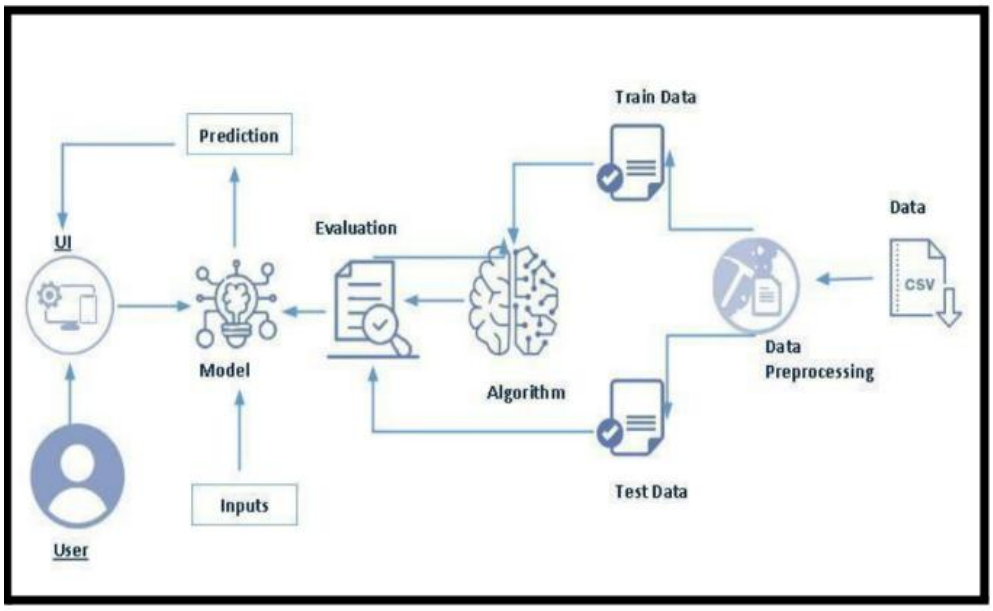
Project Design Phase
Solution Architecture

Date	27 June 2025
Team ID	LTVIP2025TMID59638
Project Name	TrafficTelligence: Advanced Traffic Volume Estimation with Machine Learning
Maximum Marks	4 Marks

Solution Architecture:

- The solution architecture of TrafficTelligence serves as a bridge between the real-world problem of traffic congestion and the technology-driven solution of predictive traffic volume estimation. It defines the structure, behavior, and core components of the system, ensuring alignment with the project's goals.
- The architecture is designed around a modular data pipeline that begins with the ingestion of historical traffic datasets, followed by preprocessing and feature engineering to clean and structure the data.
- These processed datasets are then fed into machine learning models that are trained to analyze traffic patterns and predict future congestion levels. The system includes a user-facing dashboard for stakeholders to visualize predictions and insights in an intuitive manner.
- The architecture supports scalability, allowing the solution to adapt to various regions by retraining models with localized data. Clearly defined development phases—data collection, model training, testing, and visualization—ensure a structured delivery process.
- This solution blueprint ensures that the platform is maintainable, efficient, and aligned with the goal of enabling smarter traffic management using historical data alone.

Solution Architecture Diagram:



1. **Data Collection Phase:**

    - **Inputs**: Date, Time, Climate, Temperature, Holiday.

2. **Data Preprocessing**:
    a. Handle missing data.
    b. Encode categorical variables (like Holiday).
    c. Normalize/scale numerical features (Temperature, Time, etc.).
    d. Feature Engineering (e.g., extract hour from time).

3. **Splitting the Data:**
    - Train-Test Split
4. **Model Selection and Training:**

    - **Libraries**: Pandas, NumPy, Matplotlib, Sklearn
    - **Models**: Regression models (Linear, Random Forest)
    - **Linear Regression**:

$$y=\beta0+\beta1x1+\beta2x2+\cdots+\beta nxn+\varepsilon y$$

    - **Random Forest**.

5. Evaluation:

Metrics Used:

- MAE: Mean Absolute Error

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i|$$

- MSE: Mean Squared Error

$$\mathrm{MSE} = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2$$

- R^2 Score:

$$R^2 = 1 - \frac{SS_{res}}{SS_{tot}}$$

6. Prediction and UI: (UI → Model → Prediction)

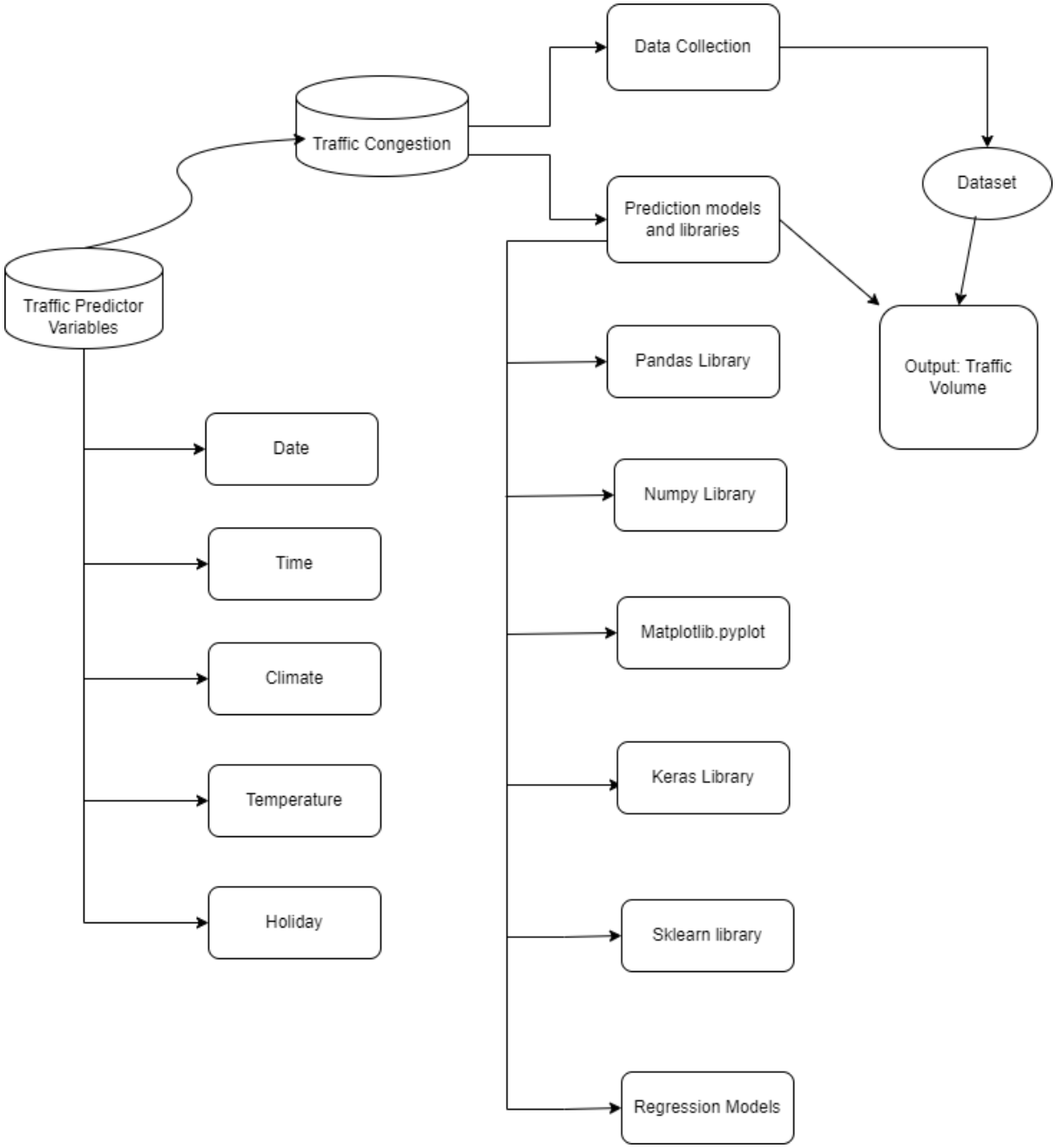
- **UI**: Takes real-time input (Date, Time, Temp, etc.)
- **Backend**: Sends it to the trained model.
- **Output**: Predicted Traffic Volume shown to the user.

7. final Output:

- Result: A numeric prediction of the expected traffic volume at the given inputs.

Data Flow Diagrams:

A Data Flow Diagram (DFD) is a traditional visual representation of the information flows within a system. A neat and clear DFD can depict the right amount of the system requirement graphically. It shows how data enters and leaves the system, what changes the information, and where data is stored.


Data Collection
Traffic Congestion
Dataset
Prediction models and libraries
Traffic Predictor Variables
Pandas Library
Output: Traffic Volume
Date
Numpy Library
Time
Matplotlib.pyplot
Climate
Keras Library
Temperature
Sklearn library
Holiday
Regression Models