Project Proposal Report

Munasinghe M.M.A.D. – IT21269820

Supervisor: Mr. H. M. Samadhi Chathuranga Rathnayake

Co-Supervisor: Ms. Supipi Karunathilaka

Group No : R25 – 034

BSc (Hons) in Information Technology Specializing in Information Technology

Department of Information Technology

Sri Lanka Institute of Information Technology

Sri Lanka

January 2025

DECLARATION

We declare that this is my own work, and this proposal does not incorporate without acknowledgement any material previously submitted for a degree or diploma in any other university or Institute of higher learning and to the best of my knowledge and belief it does not contain any material previously published or written by another person except where the acknowledgement is made in the text.

Name	Student ID	Signature
Munasinghe M.M.A.D.	IT21269820	

Date

2025/01/27

……………………………………

Signature of the Supervisor
(Mr. H. M. Samadhi Chathuranga Rathnayake)

………………………………

ABSTRACT

Driver behavior analysis is essential for improving road safety, reducing accidents, and optimizing driver performance. Traditional methods rely heavily on expensive hardware or advanced connected vehicle technologies, which limit their accessibility and affordability. This project introduces a novel and cost-effective approach to driver behavior analysis using non-invasive, widely available mobile phone sensors (e.g., accelerometer, GPS) in conjunction with the CARLA simulator. By simulating various driving scenarios in CARLA, the project can generate diverse and realistic datasets that complement data collected from real-world sources.

The system leverages the CARLA simulator to simulate driving patterns, including aggressive, moderate, and cautious behaviors, allowing for controlled experiments and comprehensive data collection. These simulated datasets are combined with mobile-sourced driving data to train machine learning models, such as decision trees and support vector machines, for classifying driving behaviors. This approach ensures high-quality training data and robust model performance, even in scenarios where real-world data might be limited or inconsistent.

The research outcomes have significant implications for road safety initiatives, driver education, fleet management systems, and insurance telematics. By integrating the CARLA simulator, this project not only ensures scalability and low-cost implementation but also establishes a foundation for replicable and reliable driver behavior analysis. This innovation bridges the gap between advanced driver monitoring technologies and accessible, user-friendly solutions for improving road safety across diverse settings.

Contents

1. INTRODUCTION

Driver behavior plays a significant role in determining road safety, fuel efficiency, and vehicle performance. Reckless driving behaviors, such as harsh acceleration, sharp braking, and erratic lane changes, are leading contributors to traffic accidents and vehicle wear and tear. Traditional methods of analyzing driver behavior often depend on expensive telematics systems or specialized in-vehicle hardware, making them inaccessible for many users, especially in low-resource settings.

Recent advancements in mobile technologies offer a promising alternative for driver behavior analysis. Mobile devices equipped with accelerometers and GPS sensors can capture driving metrics, such as speed, braking patterns, and acceleration, at a fraction of the cost of traditional systems. Furthermore, driving simulators like CARLA provide a controlled environment for generating diverse datasets and simulating real-world driving scenarios.

While previous studies have utilized mobile sensor data or simulator-based approaches, there remains a gap in integrating these two methodologies to enhance data diversity and improve model robustness. This project aims to address this gap by combining real-world data collected from mobile phone sensors with simulated data generated in the CARLA simulator. Machine learning models, such as decision trees and support vector machines, will be trained on these datasets to classify driving behavior into categories like aggressive, moderate, and cautious.

By leveraging non-invasive technologies and simulation platforms, this project seeks to create a scalable, cost-effective, and widely accessible system for driver behavior analysis. The findings have potential applications in road safety initiatives, driver education programs, fleet management, and insurance telematics, making a significant contribution to safer and more efficient road usage.

1.1. BACKGROUND & LITERATURE SURVEY

Driver behavior is a critical factor influencing road safety, fuel consumption, vehicle maintenance, and overall traffic efficiency. Unsafe driving practices such as aggressive acceleration, harsh braking, and erratic lane changes are significant contributors to road accidents and increased vehicle wear and tear. Traditional driver behavior analysis methods typically rely on expensive in-vehicle devices, such as high-end telematics systems or cameras, which limit their applicability to high-resource settings.

Recent advancements in technology, particularly the proliferation of mobile devices equipped with sensors like accelerometers and GPS, have opened new possibilities for non-invasive, low-cost driver behavior analysis. Studies have demonstrated the potential of mobile phone sensors to collect data on driving metrics, including speed, acceleration, and braking patterns. Additionally, simulators like CARLA (an open-source driving simulator) provide a powerful platform for generating realistic driving data under controlled conditions.

Existing research has primarily focused on either real-world data collection or simulator-based approaches, but few studies have combined the two to maximize data quality and diversity. This project builds prior research by integrating mobile sensor data with CARLA simulator-generated datasets to create a robust and scalable driver behavior analysis system.

1.2. RESEARCH GAP

While significant strides have been made in fuel efficiency prediction using machine learning and deep learning techniques, critical limitations persist in existing research. Research 1 (2024) focused on predictive modeling based on vehicle parameters such as engine load and speed. However, it failed to leverage IoT data or incorporate personalized driving patterns—key elements that can elevate prediction accuracy and ensure real-world applicability. Research 2 (2009) employed advanced deep learning methods, such as RNNs and CNNs, to analyze real-time driving data. Despite its innovative approach to capturing temporal and spatial patterns, it overlooked the integration of IoT-driven insights and dynamic maintenance scheduling, both essential for a comprehensive and adaptive solution.Research 3 (2014) introduced a hybrid model combining SVM with genetic algorithms, achieving notable accuracy improvements. Yet, it lacked focus on behavioral factors, such as personalized driving patterns, and did not explore the potential of IoT-based data streams to enhance predictions. Across these studies, the absence of a unified system that integrates IoT data, dynamic scheduling, advanced predictive models, and real-time behavioral inputs remains a significant shortcoming.

The proposed system directly addresses these gaps by leveraging IoT data to capture real-time insights, creating dynamic maintenance schedules, incorporating personalized driving behaviors, and refining predictive accuracy with cutting-edge machine learning models. This holistic approach not only surpasses the limitations of previous studies but also establishes a robust framework for optimizing vehicle performance and fuel efficiency in a real-world context.

Features	Research 1 (2020) [1]	Research 2 (2013) [2]	Research 3 (2022) [3]	Proposed System
Utilizes Mobile Phone Sensors	✓	✓	✓	✓
Employs CARLA Simulator for Data Generation	✗	✗	✓	✓
Integrates Real and Simulated Data	✗	✗	✗	✓
Classifies Driving Behavior Using Machine Learning	✓	✓	✓	✓
Focuses on Non-Invasive, Low-Cost Implementation	✓	✓	✓	✓

2. RESEARCH PROBLEM

Driver behavior is a critical determinant of road safety, vehicle performance, and overall traffic efficiency. Unsafe driving practices, such as aggressive acceleration, harsh braking, and speeding, contribute significantly to road accidents and vehicle wear and tear. Traditional methods of driver behavior analysis often rely on expensive in-vehicle telematics systems or connected car technologies, limiting their accessibility to resource-rich environments.

At the same time, advancements in mobile phone sensors and open-source driving simulators like CARLA provide an opportunity to analyze driving behavior in a cost-effective and scalable manner. However, the integration of real-world mobile sensor data and simulated data for robust driver behavior classification remains underexplored.

The research problem centers on developing a system that leverages non-invasive technologies, such as mobile accelerometers and GPS sensors, in combination with data generated from the CARLA simulator to classify driving behavior. The challenge lies in ensuring the accuracy, scalability, and practicality of such a system while addressing limitations like data diversity, model robustness, and user accessibility. This project aims to bridge this gap by creating a comprehensive and low-cost solution for driver behavior analysis, with the potential to impact road safety initiatives, fleet management systems, and driver training programs globally.

3. OBJECTIVES

3.1 MAIN OBJECTIVES

The primary objective of this project is to develop a scalable, cost-effective, and non-invasive system for classifying driver behavior using machine learning techniques. By focusing on affordable and widely accessible technologies, the project aims to make significant contributions to road safety, vehicle optimization, and driver performance improvements across diverse environments, including low-resource settings.

To create a non-invasive, scalable, and cost-effective system capable of classifying driver behavior into predefined categories using machine learning models.

3.2 SPECIFIC OBJECTIVES

1. Build a Machine Learning Model

- Develop and train a machine learning model to classify driver behavior into categories such as aggressive, moderate, and cautious. The model will rely on patterns identified in driving data to generate meaningful classifications that reflect real-world driving behaviors.

2. Utilize Accessible Data Sources
   - Leverage non-invasive, widely accessible technologies, such as mobile phone sensors (accelerometers, GPS), to collect driving data without the need for expensive telematics systems.

3. Integrate Simulated Data
   - Incorporate simulated driving patterns generated using the CARLA simulator to enhance data diversity and robustness. This integration will address potential gaps in real-world datasets, ensuring the system performs reliably in a variety of driving scenarios.

4. Ensure Affordability and Scalability
   - Design the system to be both affordable and scalable, allowing it to be deployed across different vehicle types, including non-connected vehicles, and making it practical for use in low-resource environments.

5. Evaluate and Validate System Performance
   - Conduct comprehensive testing and validation of the system in both real-world and simulated environments. This process will ensure accuracy, usability, and applicability for real-life driving scenarios, while identifying potential areas for improvement.

4. METHODOLOGY

The proposed methodology for driver behavior analysis integrates data collection, preprocessing, model training, and evaluation. By combining mobile sensor data with simulated data from the CARLA driving simulator, the research ensures robust and diverse datasets for training machine learning models. The steps in the methodology are as follows:

1. Data Collection
    - Real-World Data: Use mobile phone sensors (accelerometer and GPS) to collect driving metrics such as acceleration, braking, and speed.
    - Simulated Data: Generate diverse and realistic driving scenarios using the CARLA simulator to simulate aggressive, moderate, and cautious driving behaviors.
    - Data Integration: Combine real-world and simulated data to create a comprehensive dataset for model training and validation.

2. Data Preprocessing
    - Normalize and clean the collected data to ensure consistency and compatibility.
    - Extract relevant features such as average speed, braking frequency, and acceleration patterns.
    - Label data into categories (e.g., aggressive, moderate, cautious) based on predefined thresholds.

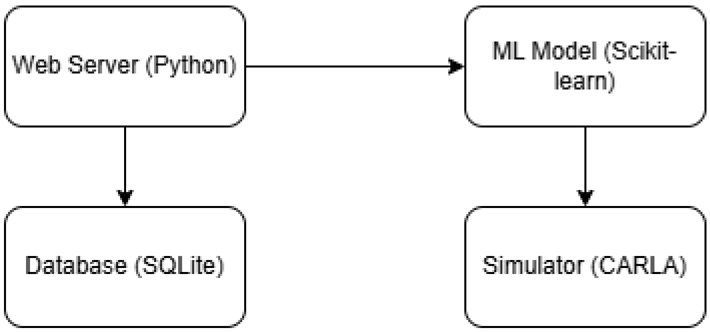
3. Machine Learning Model Development
    - Use supervised learning algorithms like Decision Trees and Support Vector Machines (SVM) for classification.
    - Train the models using the preprocessed dataset and tune hyperparameters for optimal performance.
    - Validate the model using a test dataset to ensure accuracy and robustness.

4. System Implementation
    - Build a web-based backend system using Python and SQL to manage data and run the models.
    - Implement the system to classify new driving data in real-time, providing results to end-users through a dashboard.

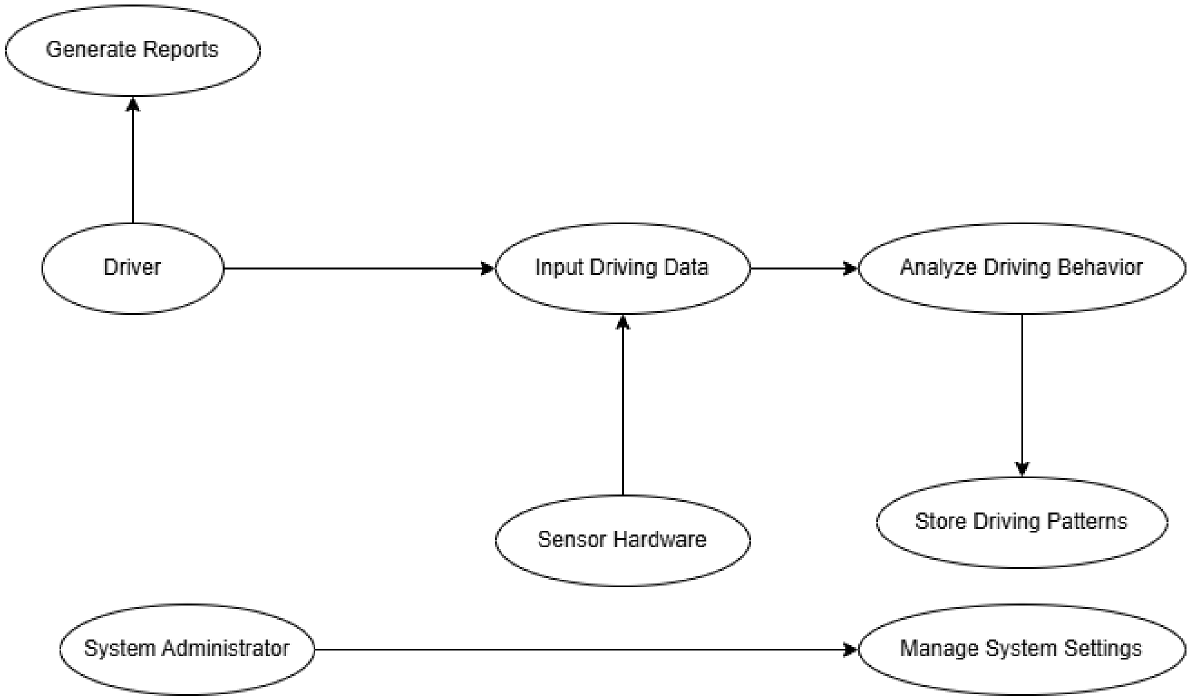
5. Evaluation and Testing
    - Assess the system's performance using metrics such as accuracy, precision, recall, and F1 score.
    - Conduct usability testing to ensure the system is user-friendly and meets functional requirements.

6. Deployment
   - Deploy the system in a scalable manner, ensuring it can handle diverse vehicle types and environments.
   - Integrate security and privacy measures to protect user data during collection, storage, and processing.

4.1 SYSTEM OVERVIEW DIAGRAM



4.2 USE CASE DIAGRAM



4.3 COMMERCIALIZATION

We hope to commercialize our driving behavior classification system by targeting a wide range of industries, including automotive manufacturers, fleet operators, and insurance companies. By providing a robust solution that uses advanced machine learning algorithms to classify driving behavior in real time, we aim to enhance vehicle safety, reduce operational risks, and improve road safety. Our revenue model includes subscription-based plans for fleet operators and insurers, licensing agreements with automotive manufacturers to integrate the system into vehicles, and a freemium model for individual consumers who want basic behavior insights with the option to upgrade for advanced analytics.

To gain market traction, we plan to initiate pilot programs with fleet operators and insurance companies to demonstrate the effectiveness of the system. Additionally, we intend to establish partnerships with automotive companies and telematics providers for seamless integration into existing vehicle platforms. Marketing efforts will include participation in industry expos, targeted advertising, and social media campaigns to generate awareness. As part of our long-term strategy, we aim to scale the solution globally by adapting it to comply with regional driving regulations and traffic conditions. Furthermore, we envision expanding into adjacent markets, such as autonomous vehicle development and smart transportation systems, while continuously enhancing the system with predictive analytics and other innovative features.

With a focus on scalability, cost-efficiency, and privacy compliance, we aim to position our system as a leading solution for driving behavior analysis, offering value to businesses and individuals alike while contributing to safer roads and better driving habits worldwide.

5. REQUIREMENTS

5.1. FUNCTIONAL REQUIREMENTS

1. Driving Behavior Classification:

   - The system should classify driving behavior into categories such as *aggressive*, *moderate*, and *cautious*.
   - Use supervised learning algorithms like decision trees or SVM for this purpose.

2. **Real-Time Analysis:**

   - Analyze driving data in real-time to provide immediate feedback.

3. **Data Collection and Integration:**

   - Collect data from sensors, GPS, and accelerometers for analysis.
   - Integrate with vehicle systems for seamless data acquisition.

4. **Behavior Insights and Reports:**

   - Generate visual reports and analytics to summarize driving behavior trends.

5. **Alert Notifications:**

   - Notify users or fleet managers in case of aggressive or unsafe driving.

5.2. NON-FUNCTIONAL REQUIREMENTS

1. **Scalability:**
   - Ensure the system is scalable for deployment across various vehicle models and fleet sizes.

2. **Efficiency:**
   - Minimize computational and sensor costs while maintaining accuracy.

3. **Security and Privacy:**
   - Ensure that all driving data collected is securely stored and complies with privacy regulations.

4. **Usability:**
   - Provide a user-friendly interface for users to access insights and analytics.

5. **Reliability:**
   - Ensure the system operates with minimal downtime and provides accurate classifications under varying conditions.

5.3. SYSTEM REQUIREMENTS

1. **Hardware Requirements:**
   - Sensors: Accelerometer, GPS, and vehicle speed sensors.
   - Computing Unit: Onboard system or edge device for real-time processing.
2. **Software Requirements:**
   - ML Libraries: Scikit-learn, Pandas, and Matplotlib.
   - Tools: CARLA Simulator, Jupyter Notebook for development.
   - Backend: Python-based web server with SQL database (SQLite).
3. **Network Requirements:**
   - Internet connection for cloud storage or remote monitoring (optional for local systems).

5.4. USER REQUIREMENTS

1. Fleet Managers:
   - Monitor driver behavior and receive alerts for unsafe driving.
   - Access detailed reports and analytics for decision-making.
2. **Individual Drivers:**
   - Receive real-time feedback to improve driving habits.
   - View summaries of their driving performance over time.
3. **Insurance Companies:**
   - Use driving behavior data for personalized premium calculations.
4. **Automotive Manufacturers:**
   - Integrate the system into vehicle models for added safety features.

5.5. USE CASES

Use Case ID	FM01 (Fleet Managers)
Name	Driving Behavior Analytics
Summary	The system provides comprehensive analytics on the driving behavior of the fleet drivers to improve safety and efficiency.
Priority	High
Preconditions	- Fleet vehicles are equipped with OBD-II devices. - Historical driving data is collected for all fleet drivers.
Postconditions	- Fleet managers receive detailed driving reports. - Recommendations are generated for optimizing fleet operations.
Primary Actor(s)	Fleet Manager
Secondary Actor(s)	System
Trigger	The fleet manager requests analytics or the system generates periodic reports.
Main Scenario	- Fleet manager accesses the dashboard or mobile app. -The system processes historical and real-time driving data. -Behavior analytics are categorized into risk levels (e.g., aggressive, moderate, cautious). -The system highlights trends and outliers. -Recommendations (e.g., training programs) are displayed.
Extensions	- If data for certain drivers is missing, the system notifies the fleet manager. - If risk levels exceed thresholds, alerts are sent proactively.
Open Issues	- Customizing reports for various fleet sizes.

Use Case ID	ID01 (Individual Drivers)
Name	Driving Insights and Feedback
Summary	The system provides personalized driving insights to individual drivers based on their behavior to promote safety and efficiency.
Priority	High
Preconditions	- The driver's vehicle is equipped with the OBD-II device. - Real-time driving data is being captured.
Postconditions	- The driver receives actionable insights to improve their driving behavior. - Data is logged for future trend analysis.
Primary Actor(s)	Individual Driver
Secondary Actor(s)	System
Trigger	The system detects driving behavior requiring feedback or the driver requests insights.
Main Scenario	-The system captures real-time driving patterns (e.g., speeding, hard braking). -Driving behavior is classified and compared to safety standards. -Insights are displayed on the driver's dashboard or app.

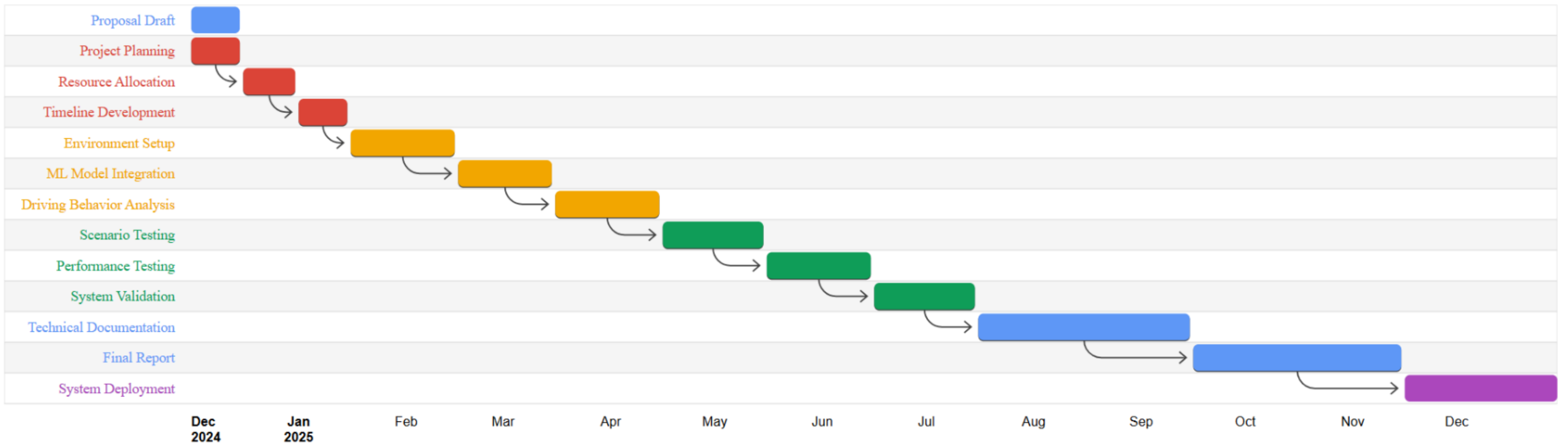
	-Recommendations are provided, such as smoother acceleration or speed reduction.
Extensions	- If connectivity is lost, insights are queued and synced later. - If the driver shows improvement, the system highlights progress for motivation.
Open Issues	- Avoiding excessive notifications to minimize distractions.

Use Case ID	IC01 (Risk Assessment Reports)
Name	Risk Assessment Reports
Summary	The system generates detailed risk reports for insurance companies based on driver behavior analytics to assess premiums.
Priority	Medium
Preconditions	Driver behavior data is shared with insurance companies (with consent).
Postconditions	Insurance companies receive driver risk profiles for premium calculation.
Primary Actor(s)	Insurance Company
Secondary Actor(s)	System
Trigger	Insurance companies request driver risk data or the system periodically generates risk reports.
Main Scenario	-The system aggregates driving behavior data (e.g., frequency of speeding or harsh braking). -Risk levels are calculated based on predefined criteria. -Reports with risk classifications are shared with insurance companies.
Extensions	- If data sharing is not consented, the system anonymizes data before sharing. - If risk criteria change, reports are updated dynamically.
Open Issues	- Addressing privacy concerns related to driver data sharing.

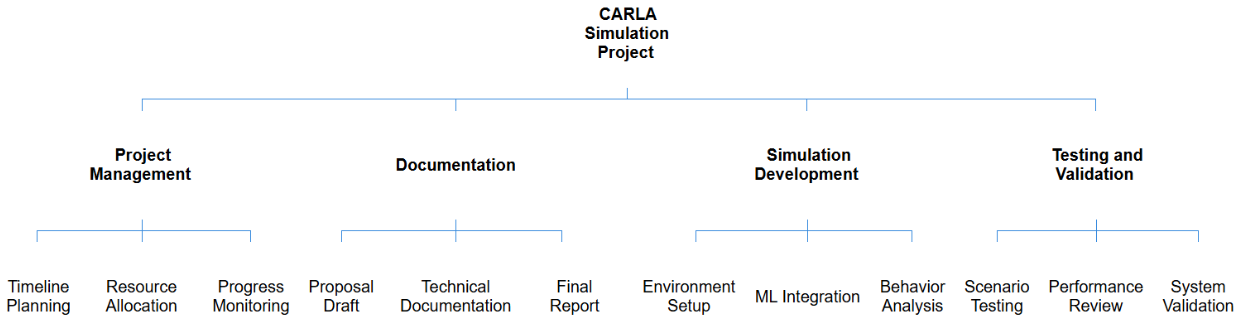
Use Case ID	AM01 (Automotive Manufacturers)
Name	Feature Enhancement Insights
Summary	The system provides aggregated insights from driving behavior data to automotive manufacturers for improving vehicle features.
Priority	Medium
Preconditions	- Manufacturers have agreements to collect anonymized driving data. - Sufficient aggregated data is available.
Postconditions	- Manufacturers receive actionable insights for R&D.
Primary Actor(s)	Automotive Manufacturer
Secondary Actor(s)	System
Trigger	Manufacturers request insights or the system generates periodic reports.

Main Scenario	-The system aggregates data across vehicles (e.g., component wear, driving trends). -Insights are generated, such as common issues with certain vehicle parts. -Manufacturers use insights to improve designs or add features (e.g., better braking systems).
Extensions	- If data from certain regions is missing, manufacturers are notified to address gaps. - Manufacturers can request tailored insights based on vehicle models.
Open Issues	- Managing the scalability of data aggregation across global markets.

6. GANTT CHART



7. WORK BREAKDOWN STRUCTURE



8. REFERENCES

[1] Lindow, F., & Kashevnik, A. (2020). *Driver Behavior Monitoring Based on Smartphone Sensor Data and Machine Learning Methods*. Link

[2] Johnson, D. A., & Trivedi, M. M. (2013). *Driving Style Recognition Using a Smartphone as a Sensor Platform*. Link

[3] Ben Brahim, S., Ghazzai, H., Besbes, H., & Massoud, Y. (2022). *A Machine Learning Smartphone-based Sensing for Driver Behavior Classification*. Link

9. APPENDIX

Tools and Technologies

1. CARLA Simulator - Used to simulate driving scenarios and collect data.
2. Python - For developing machine learning models and system implementation.
3. Scikit-learn - For supervised learning algorithms.
4. Pandas - For data analysis and manipulation.
5. Matplotlib - For data visualization.
6. SQLite - For database management.
7. Jupyter Notebook - For interactive development and testing.

Abbreviations

Abbreviation	Definition
ML	Machine Learning
OBD-II	On-Board Diagnostics II
SVM	Support Vector Machine