PROJECT REPORT

1. INTRODUCTION

1.1 Project Overview

The project "**TrafficTelligence – Advanced Traffic Volume Estimation with Machine Learning**" centers on accurately predicting traffic volume using machine learning algorithms. It leverages historical traffic data, weather patterns, public holidays, and timestamp features to build a reliable and intelligent traffic forecasting system. By integrating advanced regression models such as Random Forest and XGBoost within a Flask-powered web framework, the system provides real-time and scenario-based predictions that support dynamic traffic management, urban planning, and commuter navigation. The project includes comprehensive data preprocessing, visualization, model training, and performance evaluation, showcasing the impactful role of machine learning in solving real-world transportation challenges.

1.2 Purpose

- To develop a robust and intelligent traffic volume forecasting system using data-driven ML techniques.
- To automate traffic analysis and improve decision-making for urban planners and traffic authorities.
- To utilize ensemble learning algorithms like Random Forest and XGBoost for enhanced prediction accuracy.
- To reduce human error and manual analysis in traditional traffic estimation methods.
- To design a scalable web-based application that delivers real-time predictions to commuters and decision-makers.
- To demonstrate how supervised regression models can be adapted to smart city solutions.
- To promote efficient, responsive infrastructure development through data-informed planning and forecasting.

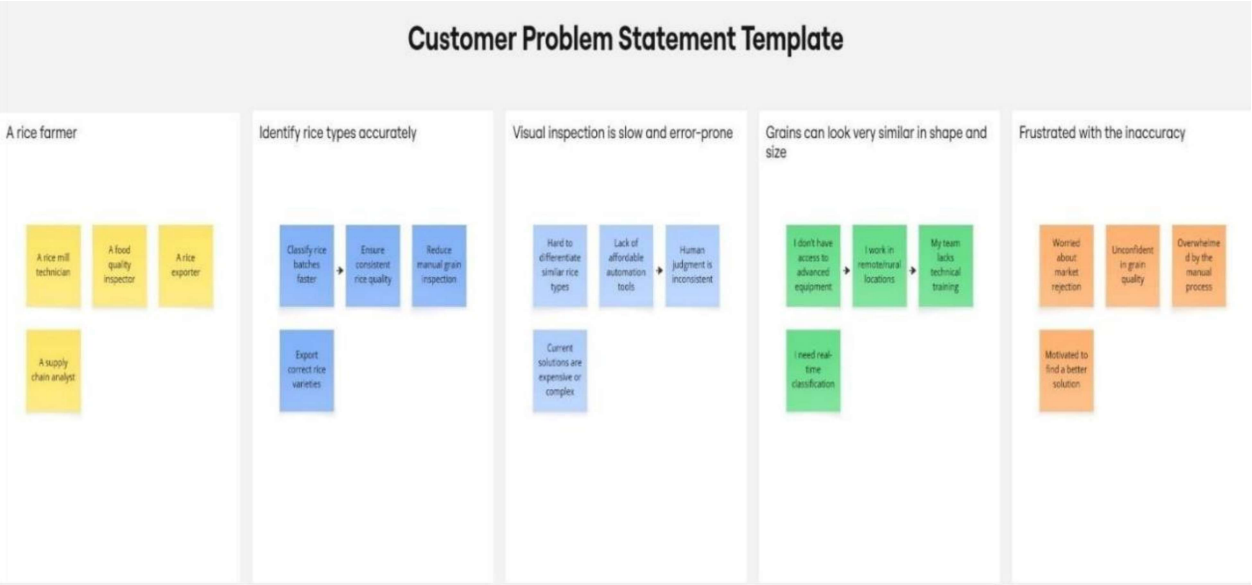
2. IDEATION PHASE

2.1 Problem Statement

Customer Problem Statement:

Create a problem statement to understand your customer's point of view. The Customer Problem Statement template helps you focus on what matters to create experiences people will love. A well-articulated customer problem statement allows you and your team to find the ideal solution for the challenges your customers face. Throughout the process, you'll

also be able to empathize with your customers, which helps you better understand how they perceive your product or service.



Reference: https://miro.com/templates/customer-problem-statement/

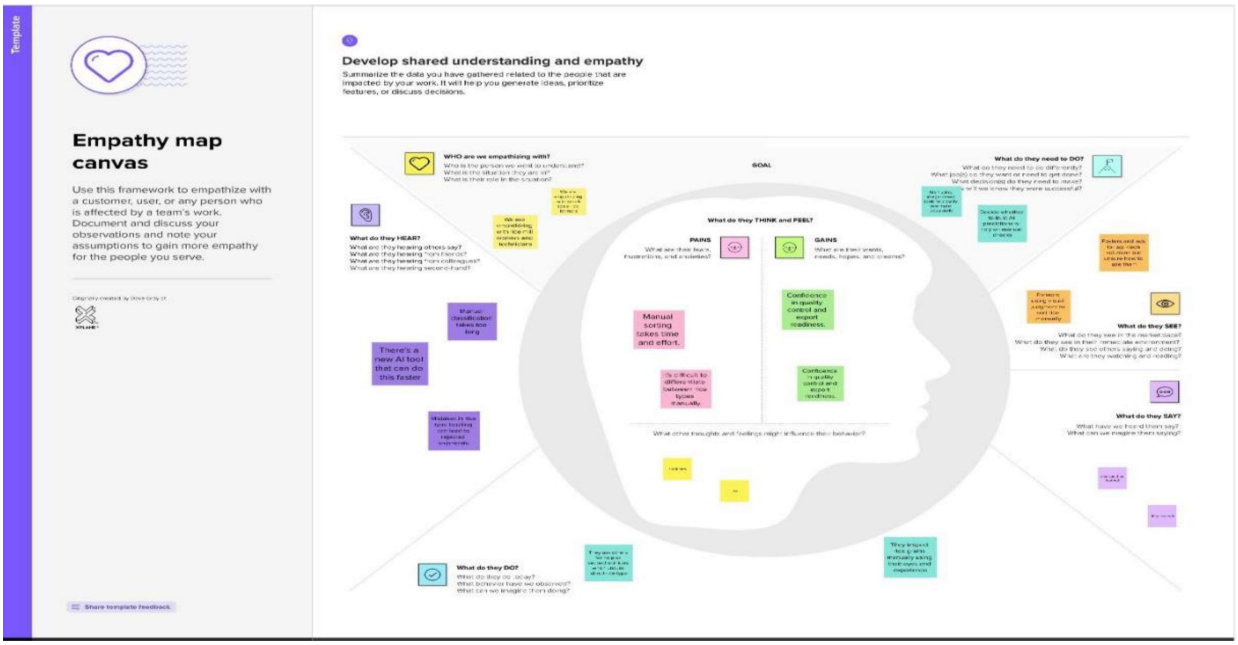
Problem Statement (PS)	I am (Customer)	I'm trying to	But	Because	Which makes me feel
PS-1	An urban planner or traffic system designer	Predict and manage traffic volume more accurately to improve road efficiency and commuter experience	I rely on manual analysis or outdated models that don't factor in dynamic variables like weather, time, or historical patterns	Traditional methods lack adaptability and cannot handle the complexity of real-world traffic behavior	Helpless, reactive, and concerned about poor traffic planning decisions that lead to congestion and dissatisfaction
PS-2	I daily commuter relying on	Plan my travel routes to avoid	Most navigation	They lack intelligent forecasting	Frustrated, late, and anxious about reaching

	city navigation apps	heavy traffic and delays	systems only react to current traffic rather than predict future congestion	based on patterns like time, weather, and historical volume trends	destinations on time

2.2 Empathy Map Canvas

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



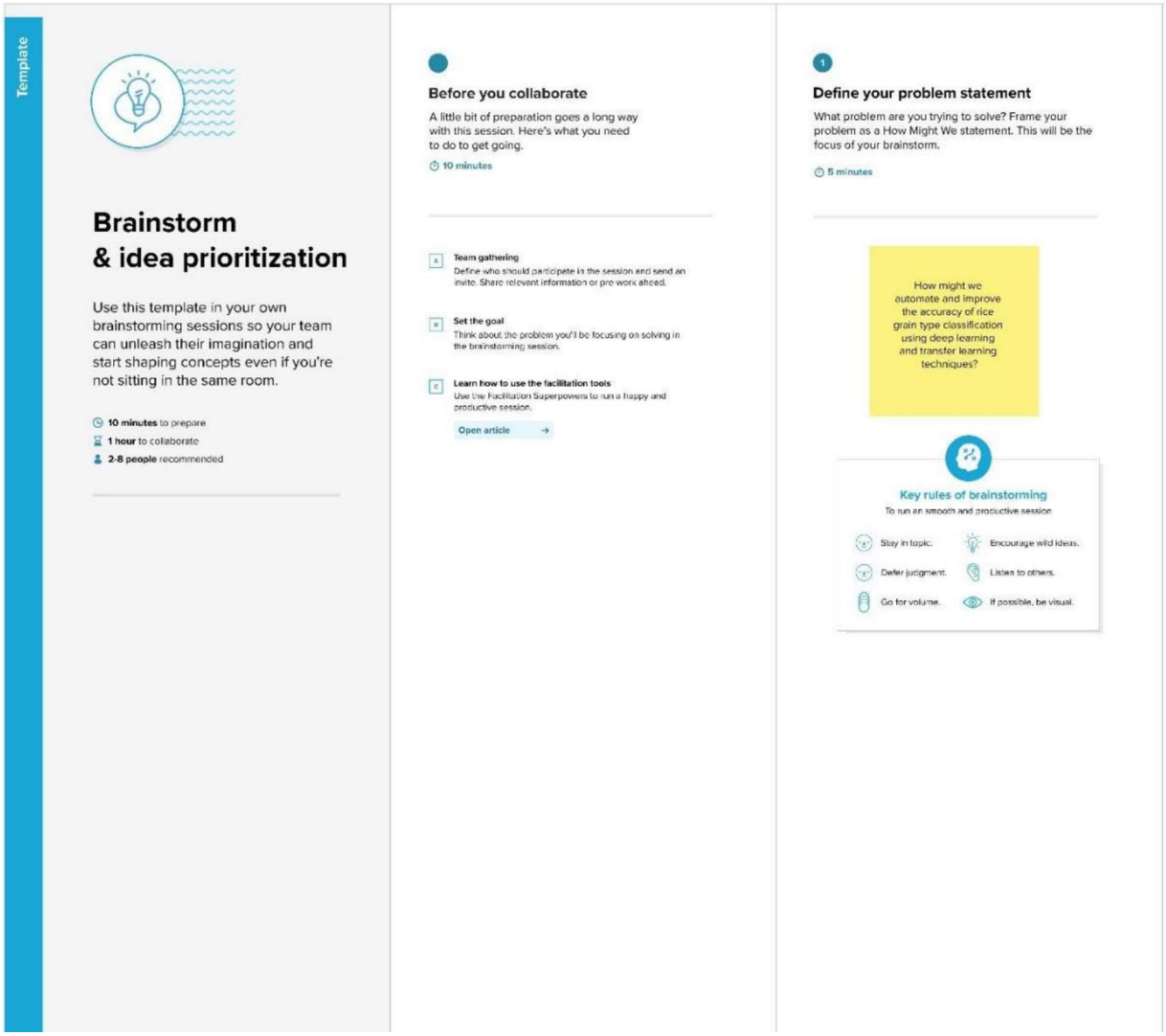
2.3 Brainstorming

Brainstorm & Idea Prioritization:

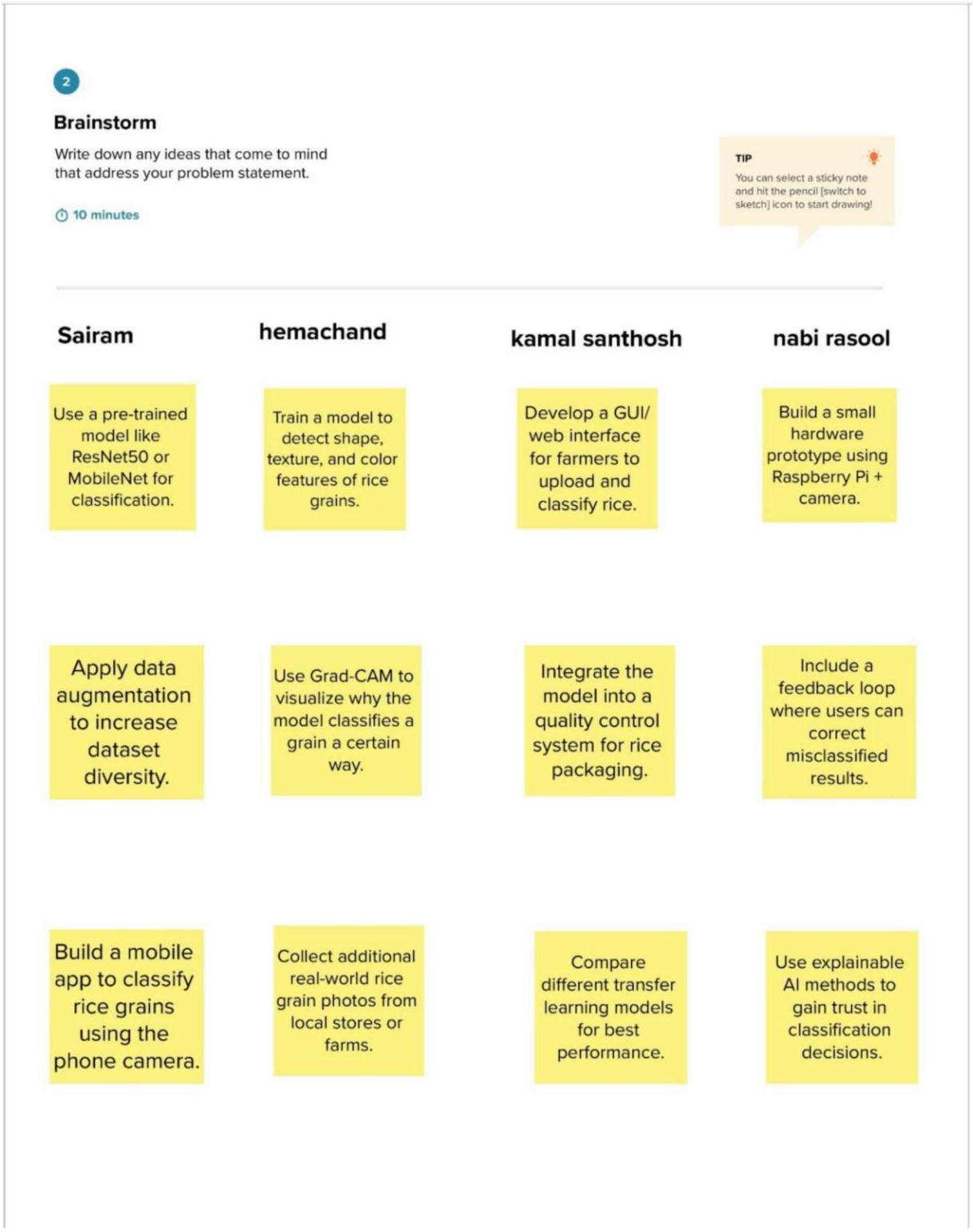
Brainstorming provides a free and open environment that encourages everyone within a team to participate in the creative thinking process that leads to problem solving. Prioritizing volume over value, out-of-the-box ideas are welcome and built upon, and all participants are encouraged to collaborate, helping each other develop a rich number of creative solutions.

Use this Reference: https://www.mural.co/templates/brainstorm-and-idea-prioritization

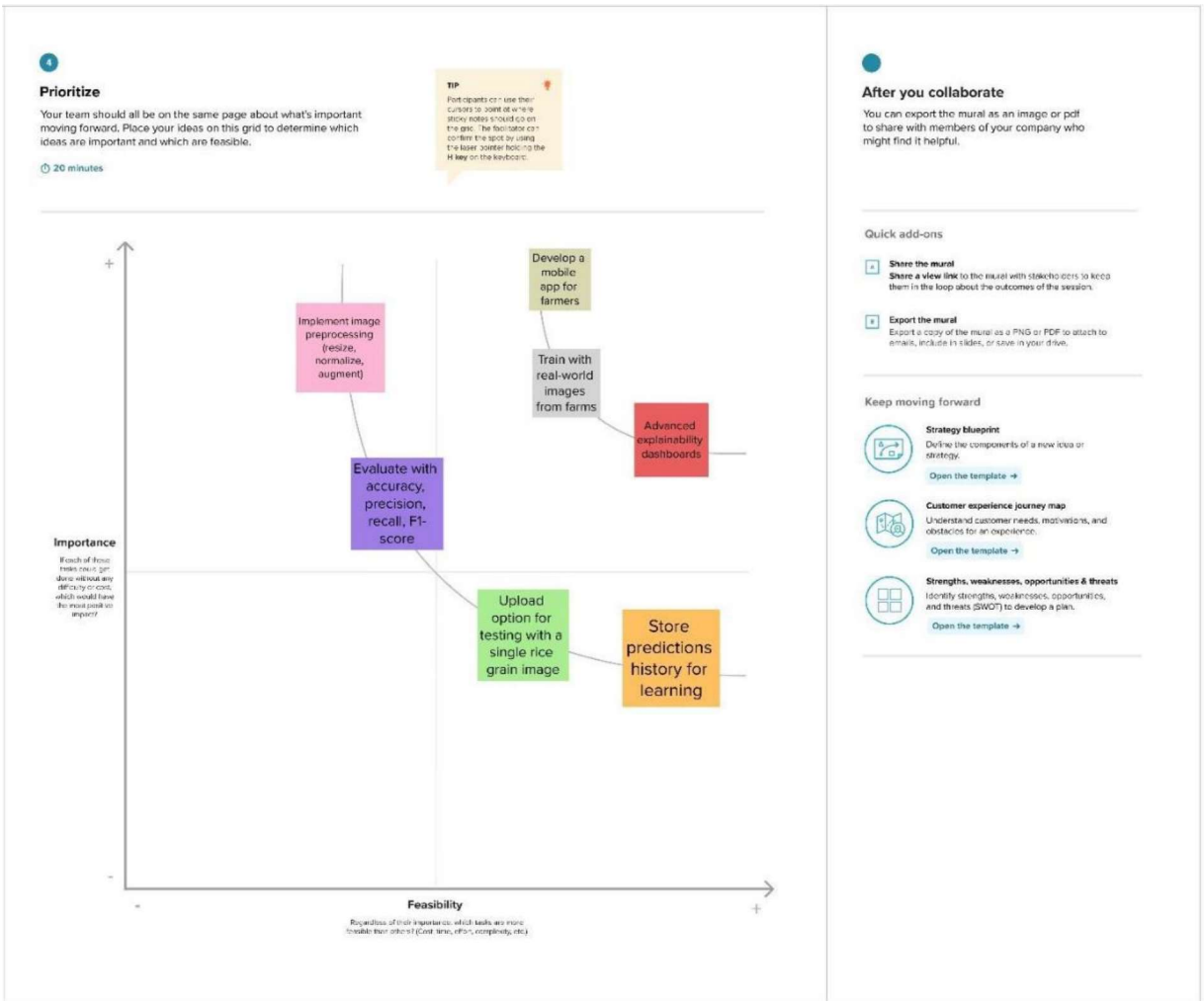
Step-1: Team Gathering, Collaboration and Select the Problem Statement



Step-2: Brainstorm, Idea Listing and Grouping



Step-3: Idea Prioritization



3. REQUIREMENT ANALYSIS

Functional Requirements:

Following are the functional requirements of the proposed solution.

FR No.	Functional Requirement (Epic)	Sub Requirement (Story / Sub-Task)
FR-1	User Registration	Registration through Form Registration through Gmail Registration through LinkedIn
FR-2	User Confirmation	Confirmation via Email Confirmation via OTP

FR-3	document upload	document contains the previous data of the traffic volume by day wise
FR-4	Traffic Prediction Interface	User inputs temporal and environmental variables - Predict traffic volume using trained ML model - Route user to appropriate page (`chance.html` or `noChance.html`)

Non-functional Requirements:

Following are the non-functional requirements of the proposed solution.

NFR No.	Non-Functional Requirement	Description
NFR-1	**Usability**	Interface is intuitive, accessible to users with varying technical backgrounds
NFR-2	**Security**	OAuth authentication; Secure storage of model and data; Protection against injection and unauthorized access
NFR-3	**Reliability**	System provides consistent predictions with validated model accuracy metrics
NFR-4	**Performance**	Predictive system gives results in <2 seconds for common input
NFR-5	**Availability**	Application uptime targeted at >99% during daytime hours; supports multiple concurrent users
NFR-6	**Scalability**	Architecture allows for cloud deployment and easy retraining with new datasets or models

4.PROJECT DESIGN

4.1Problem Solution Fit

Problem – Solution Fit Overview:

Problem Statement:

Urban traffic congestion is a growing challenge that disrupts daily commuting, increases pollution, and strains infrastructure. Traditional traffic monitoring systems are reactive and often manual, lacking the adaptability to account for dynamic factors like weather, holidays, and time-based patterns. There is a need for a predictive solution that leverages data and automation to estimate

traffic volume accurately and support smarter decisions for traffic control, urban planning, and navigation tools.

Solution:

TrafficTelligence is a machine learning–based system designed to predict urban traffic volumes using contextual data like weather, holidays, and time features. It tackles inefficiencies of traditional traffic management by providing accurate, real-time forecasts. Built with models like Random Forest and deployed through a Flask web app, it empowers commuters, planners, and city officials with actionable insights. The solution improves navigation, supports infrastructure planning, and minimizes congestion. While highly scalable and precise, it depends on quality data and regular model updates to stay effective.

4.2 Proposed Solution

Project team shall fill the following information in the proposed solution template.

S.No.	Parameter	Description
1.	Problem Statement (Problem to be solved)	Accurate traffic volume prediction is challenging due to changing weather, time patterns, and outdated methods. This project solves that by using machine learning to forecast traffic based on real-world data, helping urban planners and commuters make smarter decisions.
2.	Idea / Solution description	Our project, TrafficTelligence, is a web-based machine learning system that predicts traffic volume based on factors like time, weather, and past data. It helps users (like commuters and city planners) make better decisions to avoid traffic and plan smarter infrastructure.
3.	Novelty / Uniqueness	Unlike regular traffic maps that only show current congestion, our solution predicts future traffic using regression models. It also uses different types of input data together, which most systems don't do. This makes the prediction smarter and more reliable.
4.	Social Impact / Customer Satisfaction	The tool reduces time wasted in traffic, helps cities plan better roads, and supports eco-friendly travel by cutting down fuel usage. It's useful for drivers, planners, and tech enthusiasts who want smarter cities and smoother travel. Users will feel more confident and less stressed while commuting.

5.	Business Model (Revenue Model)	The solution can be monetized by offering premium features like: 1, Live traffic suggestions for navigation apps 2, Subscription model for urban planners and developers 3, API access for startups working on mobility solutions
6.	Scalability of the Solution	This project is scalable because it can be deployed in different cities by updating the local traffic data. It can also be extended to predict road accidents, traffic jams, or even parking availability—making it future-ready for smart cities.

4.3 Solution Architecture

Solution Architecture:

Solution architecture is a complex process – with many sub-processes – that bridges the gap between business problems and technology solutions. Its goals are to:

- Find the best tech solution to solve existing business problems.
- Describe the structure, characteristics, behavior, and other aspects of the software to project stakeholders.
- Define features, development phases, and solution requirements.

Example - Solution Architecture Diagram:

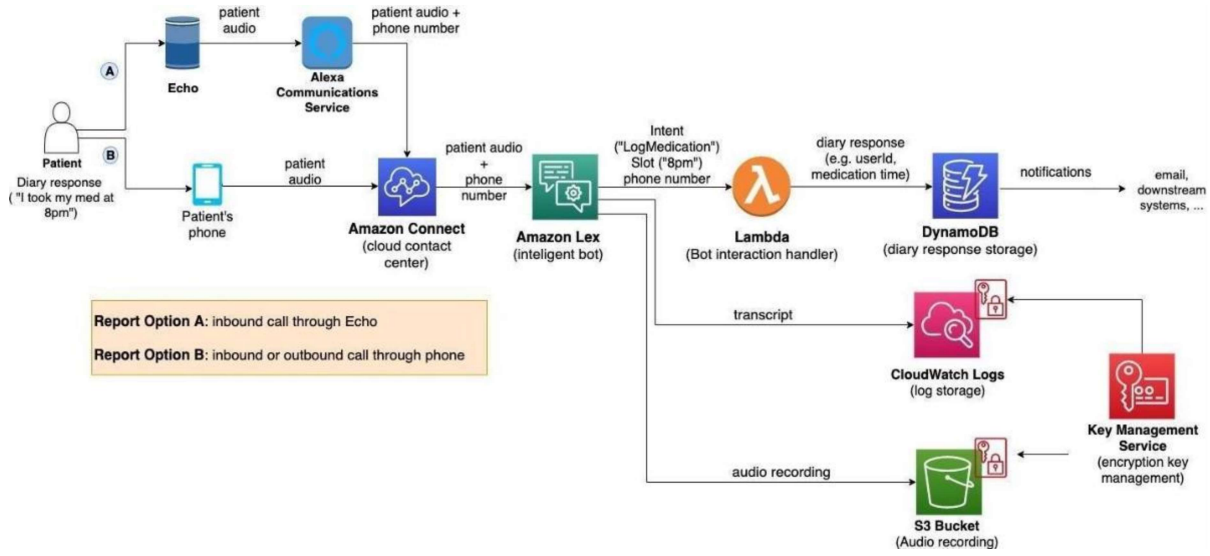


Figure 1: Architecture and data flow of the voice patient diary sample application

Reference: [https://aws.amazon.com/blogs/industries/voice-applications-in-clinical-research- powered-by-ai-on-aws-part-1-architecture-and-design-considerations/](https://aws.amazon.com/blogs/industries/voice-applications-in-clinical-research-powered-by-ai-on-aws-part-1-architecture-and-design-considerations/)

5.PROJECT PLANNING & SCHEDULING

5.1 Project Planning

Sprint	Functional Requirement (Epic)	User Story Number	User Story / Task	Story Points	Priority	Team Members
Sprint - 1	Image data collection	USN-1	Collect, organize, and label traffic-related images and dataset features from open data sources	3	High	1, N Hema Gopika Devi 2, Mohammad Kouser Ali 3, Mohammad Dastagiri Ahmed Basha 4, Mohammad Fairoz
Sprint - 1	Data preprocessing	USN-2	Clean missing values, encode categories, perform feature scaling, and extract temporal features	3	High	1, N Hema Gopika Devi 2, Mohammad Kouser Ali

						3, Mohammad Dastagiri Ahmed Basha 4, Mohammad Fairoz
Sprint - 2	Model training	USN-3	Train supervised regression models (Random Forest, XGBoost) using preprocessed data	5	high	1, N Hema Gopika Devi 2, Mohammad Kouser Ali 3, Mohammad Dastagiri Ahmed Basha 4, Mohammad Fairoz
Sprint - 2	Model evaluation	USN-4	Evaluate models using metrics like R^2 score and RMSE; select best model based on performance	2	medium	1, N Hema Gopika Devi 2, Mohammad Kouser Ali 3, Mohammad Dastagiri Ahmed Basha 4, Mohammad Fairoz
Sprint - 3	Ui development	USN-5	Build Flask-based web interface for traffic input and prediction display (index.html + templates)	3	medium	1, N Hema Gopika Devi 2, Mohammad Kouser Ali 3, Mohammad Dastagiri Ahmed Basha 4, Mohammad Fairoz
Sprint -3	Deployment	USN-6	Integrate trained model with web app, test locally, and prepare for scalable deployment	4	low	1, N Hema Gopika Devi 2, Mohammad Kouser Ali 3, Mohammad Dastagiri Ahmed Basha 4, Mohammad Fairoz

Provide specifications according to which the solution is defined, managed, and delivered.

Project Tracker, Velocity & Burndown Chart:

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	20	6 Days	16 June 2025	21 June 2025	20	21 June 2025
Sprint-2	20	6 Days	17 June 2025	22 June 2025	20	22 June 2025
Sprint-3	20	6 Days	18 June 2025	23 June 2025	20	23 June 2025
Sprint-4	20	6 Days	19 June 2025	24 June 2025	20	24 June 2025

Velocity:

Imagine we have a 10-day 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)

$$AV = \frac{sprint\ duration}{velocity} = \frac{20}{10} = 2$$

Burndown Chart:

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

https://www.visual-paradigm.com/scrum/scrum-burndown-chart/
https://www.atlassian.com/agile/tutorials/burndown-charts

Reference:

https://www.atlassian.com/agile/project-management
https://www.atlassian.com/agile/tutorials/how-to-do-scrum-with-jira-software
https://www.atlassian.com/agile/tutorials/epics **https://www.atlassian.com/agile/tutorials/sprints**

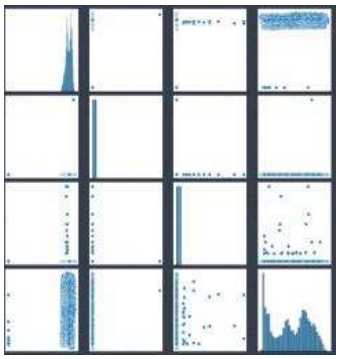
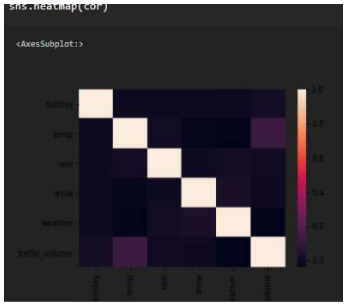
https://www.atlassian.com/agile/project-management/estimation
https://www.atlassian.com/agile/tutorials/burndown-charts

6.FUNCTIONAL AND PERFORMANCE TESTING

6.1 Performance Testing

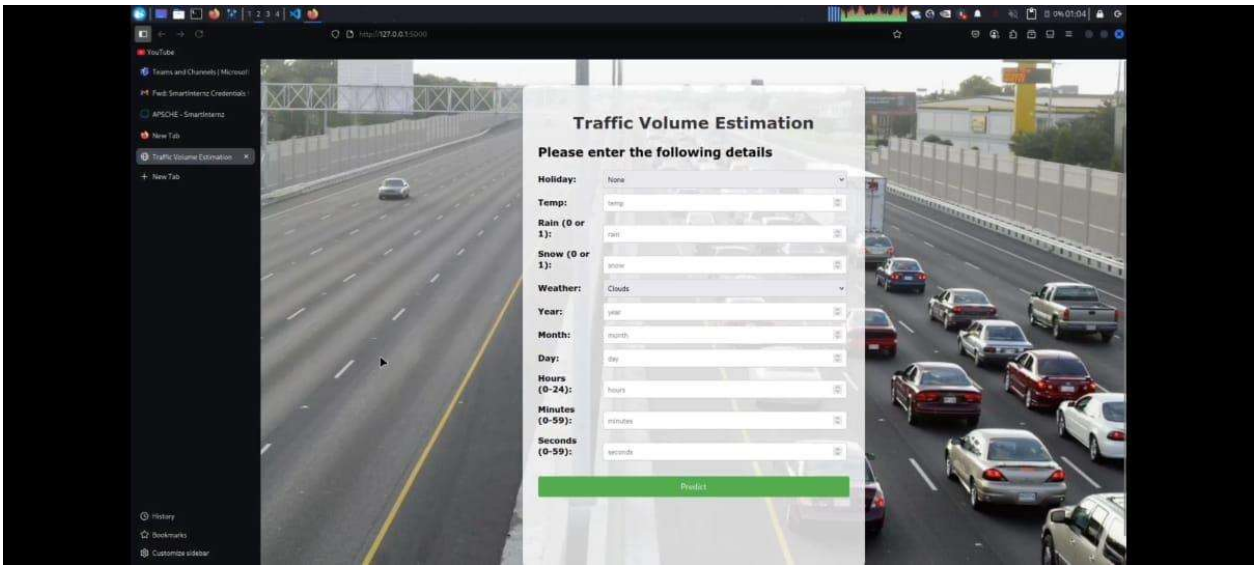
Model Performance Testing:

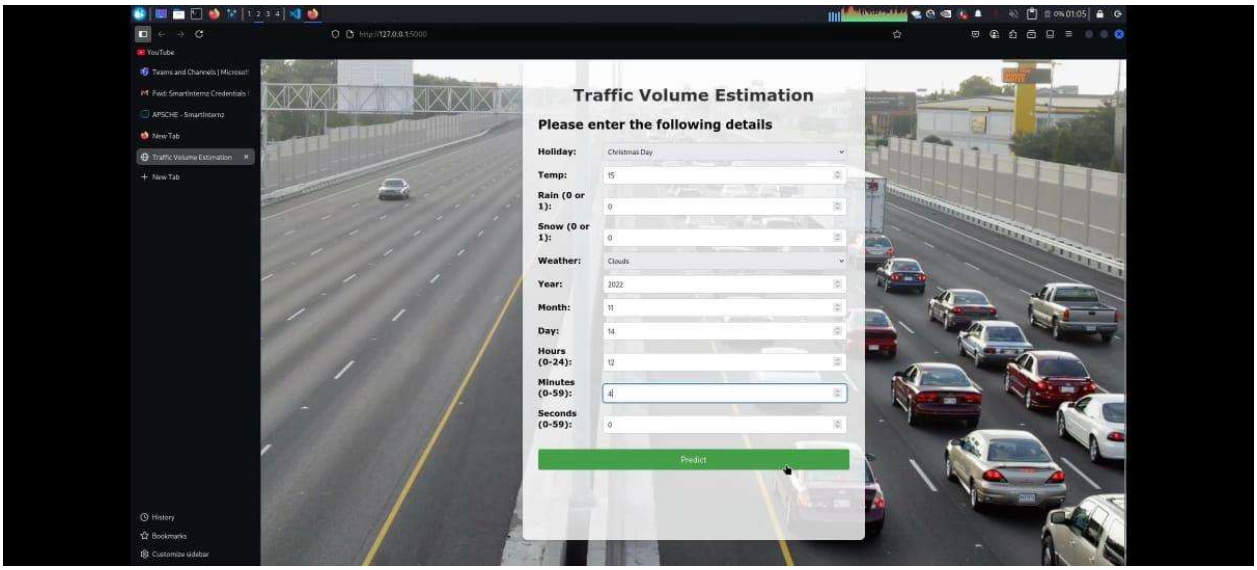
Project team shall fill the following information in model performance testing template.

S.No.	Parameter	Values	Screenshot
1.	Metrics	**Regression Model:** MAE -, MSE -, RMSE -, R2 score **ClassificationModel:** Confusion Matrix -, Accuray Score- & Classification Report -	
2.	Tune the Model	Hyperparameter Tuning - Validation Method -	

7.RESULTS

7.1 Output Screenshots:


Traffic Volume Estimation
Please enter the following details
Holiday:
None
Temp:
Rain (0 or 1):
Snow (0 or 1):
Weather:
Clouds
Year:
Month:
Day:
Hours (0-24):
Minutes (0-59):
Seconds (0-59):
Predict


Traffic Volume Estimation
Please enter the following details
Holiday:
Christmas Day
Temp:
15
Rain (0 or 1):
0
Snow (0 or 1):
0
Weather:
Clouds
Year:
2022
Month:
11
Day:
14
Hours (0-24):
12
Minutes (0-59):
4
Seconds (0-59):
0
Predict

8. ADVANTAGES & DISADVANTAGES

Here are the Advantages:

- **High Accuracy**: Ensemble models like Random Forest offer precise traffic volume predictions.
- **Scalable Framework**: Easily adaptable for new datasets, cities, or additional input features.
- **User-Friendly Interface**: The Flask web app ensures intuitive access and fast results.
- **Real-Time Forecasting**: Enables proactive traffic planning and dynamic routing.
- **Versatile Applications**: Supports traffic management, urban planning, and commuter guidance.

Here are some Disadvantages:

- **Data Dependency**: Prediction quality depends on the availability and reliability of historical data.
- **Limited Generalization**: Unseen events or anomalies may degrade model performance.
- **Manual Preprocessing**: Requires consistent data cleaning and transformation for optimal output.
- **Deployment Complexity**: Scaling the solution for cloud or mobile platforms adds infrastructure overhead.

9. Conclusion

TrafficTelligence demonstrates how machine learning can effectively address urban mobility challenges by forecasting traffic volumes using real-world contextual data. Through comprehensive data preprocessing, model evaluation, and web-based deployment, the system delivers accurate, actionable insights for both public authorities and individual commuters. Its modular architecture and predictive capabilities make it a powerful foundation for smarter city planning and real-time navigation support.

10.Future Scope

- **Real-Time Data Integration**: Connect to live traffic sensors or open traffic APIs to enable instant volume predictions and more responsive traffic control.
- **Multi-Modal Support**: Expand the model to include other transportation modes such as public transit, cycling, and pedestrian flow for smarter urban mobility.
- **Mobile App Development**: Create a companion mobile application that offers personalized, location-based traffic forecasts and alerts for commuters.
- **Cloud Deployment**: Scale the system using platforms like AWS or Azure to handle large datasets and multiple concurrent users across cities.
- **Adaptive Learning**: Introduce continuous model retraining to accommodate new data, changing trends, and evolving infrastructure patterns.
- **Advanced Visualization Dashboards**: Build interactive admin dashboards for traffic analysts to monitor predictions, usage statistics, and performance in real-time.

11. APPENDIX

Source Code:

TrafficTelligence-Advanced-Traffic-Volume-Estimation-with-Machine-Learning/notebook/traffic volume estimation1 -checkpoint.ipynb at main · Mohammad-Kouser-Ali/TrafficTelligence-Advanced-Traffic-Volume-Estimation-with-Machine-Learning · GitHub

GitHub & Project Demo Link:

https://github.com/Mohammad-Kouser-Ali/TrafficTelligence-Advanced-Traffic-Volume-Estimation-with-Machine-Learning

Project Demo Link:

https://drive.google.com/file/d/1JUyHl9JVWcszspuEZ-24j0bWIynmzAv3/view?usp=drivesdk