

Project Design Phase
Solution Architecture

Date	23 June 2025
Team ID	LTVIP2025TMID45739
Project Name	TrafficTelligence: Advanced Traffic Volume Estimation with Machine Learning
Maximum Marks	4 Marks

Solution Architecture:

- **CSV Data** – A structured file that contains historical green taxi trip data with fields like trip distance, duration, pickup hour, and more.
- **Data Preprocessing** – The step where raw data is cleaned, missing values handled, and new features like is_weekend or is_rush_hour are added for better model accuracy.
- **Train/Test Data** – The processed data is split into a training set (to build the model) and a test set (to check how well the model performs on unseen data).
- **Algorithm** – A machine learning algorithm (Linear Regression) is used to learn the relationship between input features and trip duration.
- **Evaluation** – After training, the model is evaluated using metrics like Mean Squared Error (MSE) and R^2 score to measure prediction performance.
- **Model** – The final trained model is capable of making reliable trip duration predictions based on new user inputs.
- **Inputs** – The user provides input data such as trip distance, pickup hour, and vehicle type to get a personalized duration estimate.
- **UI** – A user-friendly interface built using Streamlit that allows users to input data, view predictions, and explore visual insights.
- **Prediction** – The model processes the input and returns the estimated trip duration along with traffic level and vehicle volume indicators.

Solution Architecture Diagram:

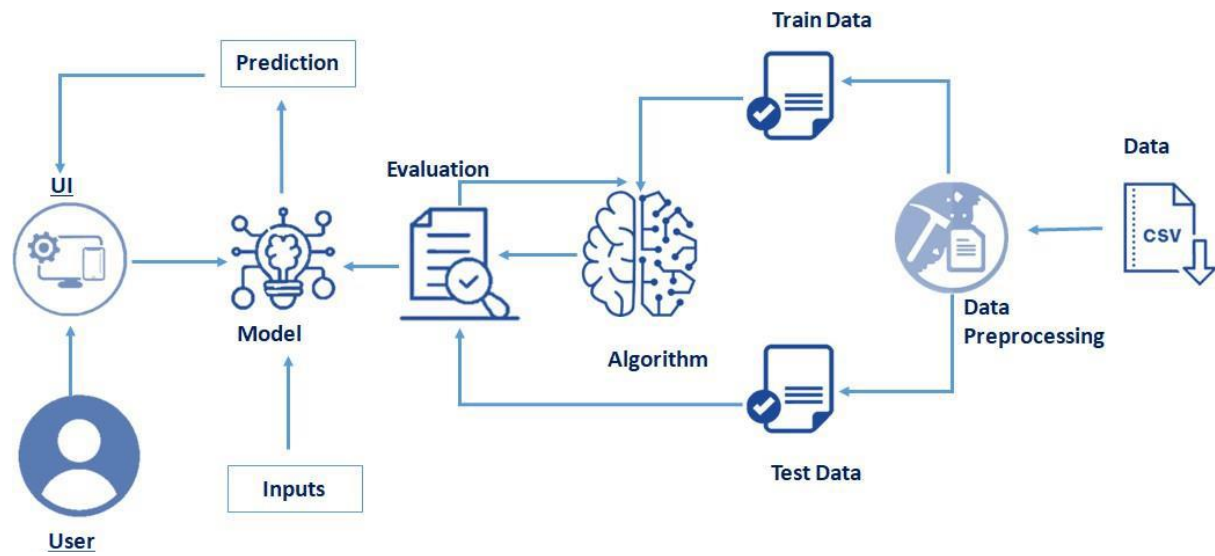


Figure 1: Architecture and data flow of the voice patient diary sample application