Machine Learning Mini Project

Traffic Flow Prediction System

<div align="right">

~Dharshini N

</div>

Project Objective:

  The aim of this project is to predict traffic volume on a highway using real-world traffic and weather data and to classify traffic flow into categories: **No/Light Traffic, Moderate Traffic, and Heavy Traffic**. This system helps commuters, traffic authorities, and city planners to anticipate traffic congestion and plan accordingly.

Dataset:

The project uses the **Metro Interstate Traffic Volume dataset** from the **UCI Machine Learning Repository**, which contains hourly traffic data with weather conditions and holiday information.

Key columns used:

- date_time → Timestamp of the observation

- weather_main → Weather condition (Clear, Clouds, Rain, Snow, Fog)

- holiday → Indicator of nearby event/holiday (used as "Nearby Event Occurring")

- traffic_volume → Target variable (vehicles per hour)

Features selected for the model (4 features):

1. **Hour of day** (hour) – extracted from date_time

2. **Day of the week** (day_of_week) – extracted from date_time

3. **Weather condition** (weather_encoded) – categorical weather encoded as numeric

4. **Nearby event occurring** (event_val) – binary feature based on holiday

Feature Engineering:

- Extracted **hour** and **day of week** from date_time.

- Encoded **weather condition** as numeric categories.

- Converted **holiday column** into a binary feature representing the presence of a nearby event.

Model Used:

- **Algorithm:** Multiple Linear Regression (MLR)

- **Reason:** Simple, interpretable, and effective for predicting continuous traffic volume based on multiple features.

- **Target:** traffic_volume (vehicles per hour)

User Interface:

- **User selects:**
  - Hour of the day
  - Day of the week
  - Weather condition
  - Nearby event occurrence (Yes/No)
- On clicking **Predict**, the app shows:
  - **Predicted traffic volume** (vehicles/hour)
  - **Traffic flow category** (No/Light, Moderate, Heavy)

Tools and Libraries:

- **Python** – Programming language
- **Pandas** – Data manipulation
- **Scikit-learn** – Machine learning (Linear Regression)
- **Streamlit** – Interactive web application for predictions
- **Pickle** – Saving and loading trained model

Key Benefits:

- Helps commuters plan travel according to predicted traffic flow.
- Assists traffic authorities in **real-time monitoring and traffic management**.

- Easy-to-use interactive interface for non-technical users.
- Can be extended to include more features like temperature, precipitation, or real-time sensor data.

Prediction Output:

- The model predicts numeric traffic volume.
- The numeric output is then mapped to a traffic flow category:

Traffic Volume (vehicles/hour)	Traffic Flow Category
≤ 2000	No/Light Traffic
2001 – 4000	Moderate Traffic
> 4000	Heavy Traffic

🚦 Traffic Flow Prediction System 🔗

Predict traffic volume using Multiple Linear Regression

Hour of Day

5	⌄

Day of Week

Sunday	⌄

Weather Condition

Rain	⌄

Is there a nearby event or holiday?

○ No
🔴 Yes

🚚 Predict Traffic Volume

📊 Predicted Traffic Volume: **2177 vehicles/hour**

🚗 Traffic Condition: **Moderate Traffic**