**PROJECT DESIGN**

**Data Collection**

|  |  |
| --- | --- |
| Date | 28 JUNE 2025 |
| Team ID | LTVIP2025TMID59766 |
| Project Name | **TrafficTelligence: Advanced Traffic Volume Estimation with Machine Learning** |
| Maximum Marks | 4 Marks |

**Data Collection:**

1. **Identify Data Requirements**
   * Define what kind of data is needed based on the problem.
   * For example, for a traffic prediction project, data may include:
     + Traffic volume
     + Time and date
     + Weather conditions
     + Road conditions
     + Accident reports
     + GPS data
2. **Data Sources**
   * **Public Datasets:** Government open data portals, Kaggle, UCI ML Repository.
   * **Sensors/IoT Devices:** Real-time data from traffic sensors, cameras, GPS, etc.
   * **APIs:** Google Maps API, OpenWeatherMap API, etc.
   * **Manual Entry:** Surveys, questionnaires.
   * **CSV Files/Databases:** Pre-recorded datasets stored locally or in cloud databases.
3. **Data Acquisition Tools**
   * **Web Scraping:** BeautifulSoup, Scrapy.
   * **APIs Integration:** REST APIs using Python (requests, pandas).
   * **Database Connections:** SQL, MongoDB, Firebase.
   * **Sensor Integration:** IoT platforms, MQTT protocols.
4. **Format of Collected Data**
   * CSV, JSON, Excel files, database tables.
   * Structured format with rows and columns.
5. **Data Storage**
   * Local file systems.
   * Cloud storage (AWS S3, Google Drive, Azure Blob Storage).
   * Databases (SQL, NoSQL).

| **Attribute** | **Description** |
| --- | --- |
| Timestamp | Date and time of observation |
| Vehicle Count | Number of vehicles observed |
| Road Name | Specific road/area name |
| Weather | Sunny, Rainy, Cloudy |
| Temperature (°C) | Ambient temperature |
| Event/Accident | Yes/No |
| Day of Week | Monday, Tuesday, ... |
| Speed Avg (km/h) | Average vehicle speed |

**Example:**