

# Problem Statement:

The objective of this project is to develop a predictive model that accurately forecasts hourly traffic volumes at various road junctions using historical traffic data and other influencing variables. Urban traffic congestion poses significant challenges for transportation efficiency, public safety, and economic productivity. Accurately forecasting traffic volumes can assist city planners, traffic control centers, and mobility platforms (e.g., Uber, Ola) in making data-driven decisions for route planning, surge pricing, and congestion mitigation.

To achieve this objective, the project involves the following tasks:

## 1. Data Cleaning and Preprocessing

- Load raw traffic datasets and clean missing, duplicated, or incorrectly formatted data.
- Convert traffic data into hourly aggregates for each junction.
- Standardize or normalize features for comparability.

## 2. Feature Engineering and Selection

- Generate time-based features such as hour of the day, day of the week, and month.
- Create lag variables to capture historical trends and temporal dependencies.
- Incorporate external features such as weather conditions (temperature, rainfall, wind speed) and local events (concerts, sports, public holidays).

## 3. Model Development

- Explore various forecasting algorithms:
  - Time series models: ARIMA, SARIMA
  - Machine learning models: Random Forest, Gradient Boosting, XGBoost

- Deep learning models: LSTM, GRU for sequential dependencies
- Use time-based train-test splitting to preserve temporal ordering.
- Perform hyperparameter tuning using grid/random search.

#### **4. Model Evaluation**

- Evaluate model accuracy using metrics such as MAE, RMSE, and  $R^2$ .
- Use time-based cross-validation to test model generalization across different time windows.

#### **5. Pattern Analysis and Peak Hour Detection**

- Identify consistent peak traffic hours across junctions using congestion metrics.
- Analyze variation by weekdays, weekends, and special event days.

#### **6. Visualization and Reporting**

- Create heatmaps and time series plots to visualize traffic trends and forecasted volumes.
- Generate actionable insights and a final report to support traffic management strategies.

#### **Goal:**

Build a robust, scalable model that can predict hourly traffic volumes across multiple road junctions with high accuracy, while uncovering key temporal and environmental factors that influence traffic flow.

Let me know if you'd like this tailored for a presentation, report intro, or business proposal format.