

# **Traffic Volume Estimation Using Machine Learning**

## **Submitted By**

**Group Members:**

KURMALA KAVYA [22F01A0582]

MIDDIBOINA NANDINI KUMARI[22F01A0590]

KOTA DINESH VENKAT[22F01A0576]

MIRIYALA PRUDHVI RAJ[22F01A0592]

**Institution:** St. Ann’s college of engineering and technology  
**Organization:** SmartBridge

## **1. Project Overview**

Traffic volume estimation is crucial for urban planning, traffic management, and reducing congestion. This project utilizes machine learning techniques to predict traffic volume based on various environmental and time-based parameters such as temperature, rain, snow, weather condition, holiday indicator, and timestamp (year, month, day, etc.). The model was deployed as a web application using Flask to allow easy interaction and input via a user-friendly form.

## **2. Objectives**

* To collect and preprocess relevant data for traffic volume prediction.
* To build and train a machine learning model for predicting traffic volume.
* To deploy the model using Flask as a web application.
* To make the UI interactive and visually appealing with a responsive background.

## **3. Tools & Technologies Used**

* **Programming Language:** Python
* **Libraries:** pandas, numpy, sklearn, matplotlib, Flask
* **Model Deployment:** Flask Web Framework
* **Frontend:** HTML, CSS
* **Data Visualization:** matplotlib (optional, for debugging and analysis)
* **Platform:** Jupyter Notebook / Anaconda / VS Code

## **4. Dataset Description**

The dataset contains the following features: - holiday: Whether the day is a holiday or not (encoded numerically) - temp: Temperature in degrees - rain: Rainfall amount - snow: Snow level - weather: Encoded weather condition - year, month, day: Date of the record

**Target Variable:** traffic\_volume (number of vehicles passing during a specific time frame)

## **5. Workflow**

1. **Data Preprocessing**
   * Handled missing values
   * Encoded categorical features
   * Feature selection
2. **Model Training**
   * Chosen algorithm: Linear Regression / Decision Tree / Random Forest (based on experimentation)
   * Trained on historical traffic data
   * Evaluated using metrics like MAE, RMSE
3. **Scaling**
   * Used StandardScaler to normalize feature values
   * Saved the scaler using pickle
4. **Model Serialization**
   * Saved trained model and scaler as model.pkl and scale.pkl using pickle
5. **Flask App**
   * Created Flask routes for home and prediction
   * Collected user input from HTML form
   * Used scaler and model to predict output
   * Displayed prediction on the same page
6. **Frontend UI**
   * Used HTML & CSS for designing the input form
   * Set background to a high-quality image of a car to enhance user experience

## **6. Screenshots**

Add screenshots of: - Home Page - Filled form before submission - Prediction result displayed

## **7. Results**

The model successfully predicts traffic volume based on the input parameters. Sample prediction:

**Input:** - Holiday: 1 - Temperature: 25°C - Rain: 0.1 mm - Snow: 0 mm - Weather: Clear - Year: 2022 - Month: 6 - Day: 15

**Output:** > Estimated Traffic Volume is: 4123 vehicles

## **8. Challenges Faced**

* Handling NaN values during scaling
* Mismatch between scaler and model input features
* Ensuring UI input matches the backend features

## **9. Conclusion**

This project provides a practical example of how machine learning can be applied to real-world traffic management systems. With further enhancement and integration with live data sources, this system can be made dynamic and scalable.

## **10. Future Work**

* Integrate real-time traffic API for live predictions
* Add more features like day of the week, time windows
* Improve accuracy with ensemble models or deep learning
* Host the model using cloud platforms like Heroku or AWS

## **11. References**

* [scikit-learn documentation](https://scikit-learn.org/)
* [Flask documentation](https://flask.palletsprojects.com/)
* [SmartBridge](https://smartbridge.in/)

**Thank You!**  
*Project guided and supported by SmartBridge.*

**Ideation Phase**

**Brainstorm & Idea Prioritization Template**

|  |  |
| --- | --- |
| **Date** | June 27,2025 |
| **Team ID** | LTVIP2025TMID39978 |
| **Project Name** | TrafficTelligence : Advanced Traffic Volume Estimation with Machine Learning |
| **Maximum Marks** | 4 Marks |

Brainstorming ideas is a creative process where a group generates a list of

potential solutions, suggestions, or concepts for a specific problem or project.

Voting in brainstorming involves participants selecting and prioritizing their

favourite or most promising ideas from the list to determine which ones should

be pursued further.

**Brainstorming for “TrafficTelligence : Advanced Traffic Volume Estimation with Machine Learning”:**

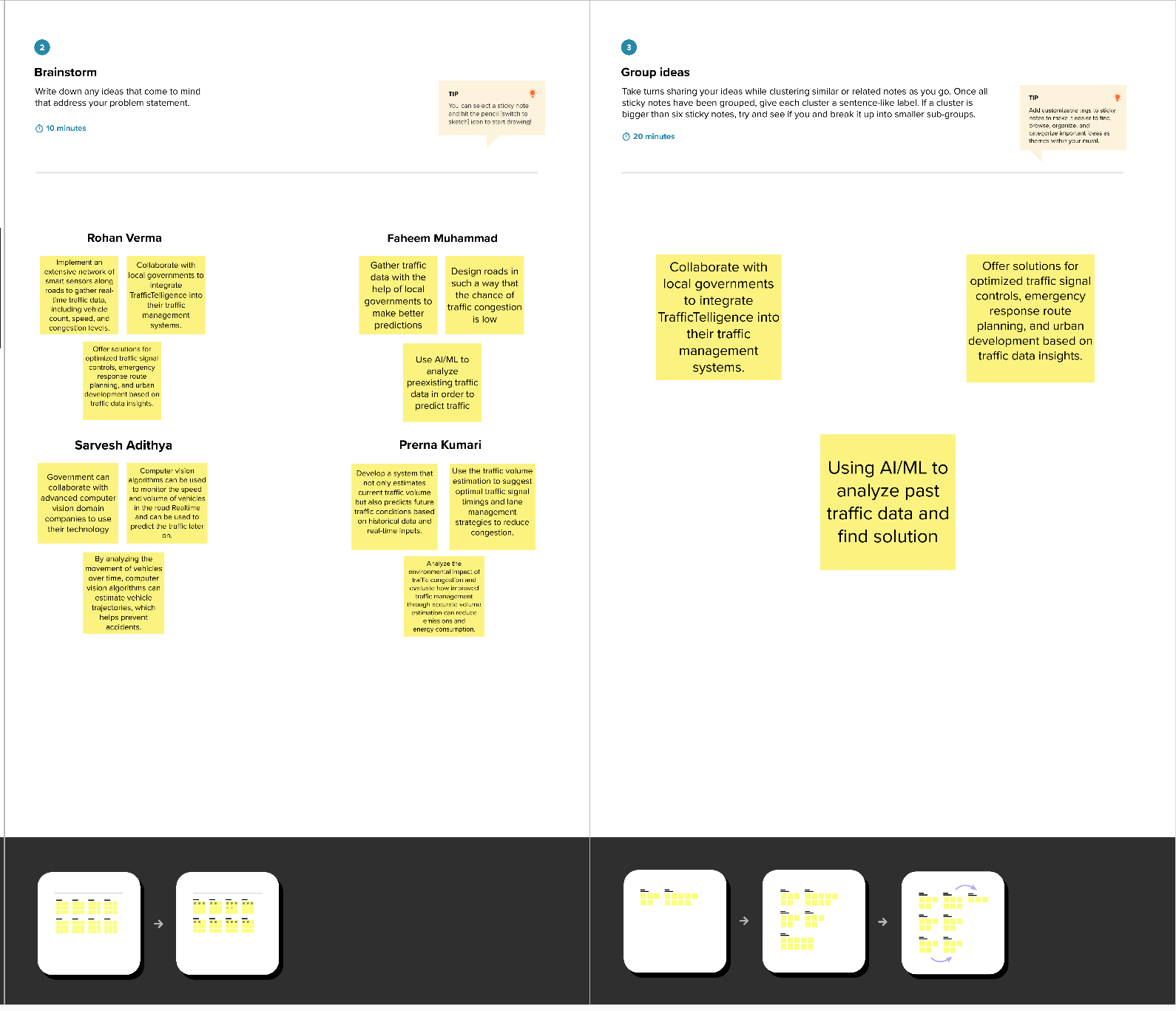
The objective of this brainstorming session is to generate creative and practical ideas to address the issue of Traffic Volume estimation effectively. We aim to help people able to plan their days better as they will have a better idea on how the traffic is going to be. It will also help traffic authorities be able to regulate traffic better.

The brainstorming session will include a diverse group of stakeholders, including public people, Traffic authorities, educators, community leaders, and technology enthusiasts. This diversity will ensure a wide range of perspectives and ideas.

**Step-1 : Team Gathering, Collaboration and Select the Problem Statement**



**Step-2: Brainstorm, Idea Listing and grouping**



**Step-3: Idea Prioritization**

Idea prioritization is the process of ranking or assessing ideas based on specific criteria such as feasibility, impact, cost, or strategic importance to determine which ideas should be implemented or pursued first.



**Here certainly we chose “Using AI/ML to analyze past traffic data and find solution” is:**

Among all of other ideas this was most important to us because, if the model is not accurate enough then the prediction may not be highly accurate. So, this was our most prioritized one.

Then comes our second most important idea such as **“Collaboration with local government to integrate TrafficTelligence into their traffic management systems”**. This was taken as our second because, if we want to give ourself a social responsibility that will be helpful, not only to use but also for others. If we work with other government or organization this might be helpful for a smooth traffic without any problems for Traffic authorities and also for people.

Then comes out our next idea **“Offer solutions for optimized traffic signal controls, emergency response route planning, and urban development based on traffic data insights.”** After fulfilling our main goal, we will scale our ML model not only to predict our main problem but also for extra features such as above-mentioned things. This will give our project more value in all ways.

**Ideation Phase**

**Empathize & Discover**

|  |  |
| --- | --- |
| **Date** | June 27, 2025 |
| **Team ID** | **LTVIP2025TMID39978** |
| **Project Name** | Project – TrafficTelligence : Advanced Traffic Volume Estimation with Machine Learning |
| **Maximum Marks** | 4 Marks |

**Empathy Map Canvas:**

An empathy map is a simple, easy-to-digest visual that captures knowledge about a user’s

behaviours and attitudes.

It is a useful tool to helps teams better understand their users.

Creating an effective solution requires understanding the true problem and the person who

is experiencing it. The exercise of creating the map helps participants consider things from

the user’s perspective along with his or her goals and challenges.

**Reference:** <https://www.mural.co/templates/empathy-map-canvas>

**TrafficTelligence : Advanced Traffic Volume Estimation with Machine Learning:**

Traffic problem is one of the major problem now a days, In the increase in no of vehicles and non –usage of public transport leading to traffic related issues, Making a eye on count of traffic at each level enables the government to take the further decisions such as building new roads, increasing infrastructure ,developing mutli-channel connectivity .

To address such problems to tracking the vehicle count in each and every place AI-ML has given a solution to such kind of traffic related issues, which are able to measure the volume of traffic, identify the violations of traffic rules etc.ML models could give early alerts of severe traffic to help prevent issues related to traffic problems.

Hence, there is needs to develop ML algorithms capable in predicting Traffic volume with acceptable level of precision and in reducing the error in the dataset of the projected Traffic volume from model with the expected observable Traffic volume.

**Empathy Canvas Map for the Project:**

**Project Development Phase**

**Model Performance Test**

|  |  |
| --- | --- |
| Date | 27 June 2025 |
| Team ID | LTVIP2025TMID39978 |
| Project Name | TrafficTelligence : Advanced Traffic Volume Estimation with Machine Learning |
| Maximum Marks | 10 Marks |

**Model Performance Testing:**

Project team shall fill the following information in model performance testing template.

|  |  |  |  |
| --- | --- | --- | --- |
| **S.No.** | **Parameter** | **Values** | **Screenshot** |
|  | Metrics | **Regression Model:** RMSE - 3.014763036569651 |  |

# Project Planning Phase

**Project Planning Template (Product Backlog, Sprint Planning, Stories, Story points)**

|  |  |
| --- | --- |
| Date | 27 June 2025 |
| Team ID | LTVIP2025TMID39978 |
| Project Name | TrafficTelligence : Advanced Traffic Volume Estimation with Machine Learning |
| Maximum Marks | 20 Marks |

Product Backlog, Sprint Schedule, and Estimation (4 Marks)

Use the below template to create product backlog and sprint schedule

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Sprint** | **Functional Requirement (Epic)** | **User Story Number** | **User Story / Task** | **Story Points** | **Priority** | **Team Members** |
| Sprint-1 | Project setup &  Infrastructure | USN-1 | Set up the development environment with the required tools and  frameworks to start the project | 1 | High | kavya |
| Sprint-2 | Data collection | USN-2 | Gather a diverse dataset of Date, time, holidays and climatic conditions. | 2 | High | Nandini kumari |
| Sprint-2 | data preprocessing | USN-3 | Preprocess the collected dataset by removing outliers and null values etc. Explore and evaluate different deep learning architectures (e.g., Regressions) to select the most suitable model for the project. | 3 | High | DineshVenkat |
| Sprint-3 | model development | USN-4 | train the selected machine learning model using the preprocessed  dataset and monitor its performance on the validation set. | 4 | High | Prudhvi raj |
| Sprint-3 | Training | USN-5 | The data set will be trained with suitable algorithms to improve the robustness and accuracy. | 6 | medium | Kavya |
| Sprint-4 | model deployment & Integration | USN-6 | deploy the trained machine learning model as a web service to make it accessible for users. Integrate the model's API into a user-friendly web interface for users to input variables such as date, time, holidays etc and receive predicted volume results. | 1 | medium | Nandini kumari |
| Sprint-5 | Testing & quality assurance | USN-7 | conduct thorough testing of the model and web interface to identify and report any issues or bugs. fine-tune the model hyperparameters and optimize its performance based on user feedback and testing results. | 1 | medium | DineshVenkat |

Project Tracker, Velocity & Burndown Chart: (4 Marks)

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Sprint** | **Total Story Points** | **Duration** | **Sprint Start Date** | **Sprint End Date (Planned)** | **Story Points Completed (as on Planned End Date)** | **Sprint Release Date (Actual)** |
| Sprint-1 | 1 | 3 Days | 8 June 2025 | 11 June 2025 | 1 | 11 June 2025 |
| Sprint-2 | 5 | 2 Days | 11 June 2025 | 13 June 2025 | 5 | 13 June 2025 |
| Sprint-3 | 10 | 3 Days | 13 June 2025 | 16 June 2025 | 10 | 16 June 2025 |
| Sprint-4 | 1 | 3 Days | 16 June 2025 | 19 June 2025 | 1 | 20 June 2025 |
| Sprint-5 | 1 | 2 Days | 19 June 2025 | 21 June 2025 | 1 | 22 June 2025 |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |

Velocity:

Imagine we have a 29-days sprint duration, and the velocity of the team is 20 (points per sprint). Let’s calculate the team’s average velocity (AV) per iteration unit (story points per day)



# AV= 19/3.8 = 5

Burndown Chart:

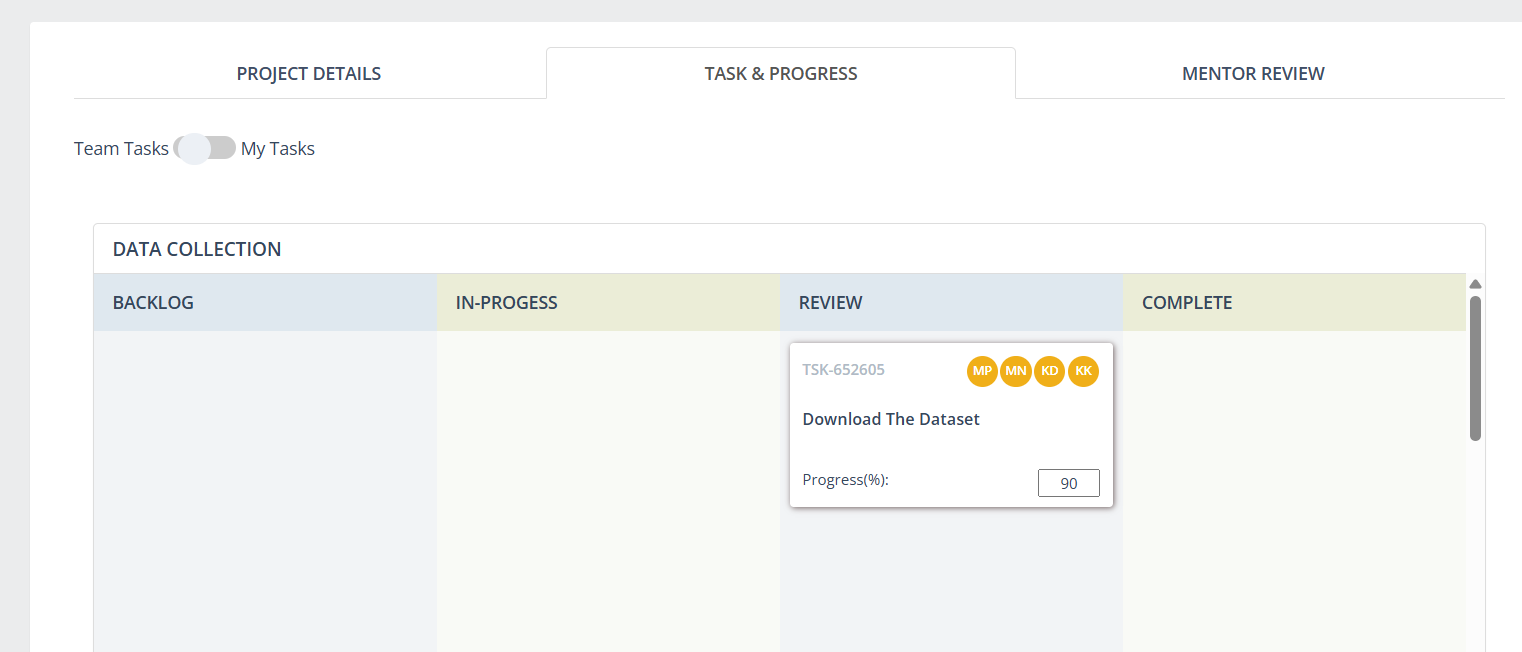
A burndown chart is a graphical representation of work left to do versus time. It is often used in agile [software development](about:blank) methodologies such as [Scrum](about:blank). However, burn down charts can be applied to any project containing measurable progress over time.

[https://www.visual-paradigm.com/scrum/scrum-burndown-chart/](about:blank) [https://www.atlassian.com/agile/tutorials/burndown-charts](about:blank)

Reference:

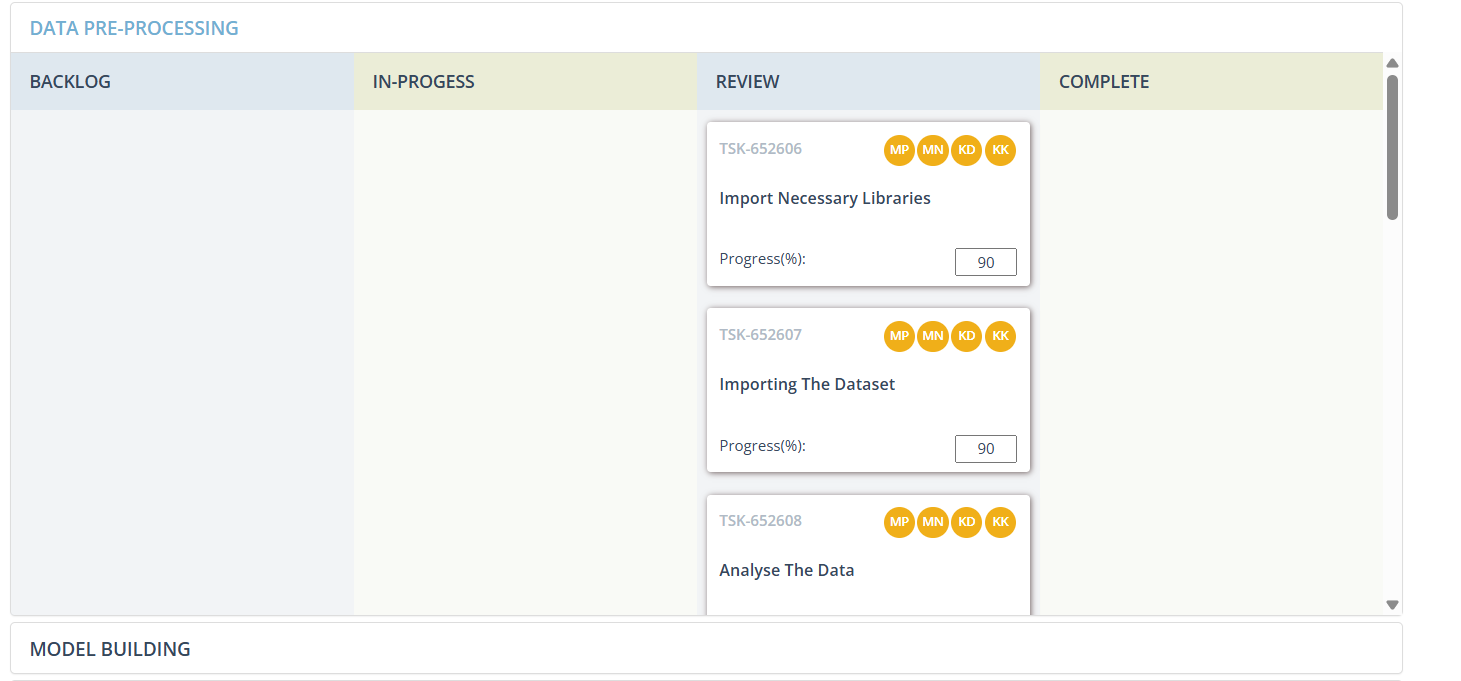
[https://www.atlassian.com/agile/project-management](about:blank) [https://www.atlassian.com/agile/tutorials/how-to-do-scrum-with-jira-software](about:blank) [https://www.atlassian.com/agile/tutorials/epics](about:blank) [https://www.atlassian.com/agile/tutorials/sprints](about:blank) [https://www.atlassian.com/agile/project-management/estimation](about:blank) [https://www.atlassian.com/agile/tutorials/burndown-charts](about:blank)

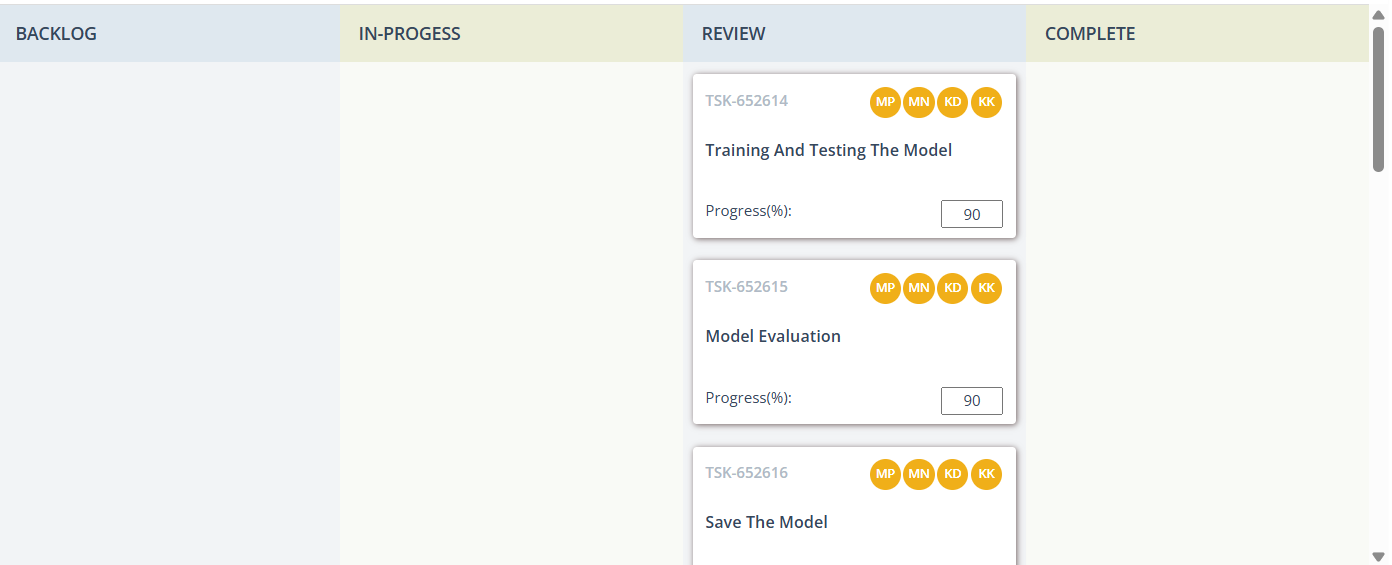
Burndown Chart: For the 5th Sprint

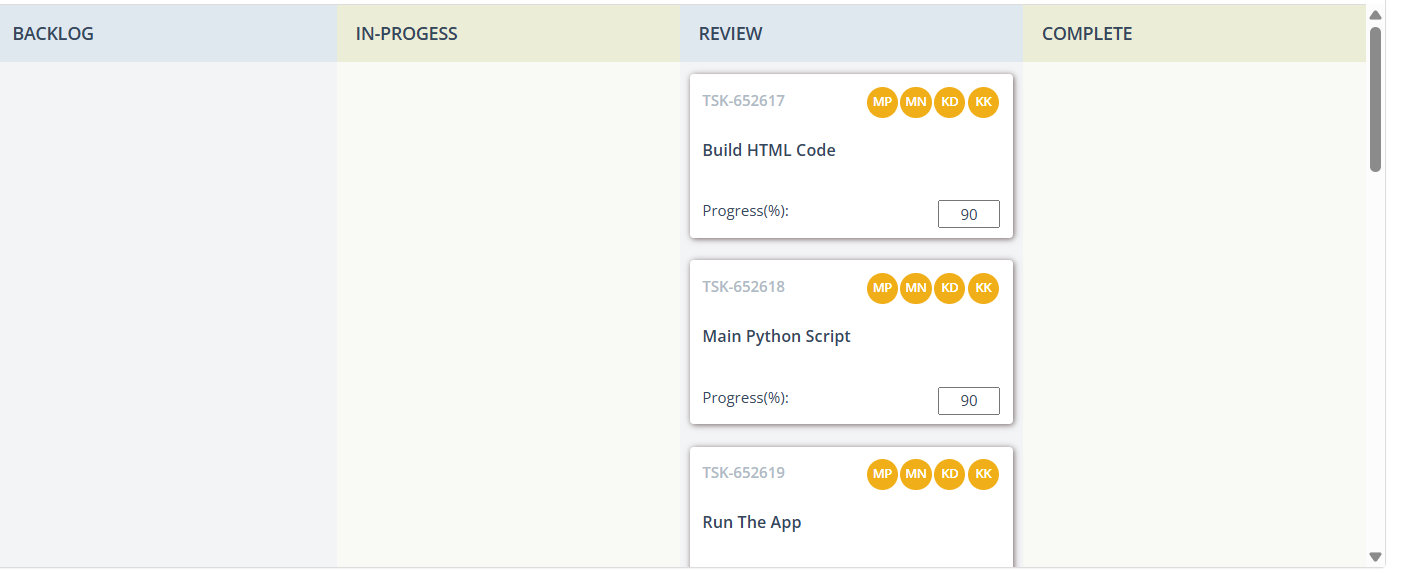


Board section.

We have completed till sprint 2.







Project Design Phase-II Technology Stack (Architecture & Stack)

|  |  |
| --- | --- |
| Date | 27 June, 2025 |
| Team ID | LTVIP2025TMID39978 |
| Project Name | TrafficTelligence : Advanced Traffic Volume Estimation with Machine Learning |
| Maximum Marks | 4 Marks |

**Technical Architecture:**

The Deliverable shall include the architectural diagram as below and the information as per the table1 & table 2



**Table-1 : Components & Technologies:**

|  |  |  |  |
| --- | --- | --- | --- |
| **S.No** | **Component** | **Description** | **Technology** |
| 1. | User Interface | Critical element designed for both Traffic Managers and everyday users, ensuring an intuitive and informative experience. | HTML, CSS, JavaScript |
| 2. | Application Logic-1 | Involves a robust backend system responsible for processing, analyzing, and managing traffic data. | Python |
| 3. | Database | Involves the storage and management of diverse traffic data for analysis. | File Manager, csv |
| 4. | File Storage/ Data | Involves managing diverse types of data, including raw traffic data, machine learning models, and configuration files. | Local System, Google Drive |
| 5. | Frame Work | It is a crucial part of our program as it is responsible for connecting the frontend with the backend. | Python Flask |
| 6. | Machine Learning Model | The machine learning model is responsible for predicting future outcomes based on available data | Machine learning model created using regression algorithms |
| 7. | Infrastructure (Server / Cloud) | Involves a combination of servers and cloud services to support the computational and storage needs of the application. | Local |

**Table-2: Application Characteristics:**

|  |  |  |  |
| --- | --- | --- | --- |
| **S.No** | **Characteristics** | **Description** | **Technology** |
| 1. | Open-Source Frameworks | Open-source frameworks can accelerate development and ensure the reliability of TrafficTelligence, contributing to a more efficient and maintainable solution. | Python’s Flask |

|  |  |  |  |
| --- | --- | --- | --- |
| **S.No** | **Characteristics** | **Description** | **Technology** |
| **2.** | Scalability | Using cameras to collect data and to make models for specific locations. | Computer vision, dynamic databases. |
| 3. | Performance | Regular performance testing, monitoring, and optimization are integral components of the development and maintenance processes, ensuring that TrafficTelligence consistently delivers timely and efficient traffic volume estimations. | R squared, Root mean squared error, Root Mean Square deviation |
| 4. | Availability | Website can be made available all time in a webserver. This makes the website running without any issues | High speed Linux based webservers. |