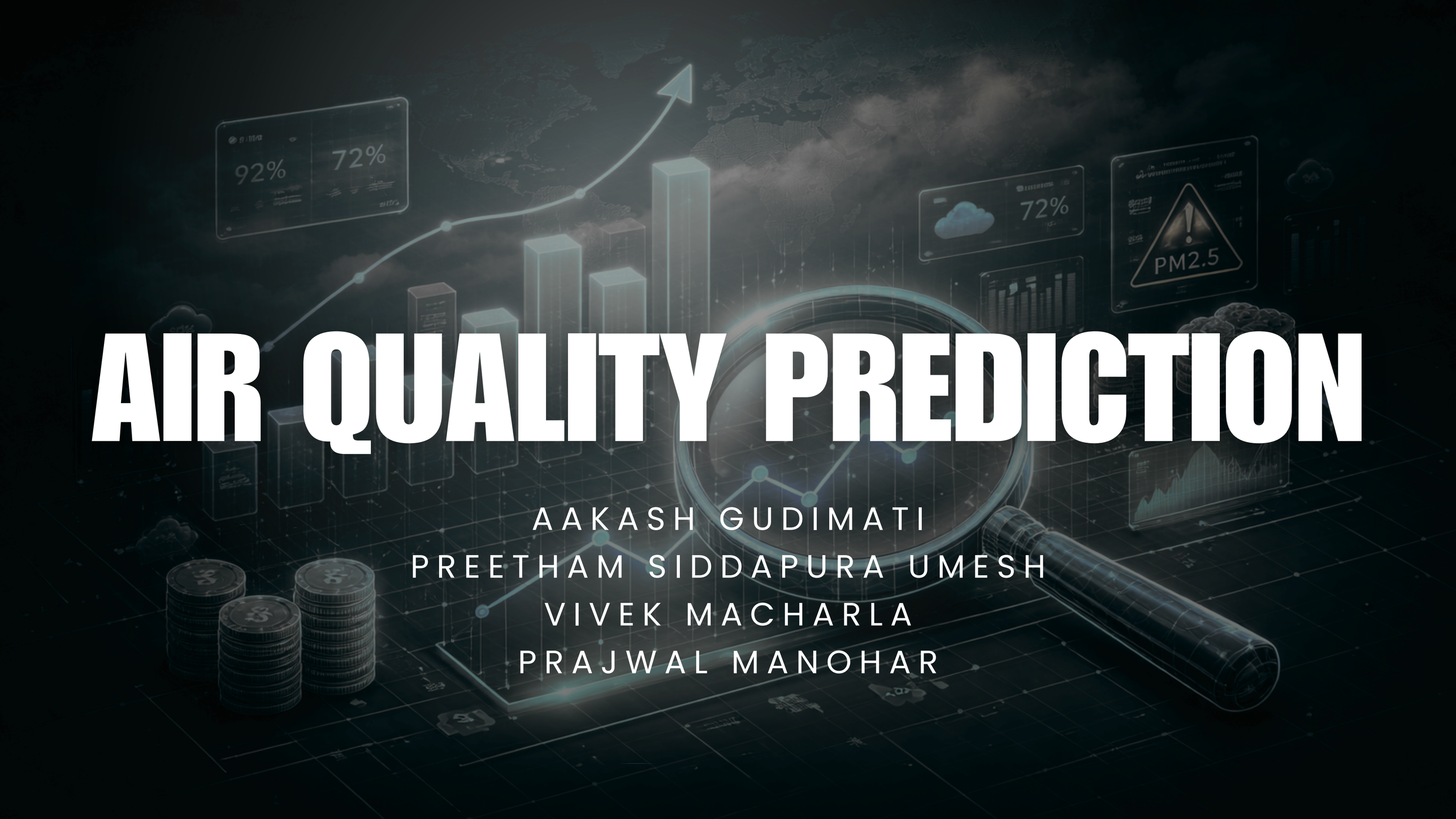
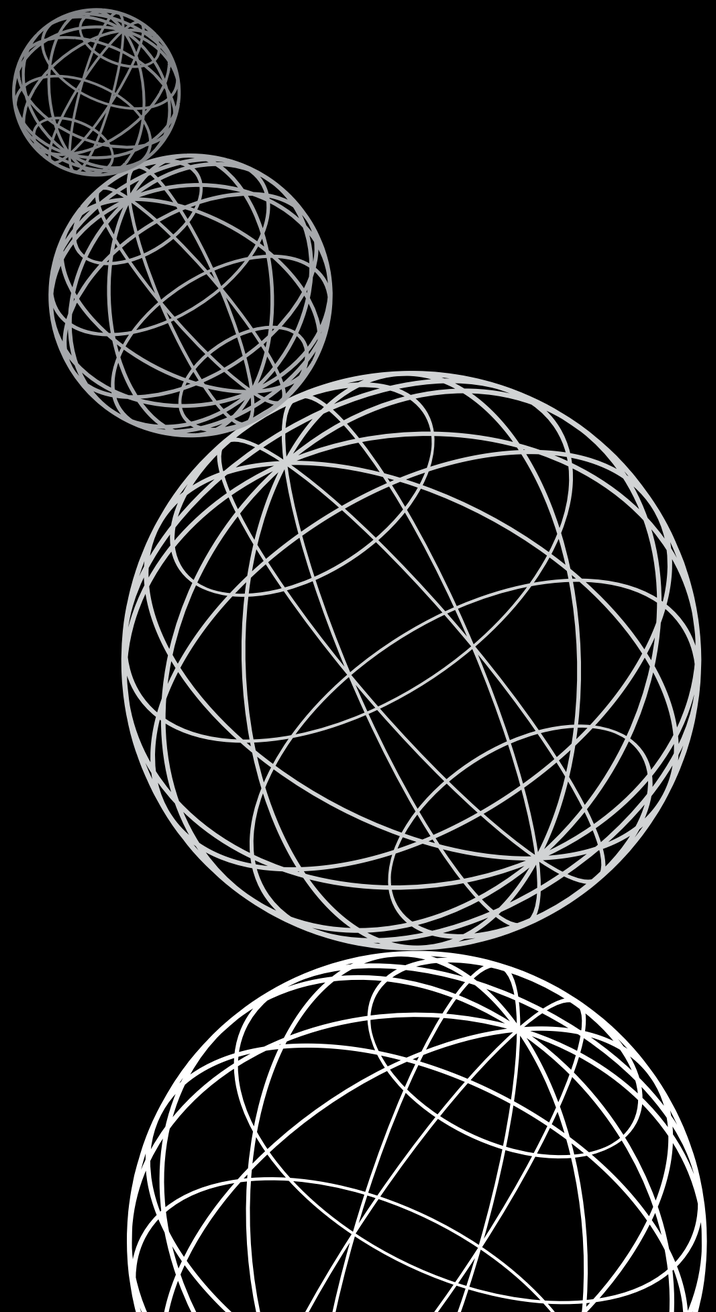


AIR QUALITY PREDICTION

AAKASH GUDIMATI
PREETHAM SIDDAPURA UMESH
VIVEK MACHARLA
PRAJWAL MANOHAR

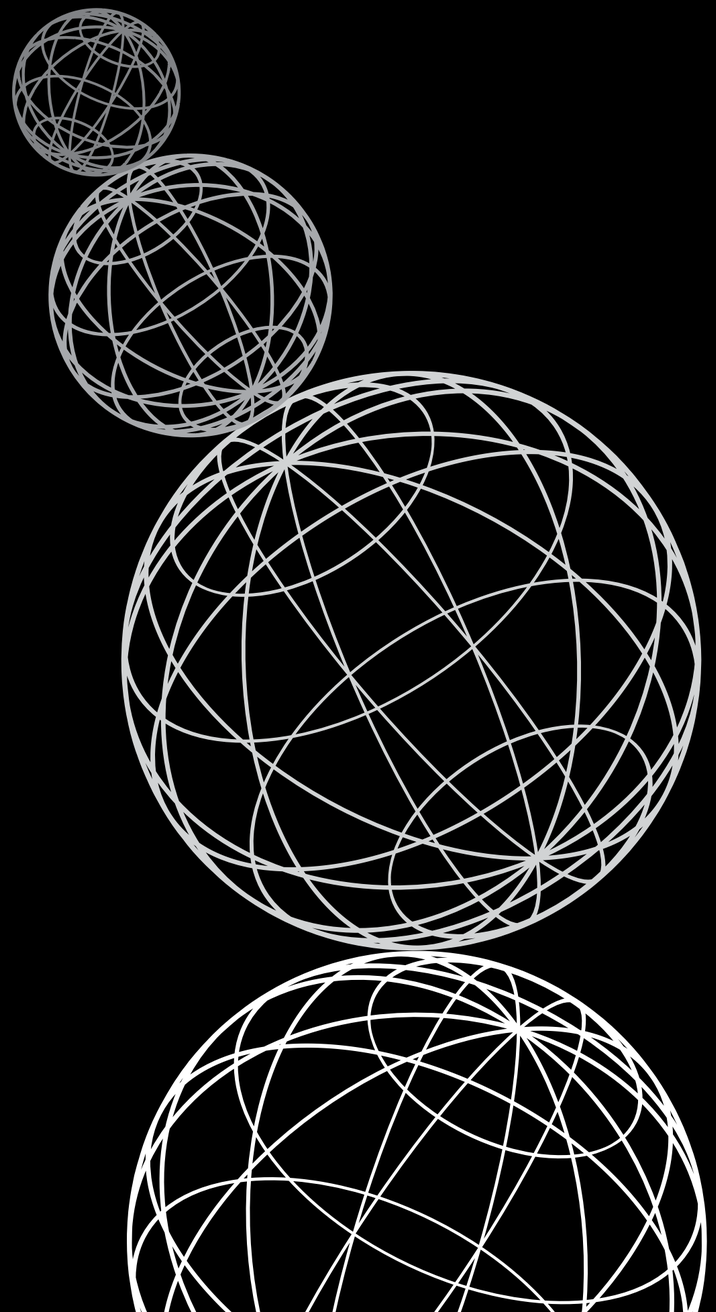
AIR QUALITY

- *Poor Air Quality is an environmental and health issue*
- *Causes: Traffic, Industries, Urbanization*
- *Predicting AQI helps goverments and citizens take preventive actions*



PROBLEM STATEMENT

- *Air quality changes dynamically over time*
- *Manual monitoring cannot predict future pollution levels.*
- *Need for an intelligent system to forecast AQI*



OBJECTIVES

- *Predict future AQI levels using historical data*
- *Apply time-series forecasting techniques*
- *Compare ML & Deep learning models*
- *Visualize trends and build monitoring system*

DATASET OVERVIEW

Dataset Source:

- *Kaggle*

Dataset Charecteristics:

- *400000+ records*
- *Time series structured dataset*
- *Collected from air monitoring stations*

Target Variable:

- *AQI (Air Quality Index)*

DATASET FEATURES

- *Date & Time*
- *PM2.5*
- *PM10*
- *NO_2 (Nitrogen Dioxide)*
- *SO_2 (Sulfur Dioxide)*
- *CO (Carbon Monoxide)*
- *O_3 (Ozone)*
- *Temperature*
- *Humidity*
- *Wind Speed*
- *AQI (Target)*

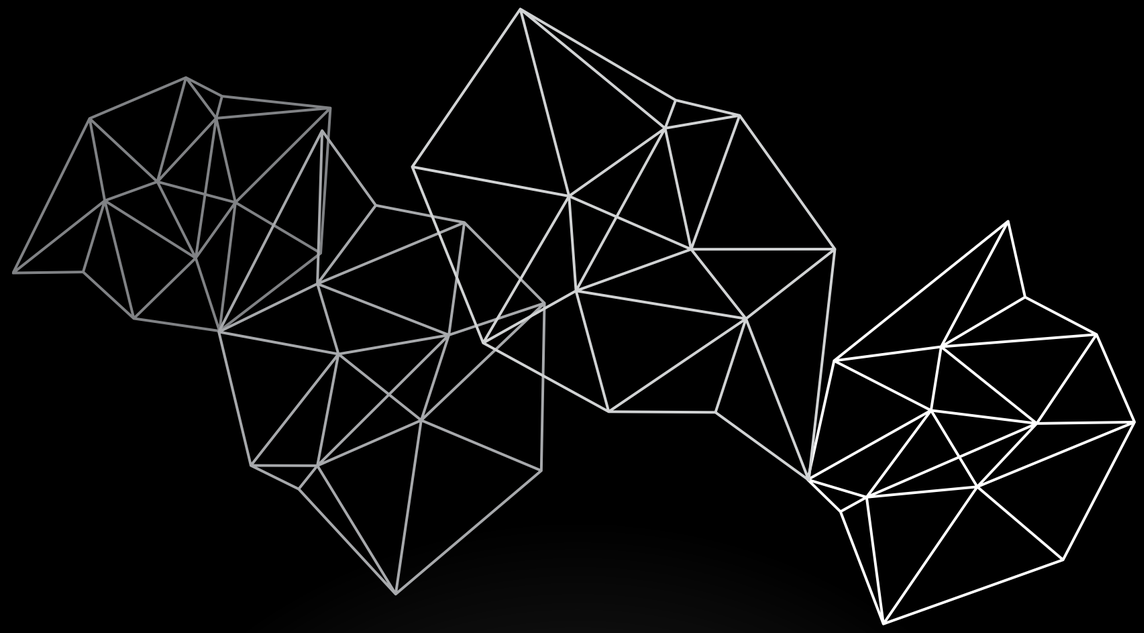
DATA PREPROCESSING

- *Removed duplicate records*
- *Handled missing values using forward fill*
- *Converted date column to datetime format*
- *Set datetime as index*
- *Normalized numerical features*
- *Split data into training and testing sets*

PROJECT METHODOLOGY

Workflow:

1. *Data Collection*
2. *Data Cleaning*
3. *Exploratory Data Analysis*
4. *Feature Engineering*
5. *Model Training*
6. *Model Evaluation*
7. *Visualization*
8. *Real-Time Prediction*



MODELS USED

Machine Learning Model:

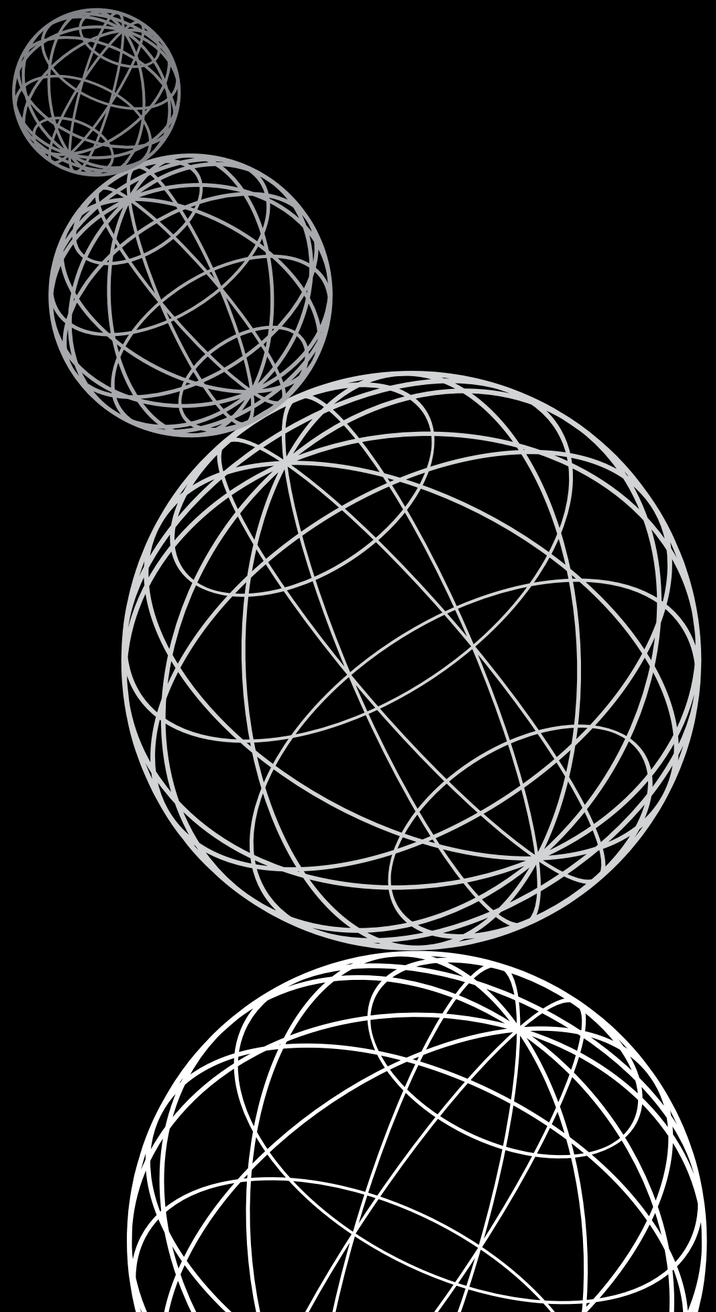
- *Linear Regression*
- *Random Forest*

Deep Learning Model:

LSTM
(Long Short-Term Memory)

FEATURE ENGINEERING

- *Created lag features (previous AQI values)*
- *Calculated rolling averages*
- *Extracted day/month from date*
- *Removed null values after shifting*
- *Improved temporal pattern learning*



MODEL TRAINING

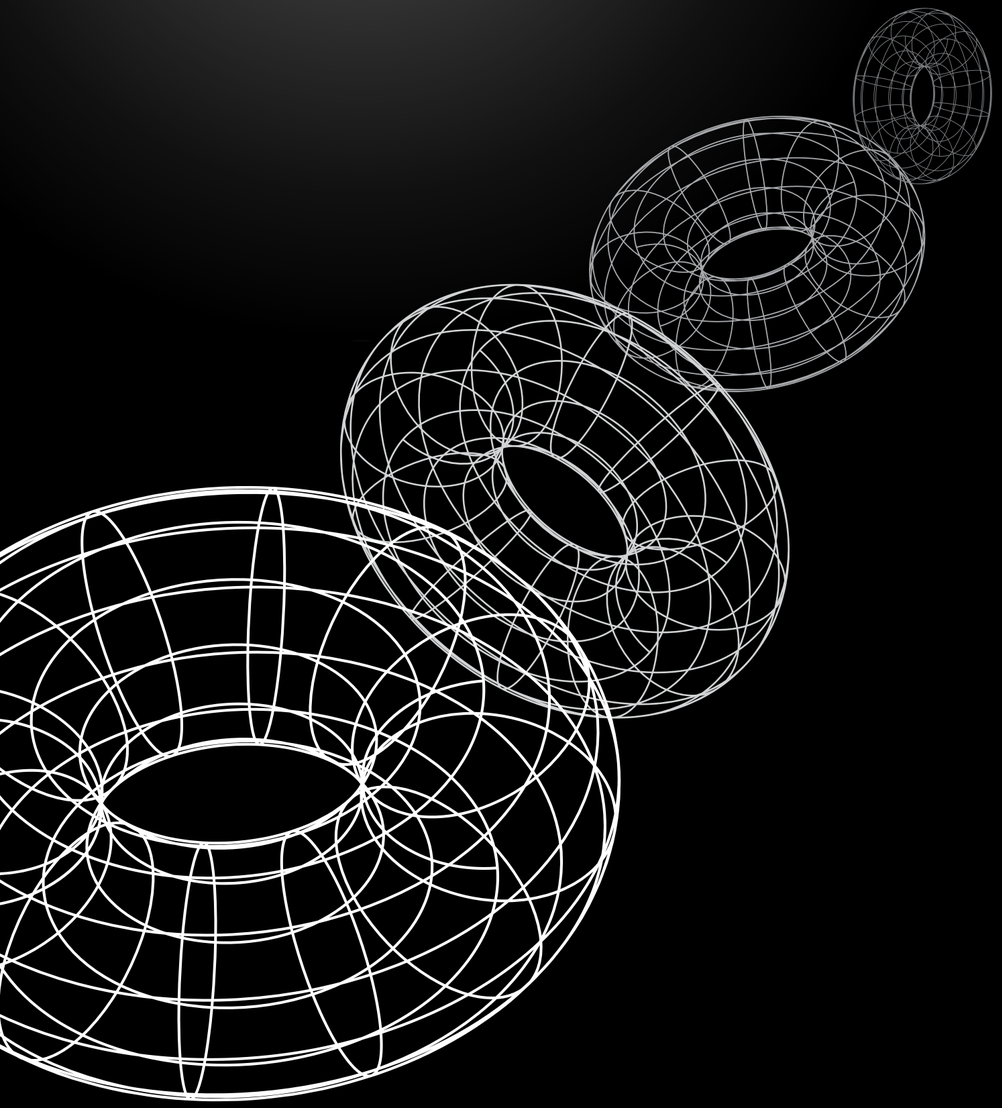
- *Split dataset into Train (80%) and Test (20%)*
- *Trained multiple regression models*
- *Used time-series sequential training*
- *Hyperparameter tuning applied*
- *Models learned pollutant–AQI relationships*

EVALUATION METRICS

- *Mean Absolute Error (MAE)*
- *Mean Squared Error (MSE)*
- *Root Mean Squared Error (RMSE)*
- *R_Square Score*

PURPOSE:

- *Measure prediction accuracy*
- *Compare model performence*



VISUALIZATION

- *AQI Trend Over Time*
- *Actual vs Predicted AQI Graph*
- *Pollutant Correlation Heatmap*
- *Seasonal Pollution Patterns*

Tools Used:

- *Matplotlib*
- *Seaborn / Plotly*

REAL-TIME MONITORING SYSTEM

- *Integrated live AQI data API's*
- *Predict next 24-hours AQI*
- *Display alerts for hazardous levels*
- *Interactive dashboard visualization*
- *Useful for smart city monitoring*

CONCLUSION

- *Machine Learning can accurately predict air quality.*
- *Time-series forecasting improves prediction reliability.*
- *LSTM outperformed traditional models.*
- *Pollution + weather factors strongly influence AQI.*
- *System can support environmental planning.*

THANK YOU