SustainAI

# SustainAI

## EXPLAINABLE INTELLIGENCE FOR A CLEANER TOMORROW

*AI PROJECT PRESENTATION*

# Our team

- Harthik M V
- Tejas
- Sridhar Madhira
- Vinayak Sharma
- Bhanu Sharma

# Our Motivation

## Why Air Quality Analysis with XAI?

Severe Global Air Pollution

Inadequate Awareness

Can drive Smarter Policy

# GAPS IN EXISTING TECHNOLGIES

Lack of Explainability in Predictions

Lack of Usable Insights

Minimal Integration of Explainability Tools

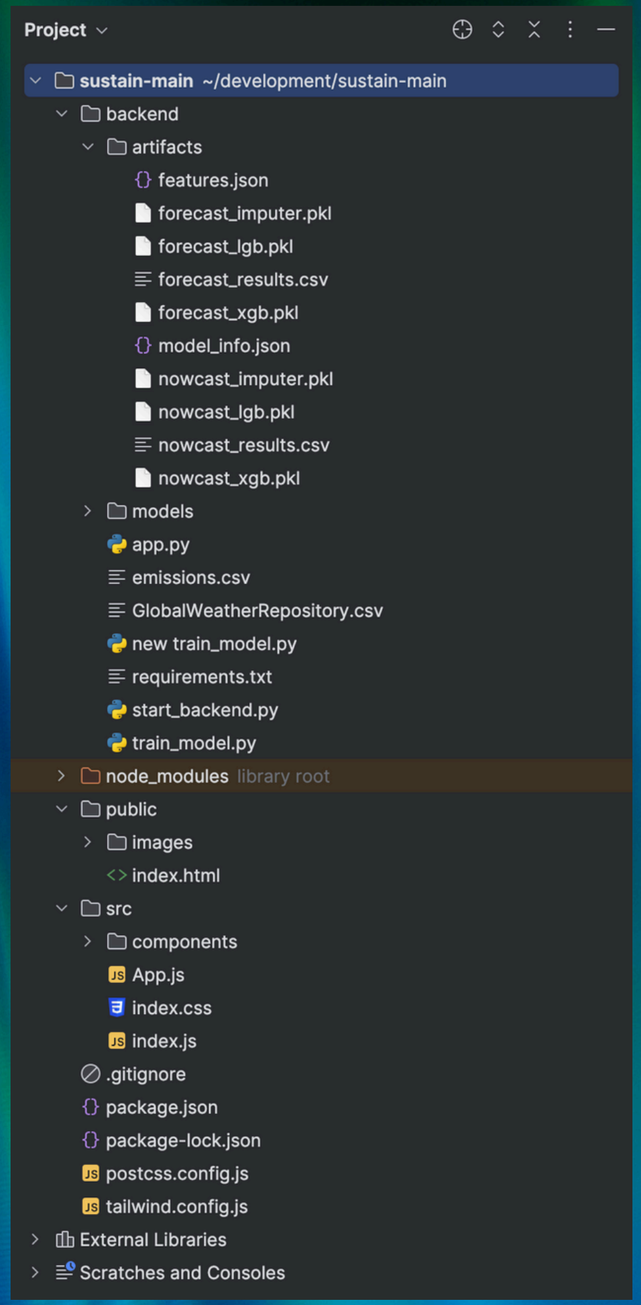
# OBJECTIVES

- Build an AI system to predict and explain air quality levels
- Provide actionable insights for cities and citizens
- Ensure environmental responsibility through carbon tracking

# PROJECT OUTCOMES


- AI-Powered AQI Forecasting & Explainability
- Actionable Insights for Stakeholders

# PROJECT STRUCTURE



# DATASET DETAILS

- *Dataset: Global Weather Repository* – Schema includes location metadata , surface meteorology , categorical weather, and pollutants/co-pollutants. Rows are global snapshots spanning varied climates *(SOURCE: kaggle)*

- Astronomy fields like moon phase, moon rise are present but not used for learning. Similarly, Time Zone fields are also excluded.

- Primary target for regression: *Air Quality PM2.5*

# MODEL TRAINING PROCESS

***PRIMARY TASK: PREDICTING PM2.5 VALUES***

- **Nowcasting: Predicting Current PM2.5 Values**
- **Forecasting: Predicting Future (+1 hour) PM2.5 Values**

# MODEL TRAINING

```
UI_FEATURES = [
    "latitude",
    "longitude",
    "temperature_celsius",
    "wind_kph",
    "pressure_mb",
    "humidity",
    "visibility_km",
    "hour",
    "month",
    "air_quality_Carbon_Monoxide",
    "air_quality_Ozone",
    "air_quality_Nitrogen_dioxide",
    "air_quality_Sulphur_dioxide"
]
```

*There are 13 features used for model training, for the ease of UI Integration*

```
# Impute missing values *within our selected features*
imp_now = SimpleImputer(strategy="median")
```

*Simple Imputer is used to fill into empty values in the features*

# MODEL TRAINING RESULTS

Here is the summary as a markdown table.

| Index | Task | Model | R2 | RMSE | MAE | Train_CO2_kg | Infer_CO2_kg |
|---|---|---|---|---|---|---|---|
| 0 | Nowcast | ExtraTrees (tuned) | 0.975697 | 6.100752 | 2.648343 | 0.0009748528 | 1.257774e-05 |
| 1 | Nowcast | RF (tuned) | 0.974562 | 6.241561 | 2.678748 | 0.002183478 | 1.372154e-05 |
| 2 | Nowcast | XGB (tuned+ES) | 0.958483 | 7.973810 | 2.845644 | 7.511767e-05 | 1.432982e-06 |
| 3 | Nowcast | LGB (tuned+ES) | 0.957642 | 8.054206 | 2.804757 | 4.359039e-05 | 7.438831e-06 |
| 4 | Nowcast | Ridge (tuned, scaled) | 0.679478 | 22.155637 | 9.645077 | 4.419177e-07 | 5.299279e-08 |

Export to Sheets

***Result: ExtraTrees has the best accuracy in Nowcasting***

# MODEL TRAINING RESULTS

Here is the FORECAST summary formatted as a table.

| Index | Task | Model | R2 | RMSE | MAE |
|---|---|---|---|---|---|
| 0 | Forecast | Ridge (tuned, scaled) | 0.789118 | 15.160917 | 7.907207 |
| 1 | Forecast | ExtraTrees (tuned) | 0.787001 | 15.236816 | 7.877531 |
| 2 | Forecast | RF (tuned) | 0.786498 | 15.254807 | 7.931574 |
| 3 | Forecast | XGB (tuned+ES) | 0.783937 | 15.346031 | 7.912966 |
| 4 | Forecast | Ensemble (LGB+XGB mean) | 0.778839 | 15.526017 | 7.917307 |
| 5 | Forecast | LGB (tuned+ES) | 0.759942 | 16.175737 | 8.006024 |
| 6 | Forecast | Naïve (persist) | 0.734271 | 17.018669 | 8.537225 |

***Result:  Ridge Regression gives the best accuracy in Forecasting***

# FRONTEND

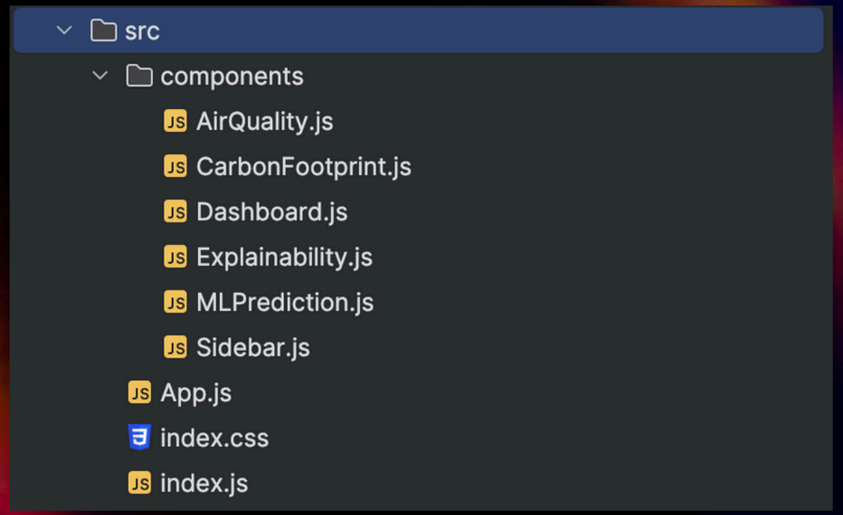
- *src/index.js: This is the JavaScript entry point for our entire app. Its one and only job is to find the index.html file (in the public folder) and tell React to render our main App component inside it.*

- *src/App.js: This is the main component for our dashboard. It's the central hub that defines the overall layout. It imports and organizes all the other components (like the input form, the gauges, and the map) into one cohesive page.*

- *src/index.css: This is the global stylesheet. Any CSS rules we write in this file (like setting a default background color or font) will apply to our entire application.*

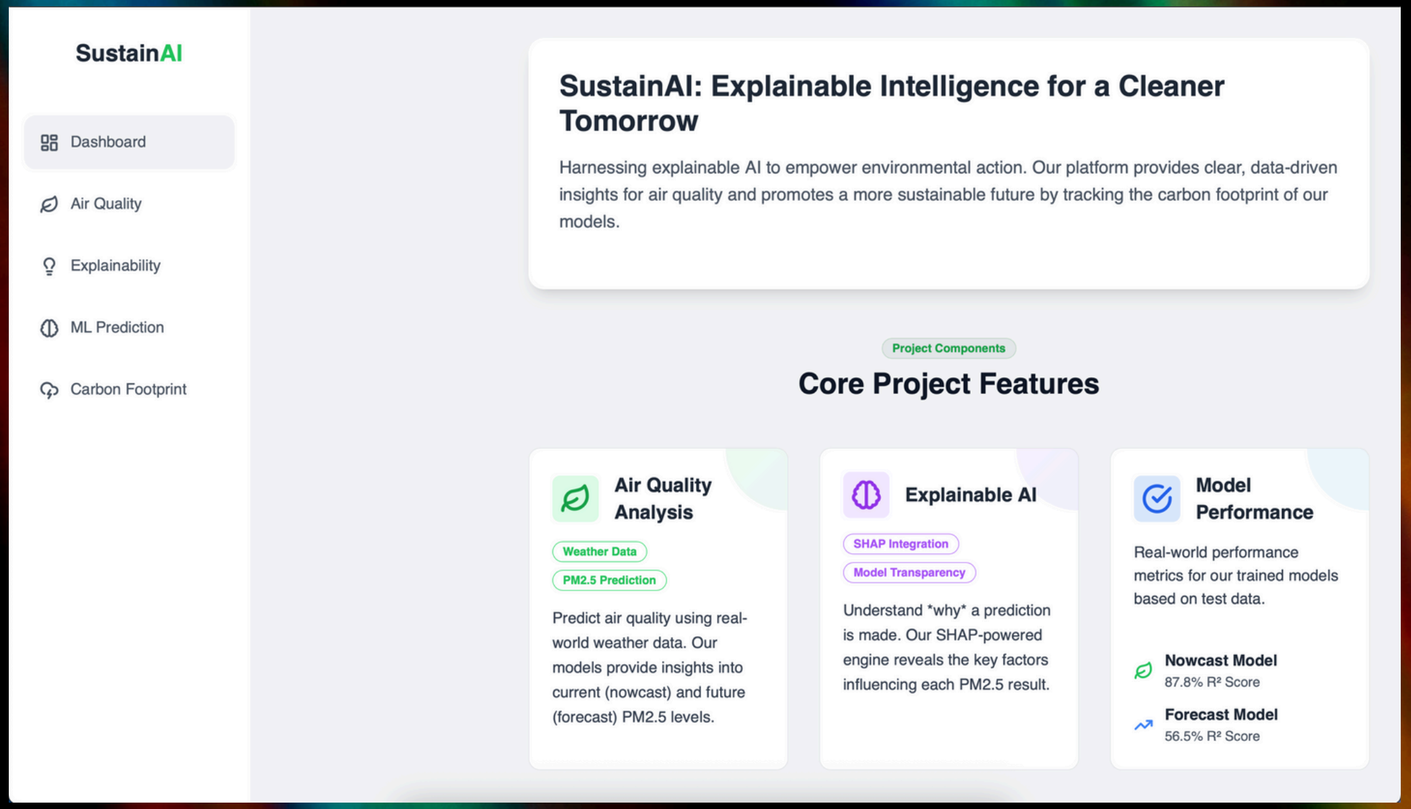
- *src/components/ (Folder): This folder is where we store all our reusable UI pieces. This is where our MLPrediction.js and other js files exist*

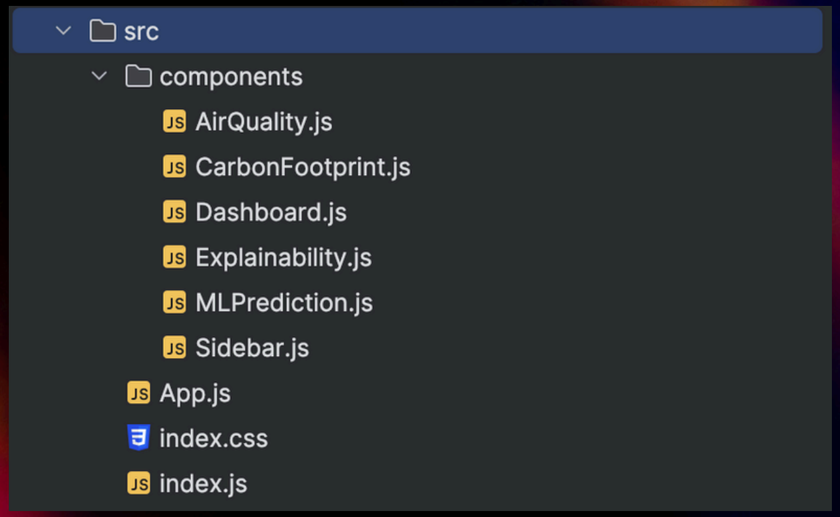
***Frontend Techniques :***
***React.js***
***CSS for Styling***



# FRONTEND

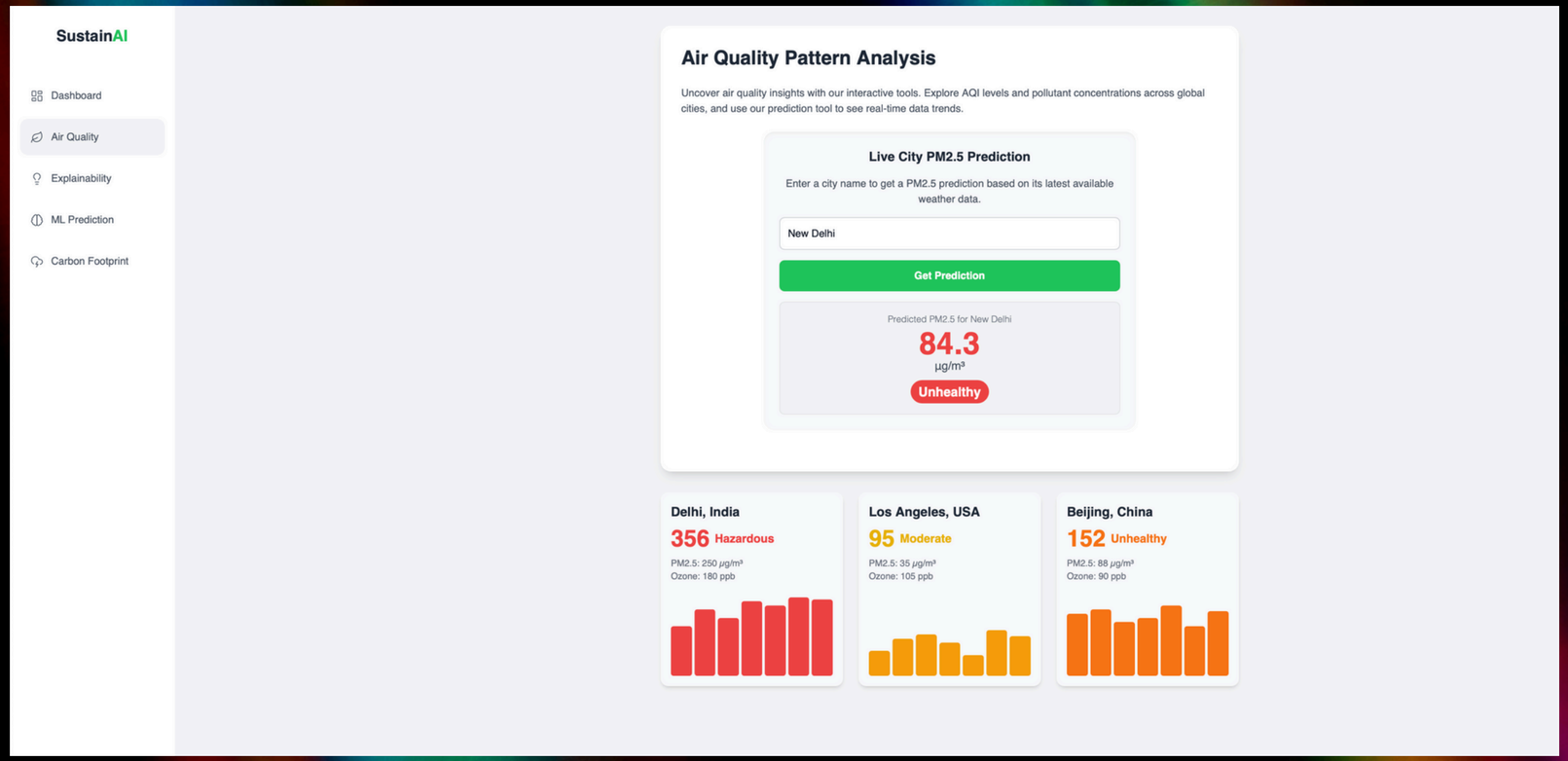
## *The DASHBOARD*

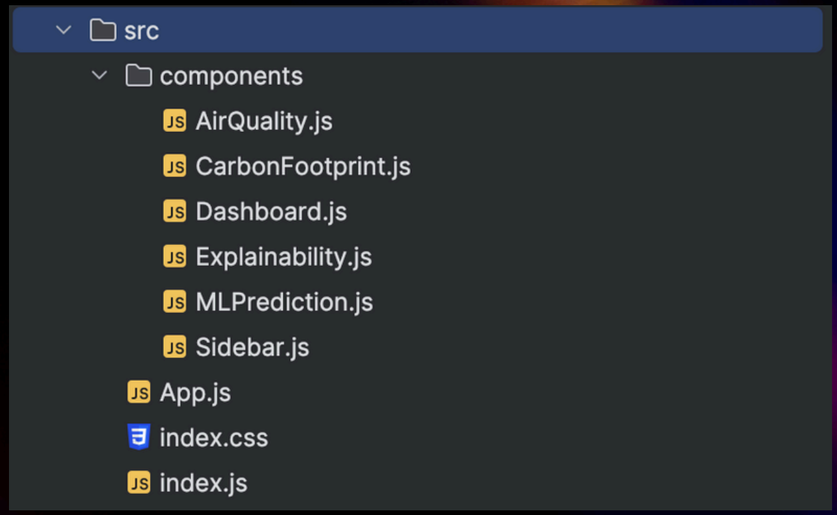


# FRONTEND

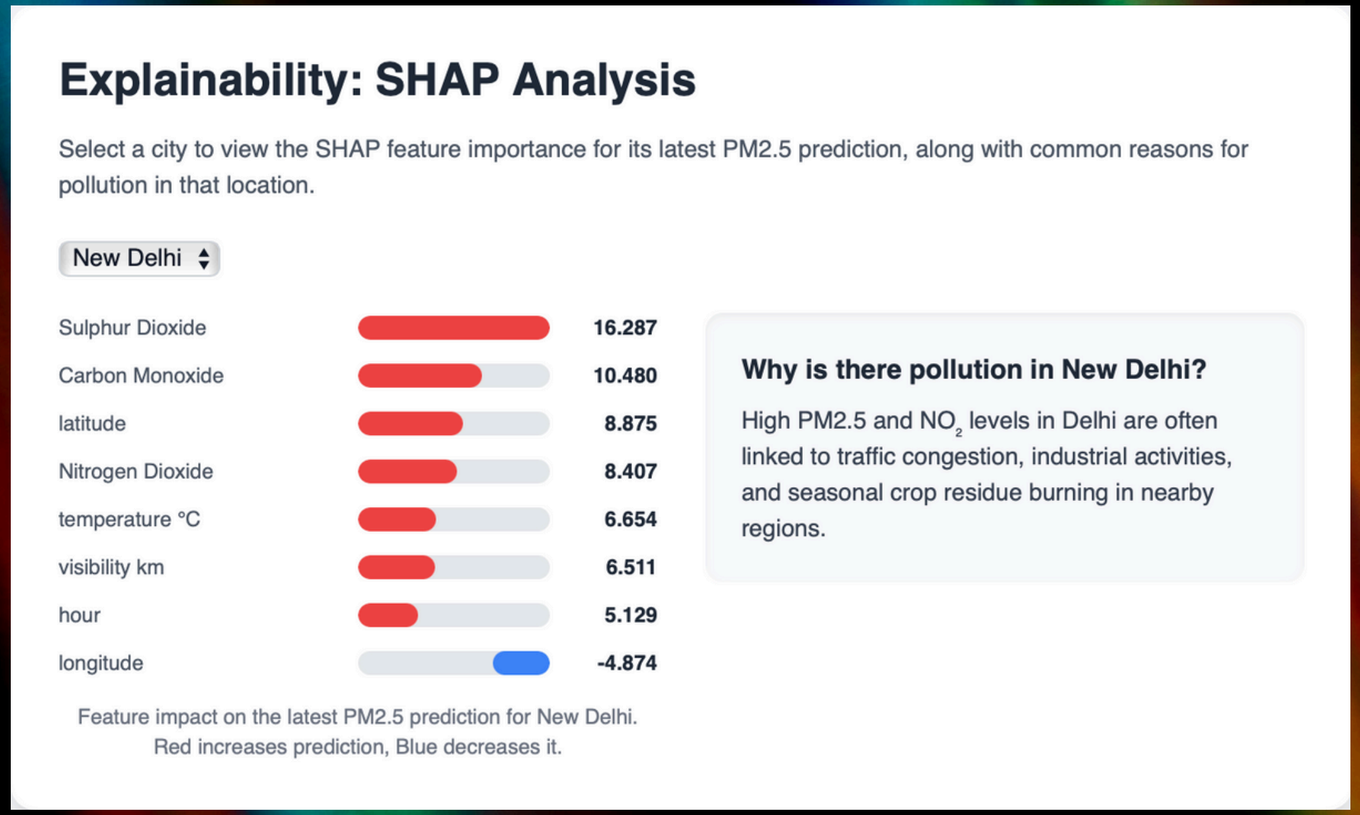
## *Air Quality*

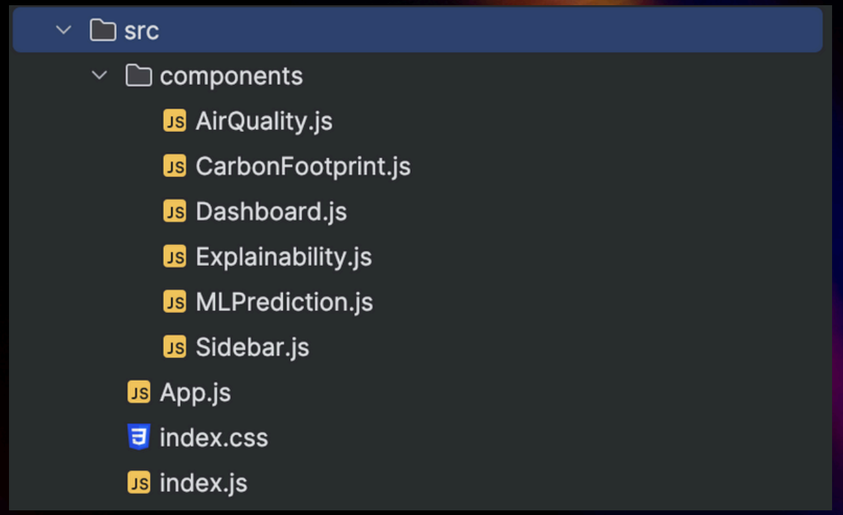


# FRONTEND

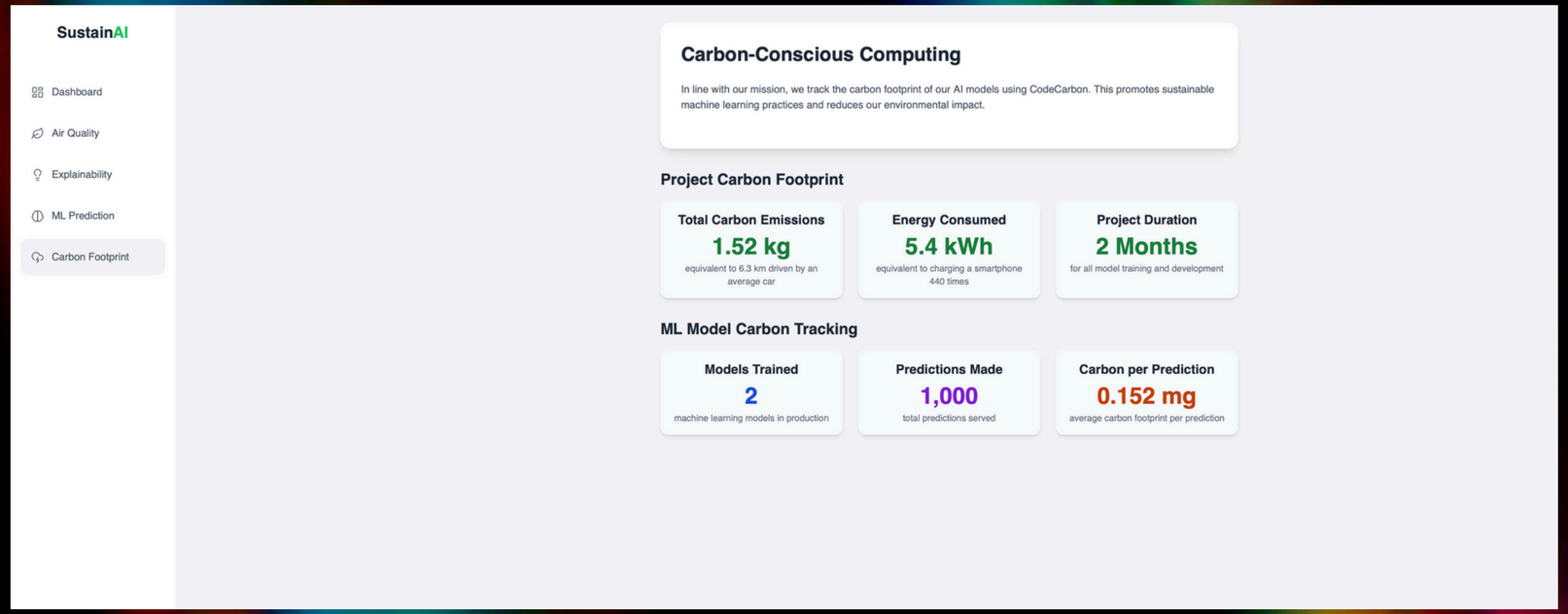
*Explainability*

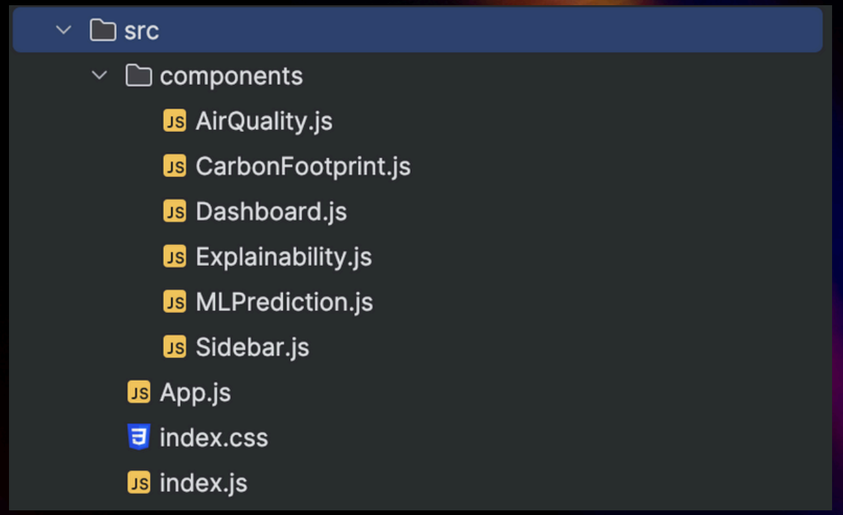


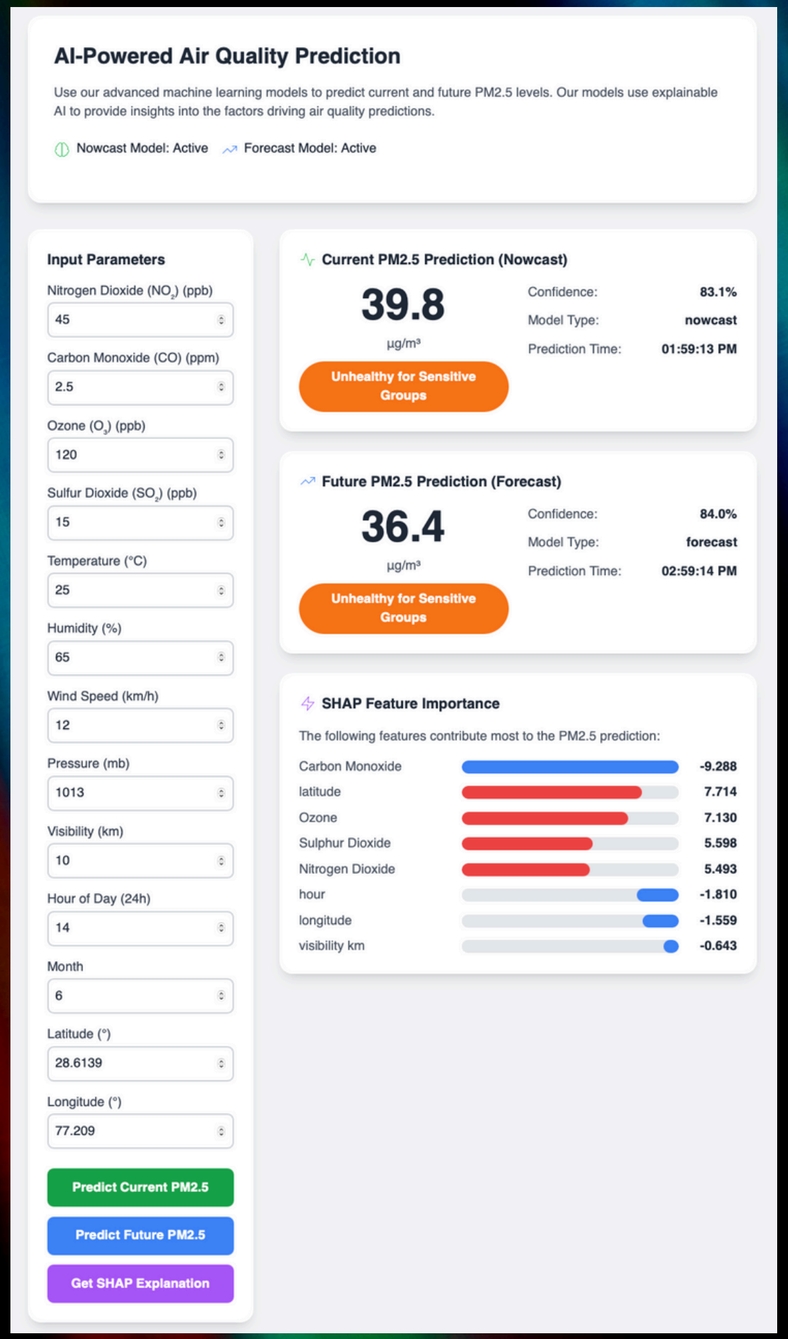



# FRONTEND

## *Carbon Footprint*






# FRONTEND

*ML Prediction - A Core Component of our project*

- *This is a primary user interaction point, with 13 inputs.*
- *These 13 inputs are the UI Features we talked about in the model training.*

# BACKEND

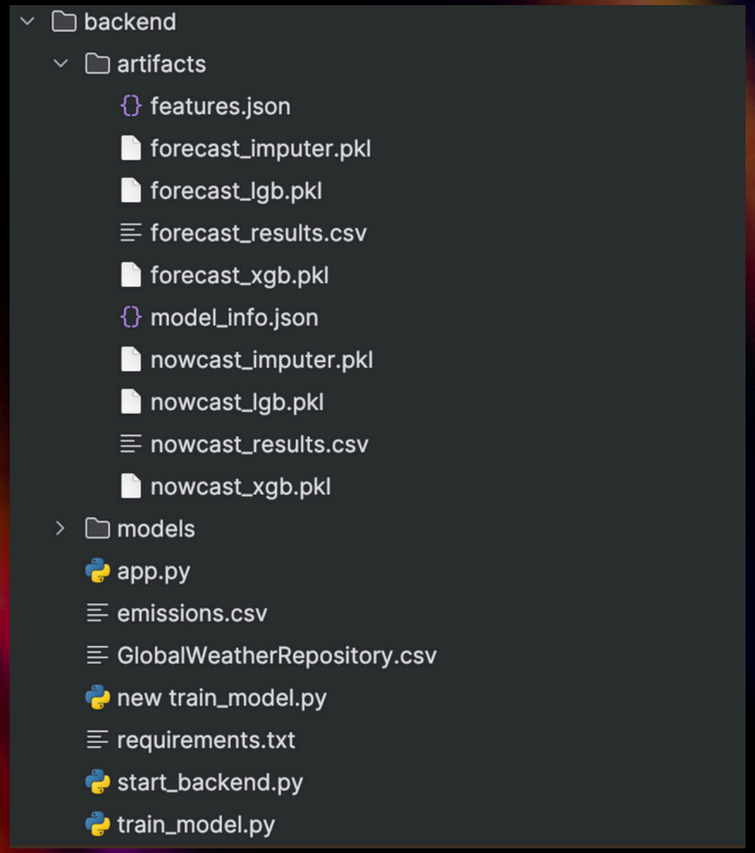
*Backend Techniques :*
*FLASK - python*

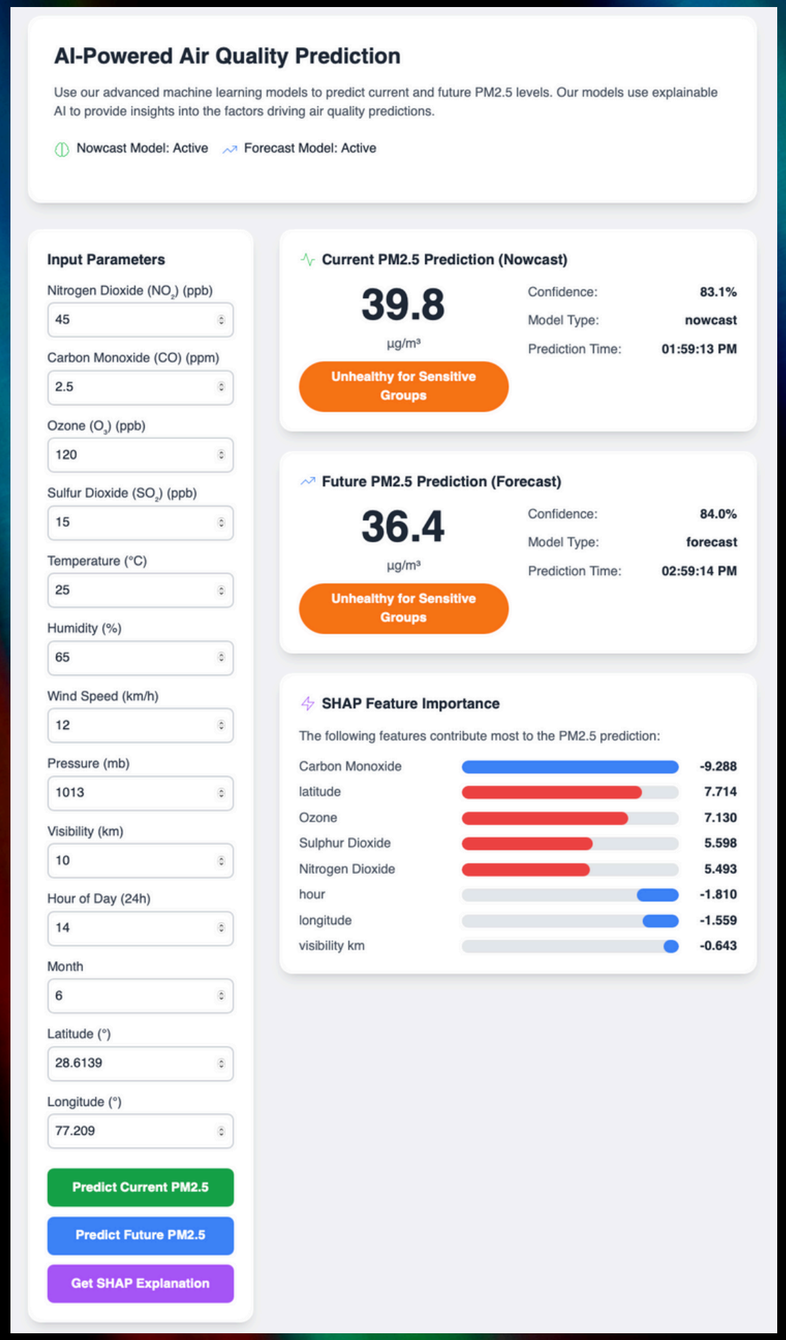
- *app.py: This is our Flask API server. It's the main file that runs the backend. It loads the saved models from /artifacts and creates the /predict API endpoint that our React frontend communicates with.*

- *train_model.py: This is our model training script. We run this file once to read the GlobalWeatherRepository.csv, train the LightGBM and XGBoost models, and generate all the files in the /artifactsfolder. This file is not part of the live server.*

- *GlobalWeatherRepository.csv: The raw data (96,000+ rows) used only by train_model.py to train and test our models.*

- *emissions.csv: This is an output log file generated by the CodeCarbon library we used in our training script. It contains the estimated CO2 emissions from our model training.*






# FRONTEND & BACKEND WORKING HAND IN HAND

***What exactly happens when we click "Predict" on the webpage?***

- ***FRONTEND: React's OnSubmit function is called, all 13 input values are bundled into JSON File.***
- ***BACKEND: FLASK receives the JSON, and loads the XGBoost Models (.pkl under artifacts), and makes the prediction and returns PM2.5 values in a JSON file.***

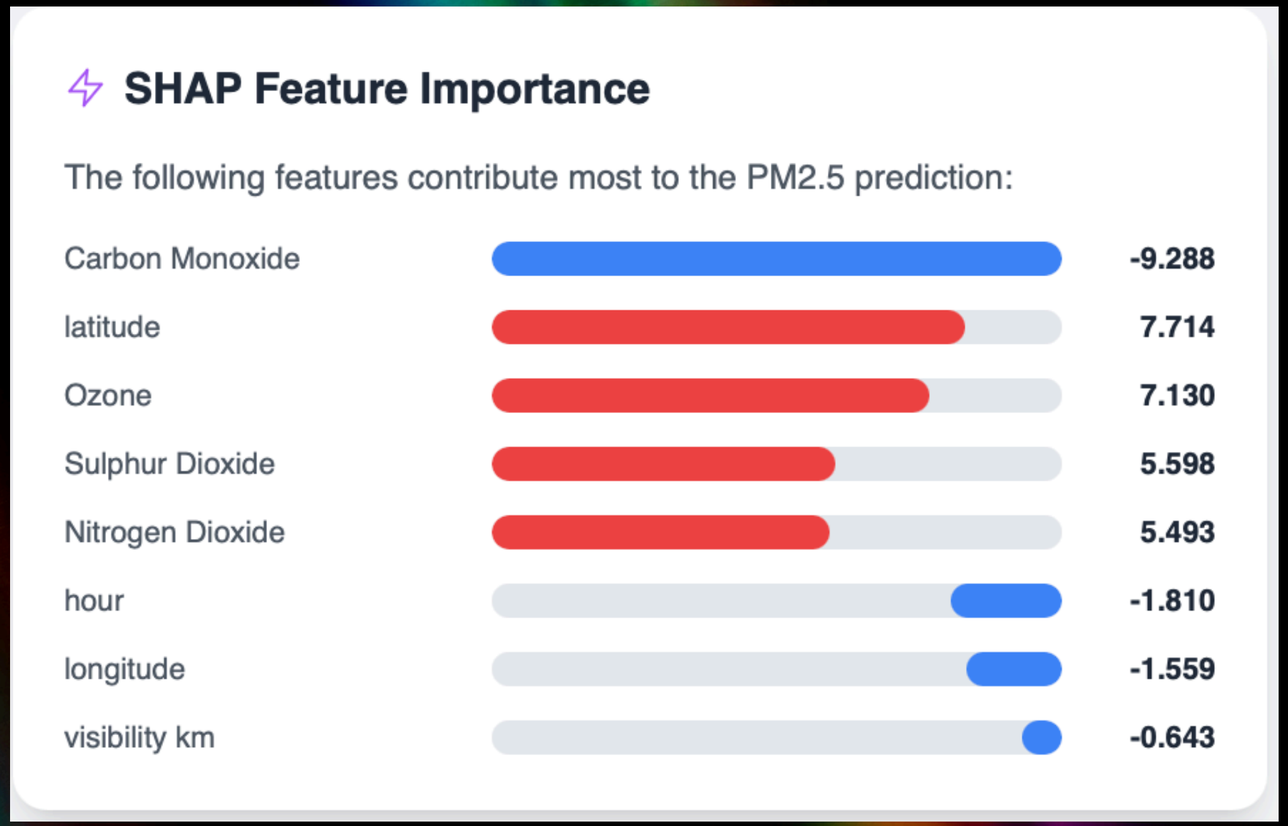
# EXPLAINABILITY

*SHAP - Shapley Additive Explanation:*

- Explainability is showcased using SHAP Graphs.
- SHAP quantifies how each input feature pushes a prediction up or down, enabling transparent reasoning for stakeholders and model debugging.
- A crucial aspect of explainability in predictions is that models can be considered “black boxes,” meaning they provide no insights into their decision-making processes or the underlying reasoning behind their predictions.

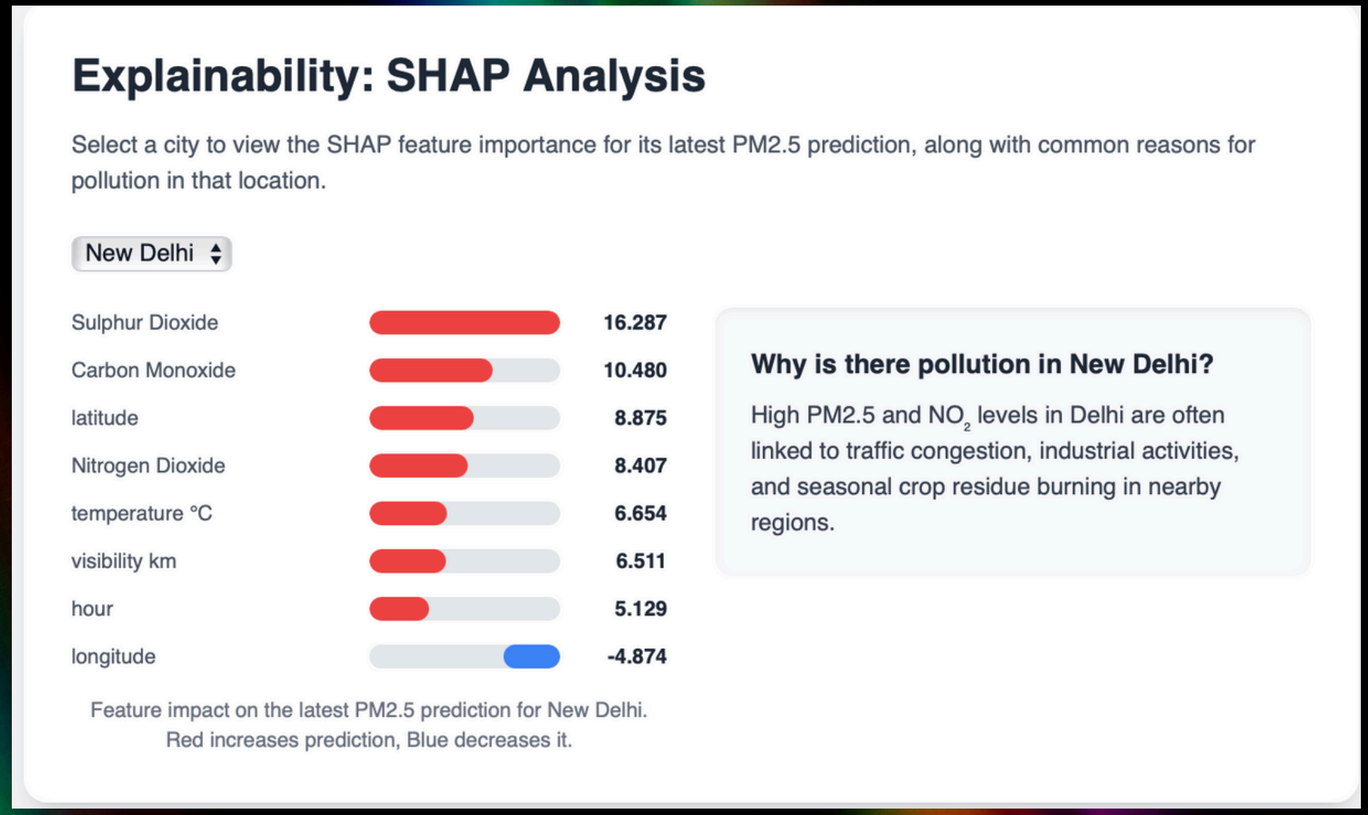
# SHAP PLOTS

**SHAP:** These Plot visualize feature importance for an ML model by showing the average magnitude of each feature's contribution across the entire dataset.



# SHAP PLOTS

**SHAP:** These Plot visualize feature importance for an ML model by showing the average magnitude of each feature's contribution across the entire dataset.



# SUSTAINABILITY

***CODECARBON:***

- Sustainability is showcased using codecarbon in python.
- This makes training footprint visible and comparable across runs, aligning the project with low-carbon ML practices.
- A relatively low carbon footprint shows that the code in itself is sustainable and can be used for large scale training and deployment.

# CODECARBON

***Emission results are automatically stored into emissions.csv whenever a model is trained***

| timestamp | project_name | run_id | experiment_id | duration | emissions | emissions_rate | cpu_power | gpu_power | ram_power | cpu_energy |
|---|---|---|---|---|---|---|---|---|---|---|
| 2025-10-19T19:38:56 | Nowcast_LGB_Infer | 59bef42e-e747-4082-941f-8b6e74bc568f | 5b0fa12a-3dd7-45bb-9766-cc326314d9f1 | 0.2211292500000040 | 2.12508262446884E-06 | 9.61013807295418E-06 | 42.5 | 0.0 | 6.0 | 2.61026883680565E-06 |
| 2025-10-19T19:39:20 | Nowcast_XGB_Infer | 4ea98cd4-b6c9-4ac0-bb31-39472c4d97c9 | 5b0fa12a-3dd7-45bb-9766-cc326314d9f1 | 0.062908415999999 | 6.04362731132435E-07 | 9.6070250939468E-06 | 42.5 | 0.0 | 6.0 | 7.424253104166E-07 |
| 2025-10-19T19:40:43 | Forecast_LGB_Infer | e67b8242-5f47-43d2-adf9-57a09c40a047 | 5b0fa12a-3dd7-45bb-9766-cc326314d9f1 | 0.1053446669999970 | 1.01224255818156E-06 | 9.60886380875443E-06 | 42.5 | 0.0 | 6.0 | 1.2433999750001E-06 |
| 2025-10-19T19:41:03 | Forecast_XGB_Infer | 045f02ac-c250-4f29-84d2-1c0a4fc7147e | 5b0fa12a-3dd7-45bb-9766-cc326314d9f1 | 0.0217529159999969 | 2.08603637953614E-07 | 9.58968618063174E-06 | 42.5 | 0.0 | 6.0 | 2.56367472916688E-07 |
| 2025-10-19T20:37:51 | Nowcast_LGB_Infer | a7f6492d-5af2-4761-a3cf-5f4803d9f01d | 5b0fa12a-3dd7-45bb-9766-cc326314d9f1 | 0.2428324999999920 | 2.33353588788643E-06 | 9.6096522824849E-06 | 42.5 | 0.0 | 6.0 | 2.86632886944437E-06 |
| 2025-10-19T20:38:17 | Nowcast_XGB_Infer | af5be1fb-ad9b-4d20-8688-453ecc30b82a | 5b0fa12a-3dd7-45bb-9766-cc326314d9f1 | 0.0799704579999911 | 7.68314750599305E-07 | 9.60748218547642E-06 | 42.5 | 0.0 | 6.0 | 9.43804972916625E-07 |
| 2025-10-19T20:39:42 | Forecast_LGB_Infer | 3835011b-7df6-43b5-ae7e-714784382057 | 5b0fa12a-3dd7-45bb-9766-cc326314d9f1 | 0.1185936250000170 | 1.13942482871304E-06 | 9.60780841898438E-06 | 42.5 | 0.0 | 6.0 | 1.39965486111117E-06 |
| 2025-10-19T20:40:01 | Forecast_XGB_Infer | 0e5982ba-9d5e-4788-8c56-3b3b71aeb7d2 | 5b0fa12a-3dd7-45bb-9766-cc326314d9f1 | 0.0401562499999954 | 3.85597927869559E-07 | 9.60243867068272E-06 | 42.5 | 0.0 | 6.0 | 4.73764322916538E-07 |
| 2025-10-19T20:44:17 | Nowcast_LGB_Infer | f3d17be2-a94d-423b-a42f-04d5ba992a6e | 5b0fa12a-3dd7-45bb-9766-cc326314d9f1 | 0.2286524579999990 | 2.19733498648723E-06 | 9.60993380830936E-06 | 42.5 | 0.0 | 6.0 | 2.6990421875001E-06 |
| 2025-10-19T20:44:18 | Nowcast_XGB_Infer | b621338d-bbaf-482a-a6d1-ea7dd393f54b | 5b0fa12a-3dd7-45bb-9766-cc326314d9f1 | 0.0760254589999931 | 7.30394471530707E-07 | 9.60723527536707E-06 | 42.5 | 0.0 | 6.0 | 8.97213863888791E-07 |
| 2025-10-19T20:46:01 | Forecast_LGB_Infer | e71cb812-1035-4189-a703-4f8e611555d5 | 5b0fa12a-3dd7-45bb-9766-cc326314d9f1 | 0.1314143750000200 | 1.26270450929678E-06 | 9.60857219232363E-06 | 42.5 | 0.0 | 6.0 | 1.55110881736121E-06 |
| 2025-10-19T20:46:02 | Forecast_XGB_Infer | aec836fc-f72b-412e-ac85-e4edb0acbdb4 | 5b0fa12a-3dd7-45bb-9766-cc326314d9f1 | 0.0377847919999965 | 3.62673124313827E-07 | 9.59838879922536E-06 | 42.5 | 0.0 | 6.0 | 4.45618402777745E-07 |
| 2025-10-19T21:08:00 | Nowcast_LGB_Infer | c3e994f9-e466-4eeb-8ee7-ab90b4309080 | 5b0fa12a-3dd7-45bb-9766-cc326314d9f1 | 0.2179471249999950 | 2.09452958251064E-06 | 9.61026479477849E-06 | 42.5 | 0.0 | 6.0 | 2.57274685E-06 |
| 2025-10-19T21:08:13 | Nowcast_XGB_Infer | 95f02615-4c50-4227-9738-c4a5d5c8e7b4 | 5b0fa12a-3dd7-45bb-9766-cc326314d9f1 | 0.0710487080000064 | 6.82595444952864E-07 | 9.60742938425851E-06 | 42.5 | 0.0 | 6.0 | 8.38515650000045E-07 |
| 2025-10-19T21:09:16 | Forecast_LGB_Infer | 923e3d5c-38cc-4791-b700-7678b598b871 | 5b0fa12a-3dd7-45bb-9766-cc326314d9f1 | 0.0573067919999914 | 5.50440324578148E-07 | 9.60514984992059E-06 | 42.5 | 0.0 | 6.0 | 6.76207961111174E-07 |
| 2025-10-19T21:09:26 | Forecast_XGB_Infer | 72305cdd-799b-4978-be85-7c59320d3c30 | 5b0fa12a-3dd7-45bb-9766-cc326314d9f1 | 0.0193709160000139 | 1.85662348064404E-07 | 9.58459311187298E-06 | 42.5 | 0.0 | 6.0 | 2.28228931249931E-07 |
| 2025-10-23T19:12:35 | Nowcast_LGB_Infer | d1f624f8-04ab-4d69-b191-44561288e62d | 5b0fa12a-3dd7-45bb-9766-cc326314d9f1 | 0.2203444580000000 | 2.11753028211813E-06 | 9.61009095185926E-06 | 42.5 | 0.0 | 6.0 | 2.60098671874988E-06 |
| 2025-10-23T19:12:53 | Nowcast_XGB_Infer | b0d1d263-ca75-4a0c-b6fe-321b0f858f3a | 5b0fa12a-3dd7-45bb-9766-cc326314d9f1 | 0.0759811249999984 | 7.29962111718653E-07 | 9.60715061429622E-06 | 42.5 | 0.0 | 6.0 | 8.96720970138818E-07 |
| 2025-10-23T19:14:07 | Forecast_LGB_Infer | c8d8fb23-d1fa-4573-9a60-055e8c314bc4 | 5b0fa12a-3dd7-45bb-9766-cc326314d9f1 | 0.060068584000021 | 5.76960962926712E-07 | 9.60503685131833E-06 | 42.5 | 0.0 | 6.0 | 7.0881588541665E-07 |
| 2025-10-23T19:14:23 | Forecast_XGB_Infer | 46cda643-73bf-4e7e-920b-a876edb6e213 | 5b0fa12a-3dd7-45bb-9766-cc326314d9f1 | 0.0294661670000960 | 2.82863804576383E-07 | 9.59961316231905E-06 | 42.5 | 0.0 | 6.0 | 3.47580146528021E-07 |

# CODECARBON

*Code Carbon Emission Results, clearly displayed on the final webpage.*

# PROJECT CONTRIBUTION

| | |
|---|---|
| TEJAS | BACKEND |
| HARTHIK | DATASET DETAILS AND MODEL TRAINING |
| BHANU SHARMA | FRONTEND |
| VINAYAK | EXPLAINABILITY |
| SRIDHAR | SUSTAINABILITY |


Thank You
Github repository link:- https://github.com/yoursanonymous/SustainAI/tree/main