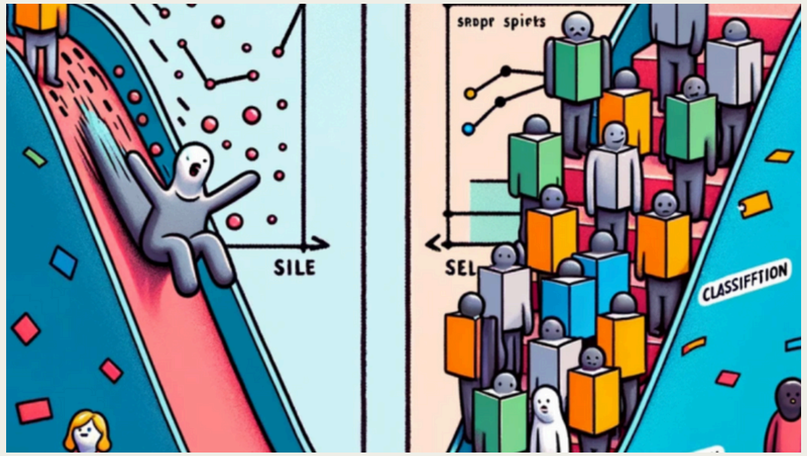
BACKGROUND

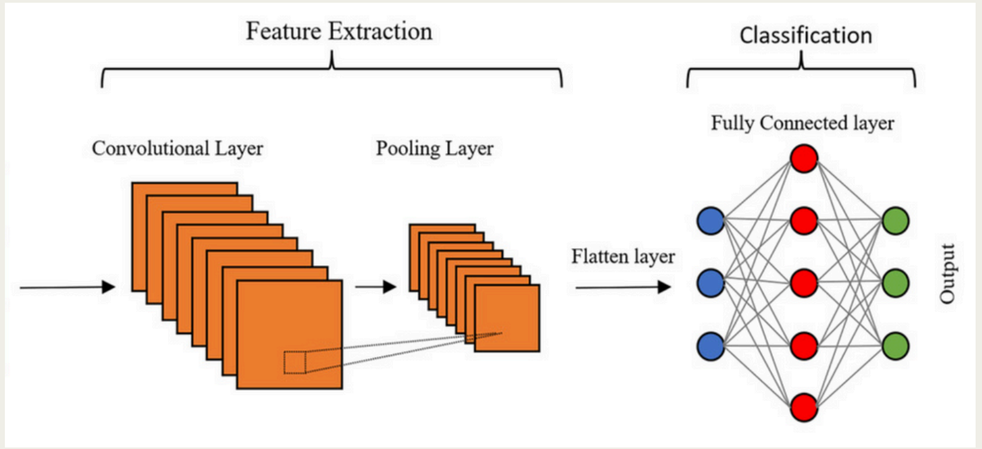
Air pollution is a pressing global issue with detrimental effects on human health and the environment. Understanding the concentration of pollutants in different cities over time is crucial for implementing effective mitigation strategies. Therefore, we developed a web application to visualize the concentration of various pollutants in different cities over a span of 9 years.

OBJECTIVE

- Develop a web application for visualizing the concentration of pollutants in different cities.
- Implement a neural network model to predict the current AQI based on pollutant data.
- Allow users to upload images denoting pollution and provide severity results.







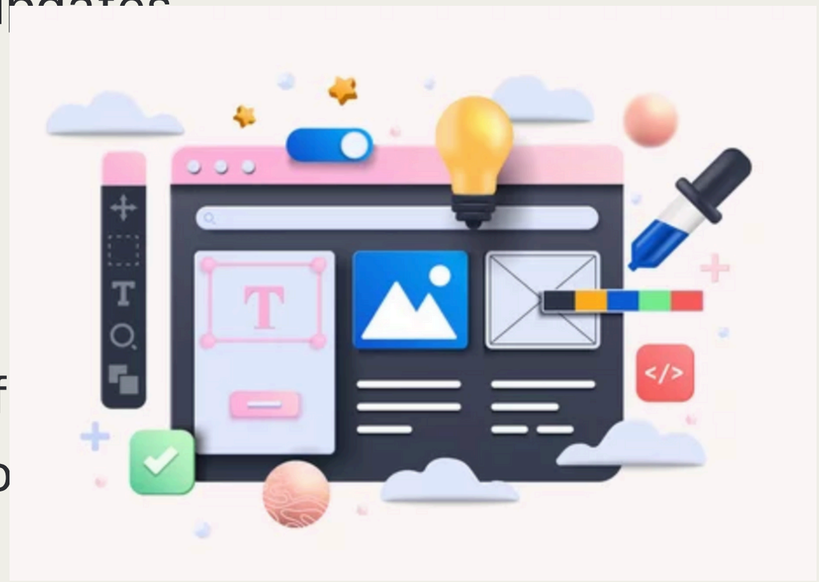
SCOPE

The scope of this project is to provide users with a user-friendly platform to visualize and understand the trends of air pollution in different cities. Additionally, we aim to predict the current Air Quality Index (AQI) by considering various pollutants as factors.

METHODOLOGY

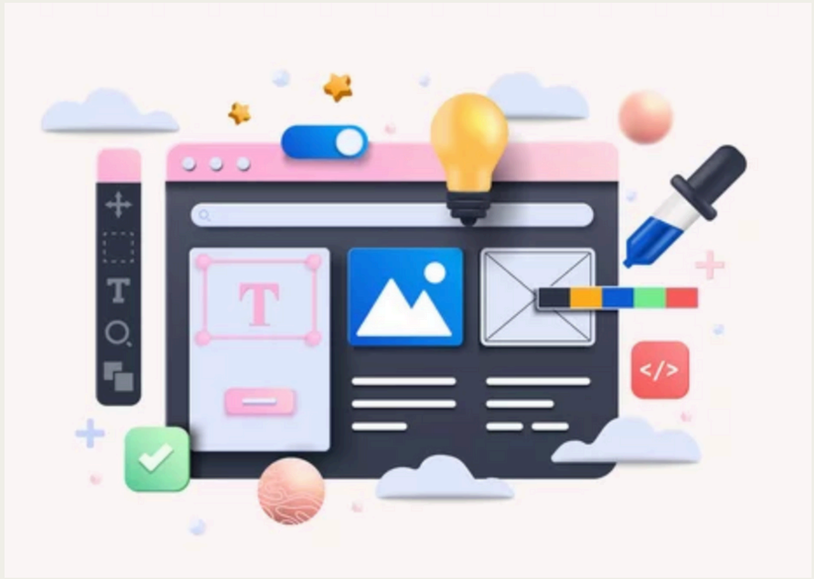
Data Collection :

1. Historical Data Collection:
   - Gathered pollutant concentration data from various cities, including state capitals and spanning 9 years (2015 - 2023).
2. Real-Time Data Collection:
   - Used Ambee-Data External API which fetches live AQI and pollutant concentrations integrated into project's pipeline for continuous updates
3. Analysis and Comparison:
   - Employed Kaggle datasets for historical analysis which are useful in identifying trends and seasonal patterns.
4. Government Sources:
   - Accessed official air quality data which is used as Referenced f regulatory information and utilized for comprehensive AQI and po level datasets.



METHODOLOGY

- Data Availability: Government websites provided comprehensive datasets on AQI, pollutant levels, and related environmental policies and regulations.
- Data Preprocessing: Cleaned and processed the collected data for analysis and visualization.
- Web Application Development: Utilized web development technologies to create an interactive platform for users.
- Neural Network Model: Implemented a neural network with 4 layers to predict the current AQI based on pollutant factors.
- Image Processing: Employed image processing techniques to analyze user-uploaded images and determine pollution severity.



RESULTS

The web application successfully visualizes the concentration of pollutants in different cities over the specified time period. Additionally, the neural network model accurately predicts the current AQI based on pollutant data. Users can upload images, and the application provides severity results indicating the level of pollution.

CATEGORICAL LOSS IN CLASSIFICATION

Categorical loss measures the performance of a classification model by comparing the model's predicted probabilities to the actual categorical labels.

Cross-Entropy Loss: It is also known as logarithmic loss or log loss, is a popular loss function used in machine learning to measure the performance of a classification model.

$$L\left(y,\hat{y}\right) = -\frac{1}{N}\sum_{i}^{N}\sum_{j}^{C} y_{ij} log\left(\hat{y}_{ij}\right)$$

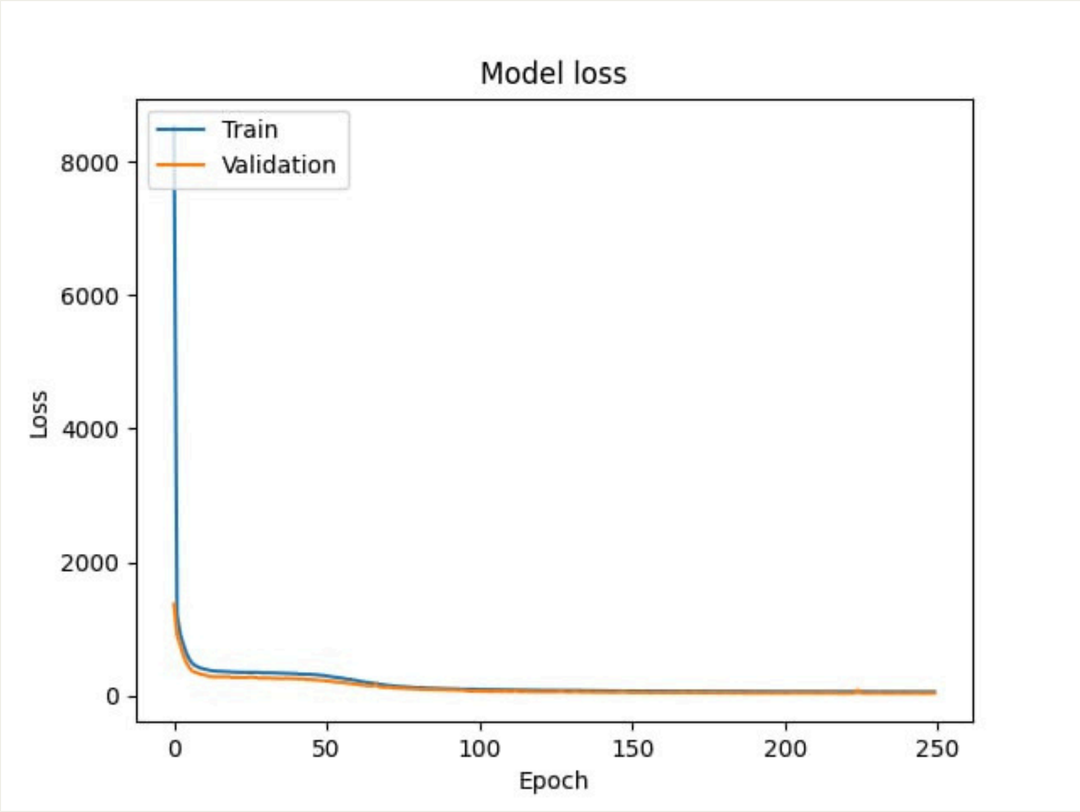
RIDGE REGRESSION

Ridge regression is a method of estimating the coefficients of multiple-regression models in scenarios where the independent variables are highly correlated

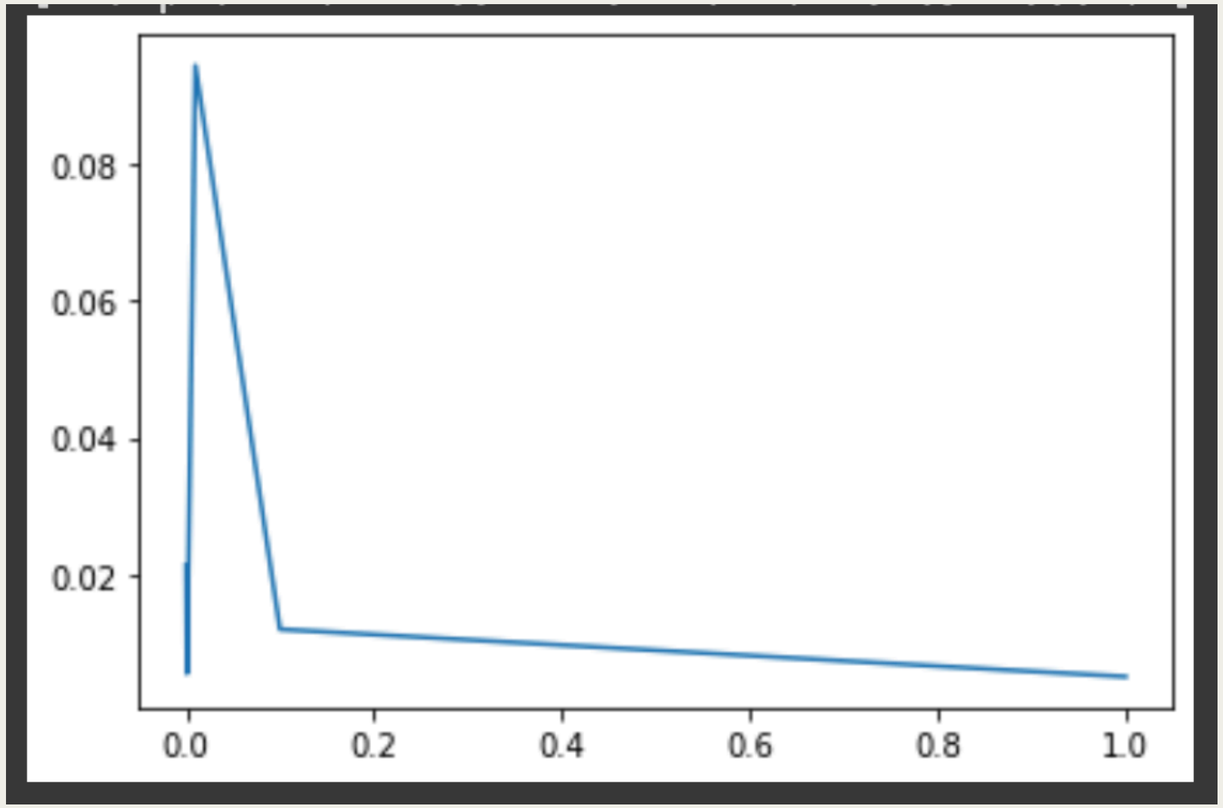
$$RSS = \sum_{i=1}^{n} (y_i - f(x_i))^2$$

$$= \sum_{i=1}^{n} (y_i - \beta_0 - \sum_{j=1}^{p} \beta_j x_{ij})^2$$

RESULTS



LOSS / EPOCH GRAPH
SIMPLE NEURAL NETWORK



RIDGE REGRESSION
x - axis : learning rate
y - axis : mean squared error

DEMONSTRATION SCREENSHOTS

BEYOND THE AQI

AQI DATA | AQI PREDICTION | LIVE AQI DATA | IMAGE PREDICTION

AQI PREDICTION

Select city

Ahmedabad

Current CO value: 0.672

Current O3 value: 7.785

Current PM2.5 value: 67.007

Current PM10 value: 140.542

Current NO2 value: 10.643

Actual AQI is: 157

Predicted AQI through neural network is: 164.52153

Predicted AQI through regression model is: 171.23761

BEYOND THE AQI

AQI DATA | AQI PREDICTION | LIVE AQI DATA | IMAGE PREDICTION

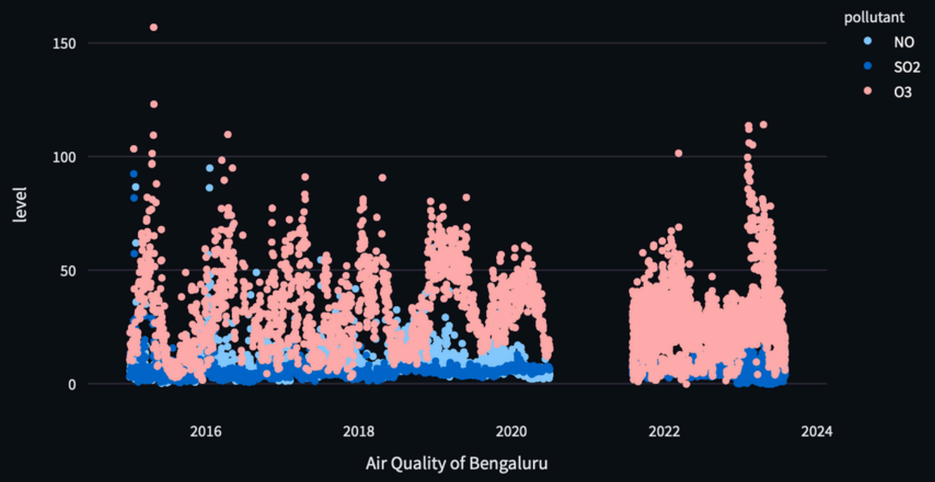
AQI DATA

Select city

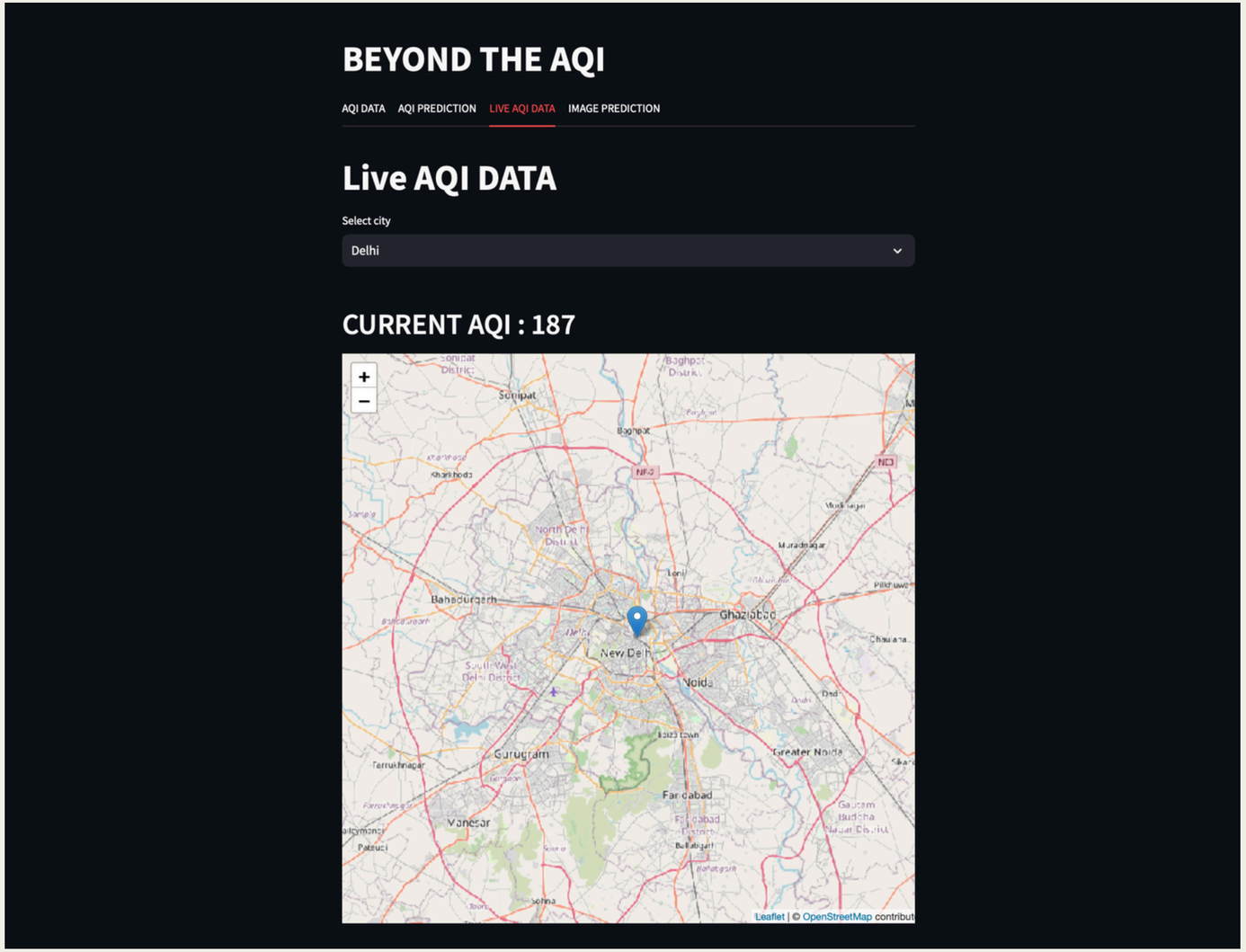
Bengaluru

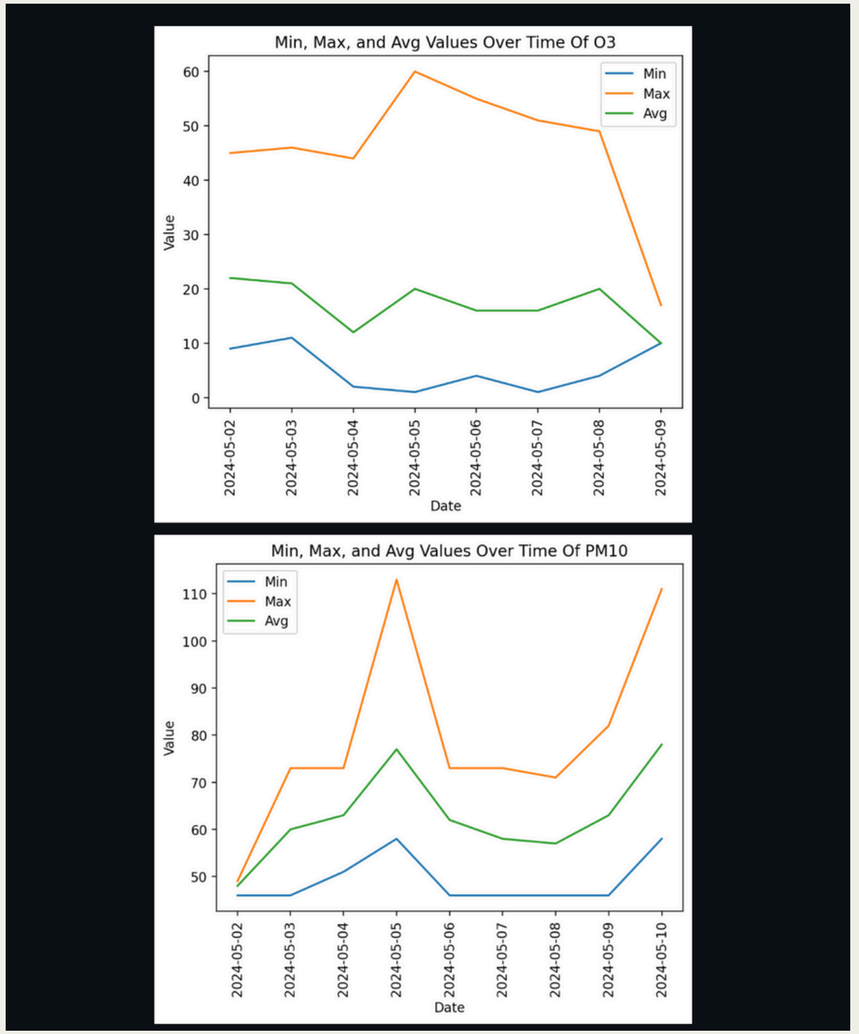
Select pollutants

NO × SO2 × O3 ×

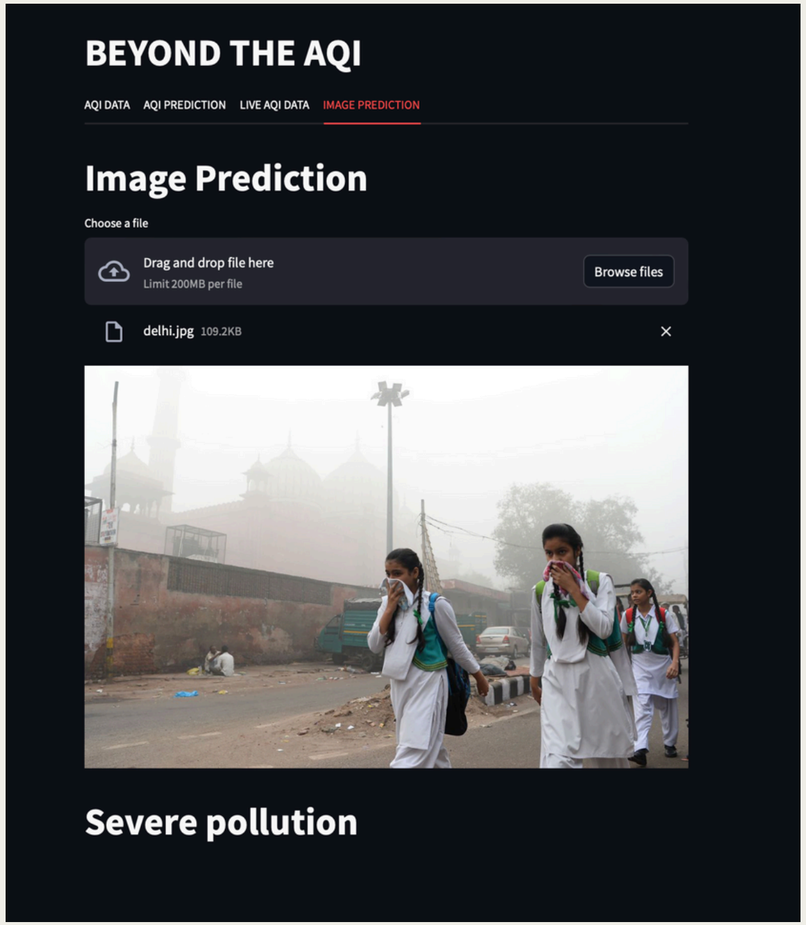


DEMONSTRATION SCREENSHOTS

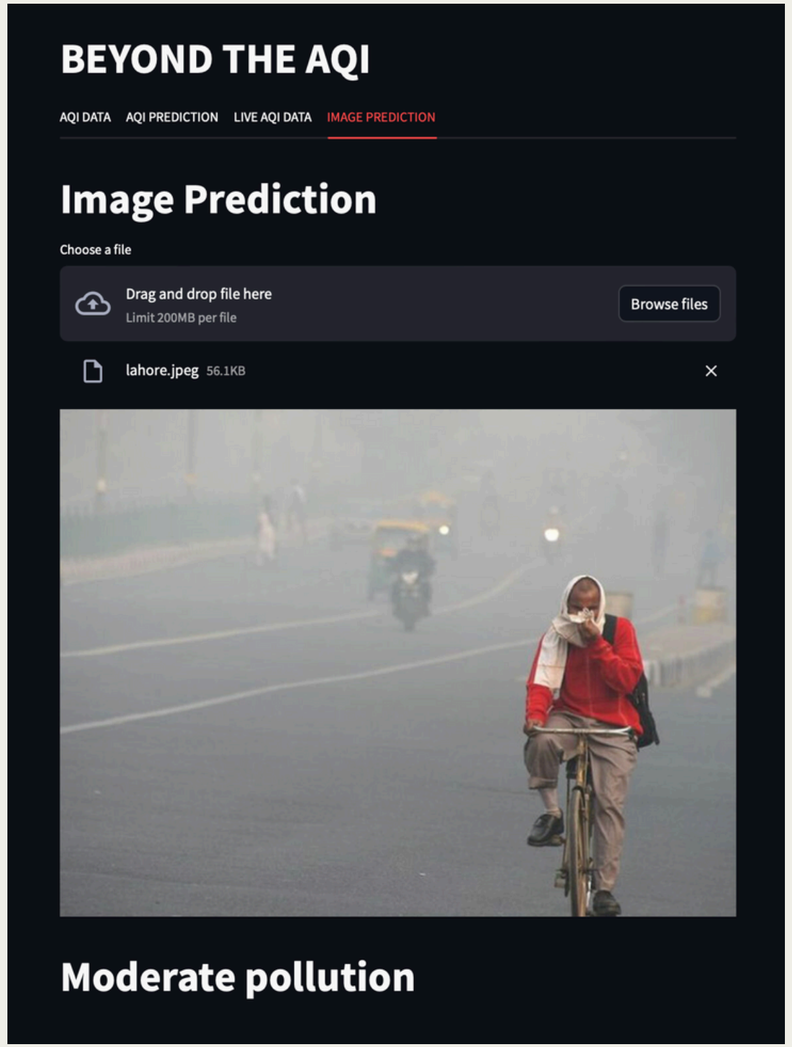




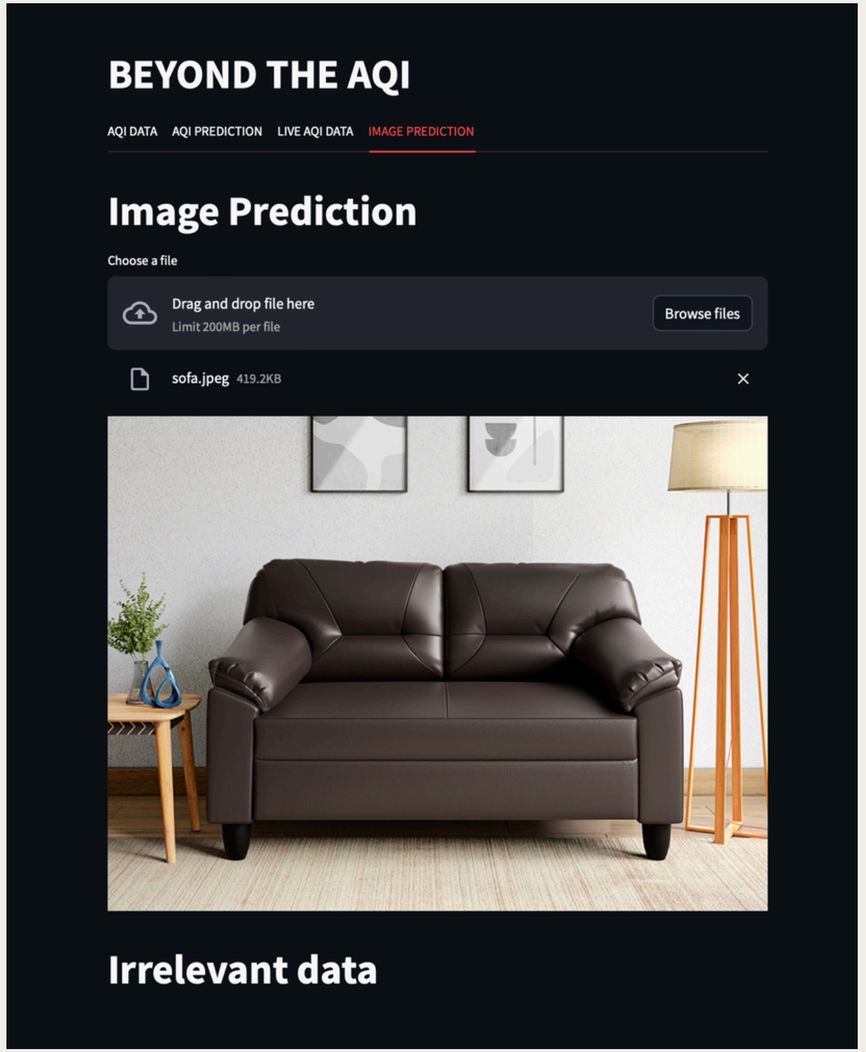
DEMONSTRATION SCREENSHOTS



Severe pollution



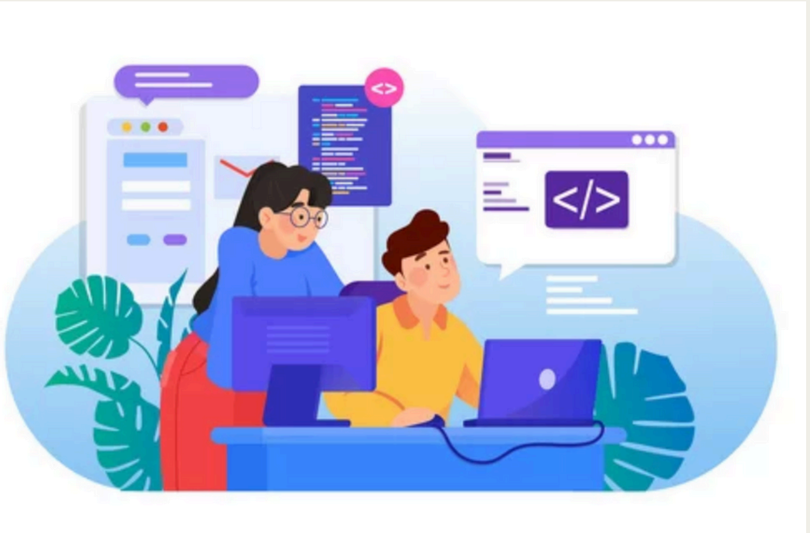
Moderate pollution



Irrelevant data

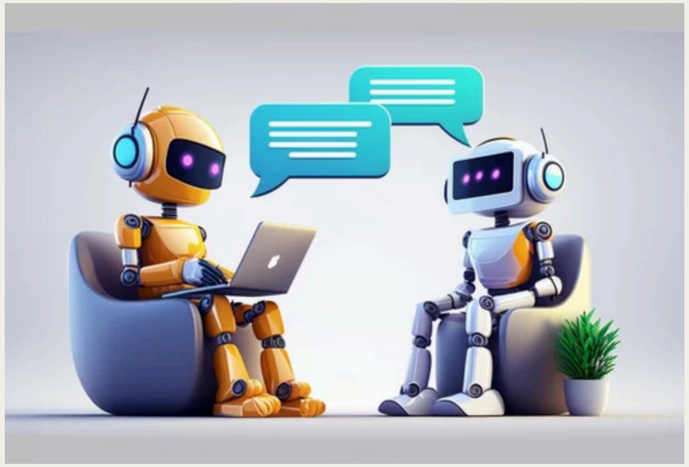
CONCLUSIONS

Through this project, we have developed a comprehensive web application for visualizing air pollution data and predicting the current AQI. This tool can be valuable for policymakers, researchers, and the general public in understanding and addressing air quality issues.



FUTURE WORKS

1. Enhance the accuracy of AQI predictions by refining the neural network model and incorporating more factors.
2. Expand the database of pollutant data to include more cities and a longer time span.
3. Incorporate real-time data feeds for up-to-date information on air quality.
4. Integrate machine learning algorithms for image classification to improve the accuracy of pollution severity assessment.
5. Collaborate with environmental agencies and organizations to validate the effectiveness of the web application in real-world scenarios.



REFERENCES

- https://en.wikipedia.org/wiki/Air_quality_index
- https://www.ncbi.nlm.nih.gov/books/NBK574581/
- https://www.tropmet.res.in/~lip/Publication/RR-pdf/RR-127.pdf
- https://www.sciencedirect.com/science/article/pii/S004565352301785X
- https://ieeexplore.ieee.org/document/9362912
- https://www.sciencedirect.com/science/article/abs/pii/S016604622030281 7
- https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2698272/
- https://www.nature.com/articles/s41598-024-54807-1

REFERENCES

- https://www.kaggle.com/code/momo88/vgg16-translearning-for-image-based-aqi-estimation/output
- https://www.kaggle.com/datasets/adarshrouniyar/air-pollution-image-dataset-from-india-and-nepal https://www.kaggle.com/datasets/fedesoriano/air-quality-data-set
- https://www.kaggle.com/code/arjunprasadsarkhel/time-series-analysis-indian-cities-aqi/notebook
- https://www.kaggle.com/code/aishwaryasarkar/india-s-aqi-eda-forecasting-sarima/notebook
- https://www.kaggle.com/code/aishwaryasarkar/india-s-aqi-eda-forecasting-sarima/input
- https://www.kaggle.com/datasets/rohanrao/air-quality-data-in-india/data?select=stations.csv
- https://www.kaggle.com/datasets/gokulrajkmv/indian-statewise-data-from-rbi
- https://www.kaggle.com/code/shubhendra7/indian-cities-dataset/input
- https://www.kaggle.com/datasets/rohanrao/air-quality-data-in-india/code
- https://www.kaggle.com/datasets/rudravpatel/aqi-data-of-india-2021-2023
- https://www.kaggle.com/datasets/abhisheksjha/time-series-air-quality-data-of-india-2010-2023/data?select=BR012.csv

Thank you!