**Machine Learning**

**Project Documentation format**

**1. Introduction**

**Project Title:** Traffic intelligence Advanced Traffic Volume Estimation

**Team Members:** George Mullar,Kambhampati Sruthi, Pravalika

**2. Project Overview**

**Purpose:** Estimate traffic volume using machine learning based on historical and real-time data to help manage urban traffic flow.

**Features:**

Real-time and historical data processing

Regression-based traffic prediction

Performance metrics (MSE, R²)

Output visualization through dashboard/API

**3. Architecture**

**Data Collection:** Describe sources (e.g., traffic cameras, IoT sensors, CSV datasets).

**Preprocessing:** Steps like data cleaning, feature extraction, normalization.

**Model:** Algorithms used (e.g., Linear Regression, Random Forest, XGBoost).

**Deployment:** Tools used for serving the model (e.g., Flask API, Streamlit, FastAPI).

**Visualization:** Dashboard tools or reporting methods used.

**4. Setup Instructions**

**Prerequisites:** Python, Jupyter, scikit-learn, pandas, Flask/Streamlit.

**Installation:**

1. Clone the repo
2. Create virtual environment
3. Install dependencies (pip install -r requirements.txt)
4. Run notebook/API

**5. Folder Structure**

/data — Raw and processed datasets

/notebooks — Jupyter notebooks

/models — Trained model files

/api — Model API scripts

/visualization — Dash or Streamlit dashboard

**6. Running the Application**

**Notebook:** Open Jupyter and run main\_model.ipynb

**Run:** python app.py

**Access at:** http://localhost:5000/predict

**7. Model Documentation**

**Endpoints:** /predict, /train, /metrics

**Input:** Sample features (e.g., hour, weather, lane count)

**Output:** Predicted traffic volume

**8. Authentication**

Authentication is not required for local use.If deployed online, basic API key authentication can be added to secure the prediction endpoint.

**9. User Interface**

A simple and user-friendly dashboard built using Streamlit displays:

Input fields (e.g., time, day, weather)

Predicted traffic volume output

Graphs and charts for trend visualization

**10. Testing**

**Tools Used:**

Python

Scikit-learn

Jupyter notebook

**Validation Methods:**

**Train-Test split:** Data was divided using train\_test\_split() from scikit-learn’

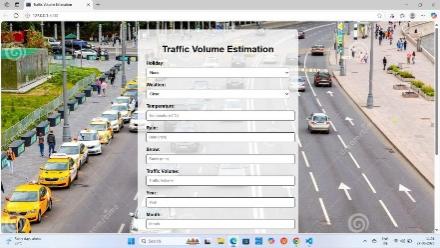
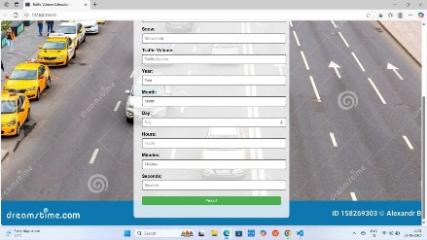
**Evaluation Metrices:**

MSE: 3357277.50

MAE:1620.45

R2 score**:** -88.9427

**11. Screenshots or Demo**

**12. Known Issues**

Limited dataset may affect model accuracy.

Weather or seasonal changes can impact prediction results.

Real-time sensor data may be missing or inaccurate.

Currently uses static data; live data not yet integrated.

Minor delay in prediction due to model/API processing.

Model may not generalize well to all locations.

**13. Future Enhancements**

**Include:**

Live sensor integration

Real-time alerts

Model improvement using deep learning