DEPARTMENT OF COMPUTER SCIENCE

MSc PROJECT(COMP702)

Dense reinforcement learning for safety validation of autonomous vehicles (#4)

SPECIFICATION AND PROPOSED DESIGN

Student
Vikhyat Bisht
201687684

Primary Supervisor
Wiley Ruan

Secondary Supervisor
Xiaowei Huang

PROJECT DESCRIPTION

NON-TECHNICAL SUMMARY:

The goal of this research is to improve the safety of autonomous vehicles by building a reinforcement learning model that teaches autonomous vehicles to make safer judgements. The reinforcement learning model will interact with the self-driving car in CARLA, a virtual simulation environment that faithfully simulates real-world driving situations.

By interacting with the environment and getting feedback, the reinforcement learning model will learn to make safer judgements. It will also observe other cars and pedestrians in the simulation to learn how to interact with them properly. The project's ultimate objective is to develop a safer and more dependable autonomous car system.

TECHNICAL SUMMARY:

The major goal of this research is to use reinforcement learning (RL) models to solve the safety problems connected with autonomous cars. RL is a sort of machine learning in which an agent learns to make decisions through interaction with its surroundings. In our scenario, the agent is an autonomous car navigating the CARLA simulation environment, which accurately simulates real-world driving situations.

The RL model will be trained to manage the self-driving car using a reward mechanism that encourages safe driving habits. This involves observing traffic laws, avoiding crashes, maintaining lane discipline, and staying under speed limits. The agent examines the surroundings, including the vehicle's speed, neighboring cars, pedestrians, and traffic signals, before taking actions such as accelerating, decelerating, or turning.
Simultaneously, we will use CARLA's extensive sensor suite, which includes LIDAR, RADAR, and cameras, to sense the surroundings. With this precise vision and our reinforcement learning model, the autonomous car will be able to learn from its encounters and enhance its safety measures over time.

This effort has made a substantial addition to the larger body of knowledge. It highlights the practical use of reinforcement learning in enhancing the safety of autonomous cars, a topic of great relevance given their potential role in our transportation networks. The outcomes of the initiative might give useful insights to researchers, the automotive industry, and regulatory agencies, therefore advancing autonomous vehicle technology.

AIMS AND OBJECTIVES

AIMS:

1. Creating an efficient reinforcement learning model for simulated autonomous vehicle control.
2. Decrease the frequency of traffic infractions and crashes to increase the safety of autonomous cars.
3. Examine how different traffic circumstances affect the effectiveness of the reinforcement learning model.
4. Contribute to the continuing research on reinforcement learning and autonomous cars, offering insightful information to both businesses and academics.

OBJECTIVES:

1. Setting up the Simulation Environment:
   Use the CARLA simulator, an open-source, highly accurate platform for autonomous vehicle research. Within the simulator, link the pedestrian agents with the autonomous cars.
2. Developing the Reinforcement Learning Model:
   Use a reinforcement learning technique that is appropriate for teaching autonomous cars.
   Establish a suitable incentive system that promotes defensive driving.
   In the CARLA simulation environment, train the model.
3. Sensor integration implementation:
   LIDAR, RADAR, and camera data should all be included to provide the learning model a thorough understanding of the surroundings.

4. Evaluation and testing:
   Test the trained model under various traffic conditions in the simulator to gauge how well it performs. Success is determined by the decline in traffic infractions, the prevention of crashes, and adherence to traffic regulations.

5. Reporting and Documentation:
   Keep track of the project's steps, obstacles, fixes, and outcomes. For both academic and business audiences, provide the findings in a thorough and understandable way, including a complete analysis and performance indicators.
   We anticipate having a fully trained reinforcement learning model that improves the safety of autonomous cars in the simulation environment at the conclusion of this research. The effectiveness of this model will be assessed using a variety of metrics and situations, and the results will show how well-suited it is for practical use.

KEY LITERATURE AND BACKGROUND READING

The area of autonomous vehicle (AV) safety validation is wide and complex, spanning multiple fields like computer science, robotics, artificial intelligence, machine learning, and others. A wide range of literature has had a considerable impact on our knowledge and approach to this topic.

"Dense Reinforcement Learning for Safety Validation of Autonomous Vehicles" (2023) is a seminal study that has greatly influenced the technique of our research. The study emphasizes the high costs of verifying the safety of autonomous cars, owing to the rarity of safety-critical occurrences in realistic driving conditions. The authors created an intelligent testing environment by using artificial intelligence (AI)-based background agents that were trained to perform adversarial maneuvers using a Dense Deep-Reinforcement-Learning (D2RL) approach, thereby accelerating the safety validation process without sacrificing unbiasedness.

"Deep Reinforcement Learning for Autonomous Driving: A Survey" (Li, 2019) provides a thorough examination of the most recent deep reinforcement learning approaches used in autonomous driving. The survey thoroughly addressed all essential components, from decision-making to control levels, guiding our decision-making process and assisting in the creation of our project's RL model.

"CARLA: An Open Urban Driving Simulator" (Dosovitskiy et al., 2017) provided a comprehensive overview of CARLA, the simulator we're employing for this study, as well as its relevance and usefulness in AV research. This paper's knowledge has considerably informed our approach to the project, providing us with a clear image of the simulator's capabilities and possibilities.

The paper "Physics-based Simulation of Continuous-Time Trajectories for Large-Scale Traffic Scenario Validation" (Noe, 2020) demonstrated how physics-based simulations may be used to validate large-scale traffic scenarios. The approaches and ideas mentioned in this work have aided in the development of our approach to incorporating realistic physics into our simulations.

The paper "Virtual to Real Reinforcement Learning for Autonomous Driving" (Zhang et al., 2020) described how reinforcement learning models developed in a simulated environment were transferred to real-world testing. The technique provides guidance on how to develop our reinforcement learning algorithm while considering both the complexity of the real-world environment and the constraints of simulation.

The paper "Testing and Validating High-Definition Maps for Autonomous Driving" (Yan et al., 2018) emphasized the necessity of HD maps in AV navigation and testing. This literature taught us how to add HD maps into our simulation, boosting the realism and autonomous driving capabilities of our agent.

In addition, the Waymo Open Dataset: Building, Testing, and Validating Self-Driving Algorithms (Waymo, 2020) supplied us with high-resolution sensor data from numerous self-driving car scenarios. The data collection, which included LiDAR and radar data, helped us better learn how to manage, interpret, and integrate sensor data into our project.

SUMMARY OF ACTIVITIES SINCE THE START OF THE PROJECT:

Our early actions included creating the simulation environment and setting up the environment for development and testing, which included selecting the appropriate tools and platforms and configuring the simulation environment. This necessitated extensive investigation and evaluation since the tools and platforms used might have a considerable impact on the efficacy and speed of the validation process.

Following that, we investigated cutting-edge methodologies in AV safety validation. This method entailed reading and comprehending critical literature before applying this information to our conclusions. The literature has played an important role in developing our approach to building and training our AI agents, as well as conceptualizing the development of an expedited testing environment.

Finally, we started creating and implementing scripts for configuring and controlling the AI agents in the simulation environment. defining the AI agents' behavior, defining their interactions with the autonomous vehicle, and implementing the reinforcement learning technique for training the AI agents were all part of this. Our actions have resulted in an effective and efficient approach for AV safety validation, guided by the results of the literature.

DEVELOPMENT AND IMPLEMENTATION SUMMARY

The main goal of our project is to make a system for autonomous vehicles (AVs) that is efficient and reliable. We will do this by using a type of reinforcement learning method called Advantage Actor-Critic (A2C) in an urban driving simulator called CARLA. In the parts that follow, we'll talk more about our concept, our working environment, and the process of making it happen.

SYSTEM OVERVIEW:

The most important part of our project is a complex system that lets an AV connect with its surroundings in a way that fits its needs, make decisions based on what it knows, and learn from the results of its actions. This high-level connection between different parts of the system can be shown as follows:

DEVELOPMENT ENVIRONMENT AND PROGRAMMING LANGUAGE:

We chose Python for implementation because it is easy to read and use and has a lot of tools for machine learning and data analysis. Also, Python is a great choice for making our A2C code because it works well with the PyTorch system. The CARLA simulator was chosen because it is reliable and has a lot of pre-made urban situations that make it a very useful tool for study into self-driving cars.

WORKFLOW ORGANIZATION:

The project's completion is broken down into well-defined steps:
1. Data collection: During this step, the CARLA model is used to gather information about the senses. Images, lidar readings, and other sensor data from different traffic situations make up most of the data.
2. Model development: At this point, we use PyTorch to make our A2C method work. The main goal is to build an AI model that can use the collected data to make good decisions about driving.
3. Training: During the training phase, the A2C model gradually improves its performance based on feedback from the surroundings. This is done through a process called "iterative learning."
4. Testing: After the training, we use CARLA to try our A2C model in different traffic situations to see how flexible and effective it is.
5. Evaluation: In the last step, we use different measures, such as the number of accidents, the number of times traffic rules are followed, and the general efficiency, to evaluate how well our model works.

ALGORITHM PSEUDOCODE:

The A2C learning algorithm can be presented as the following pseudocode:
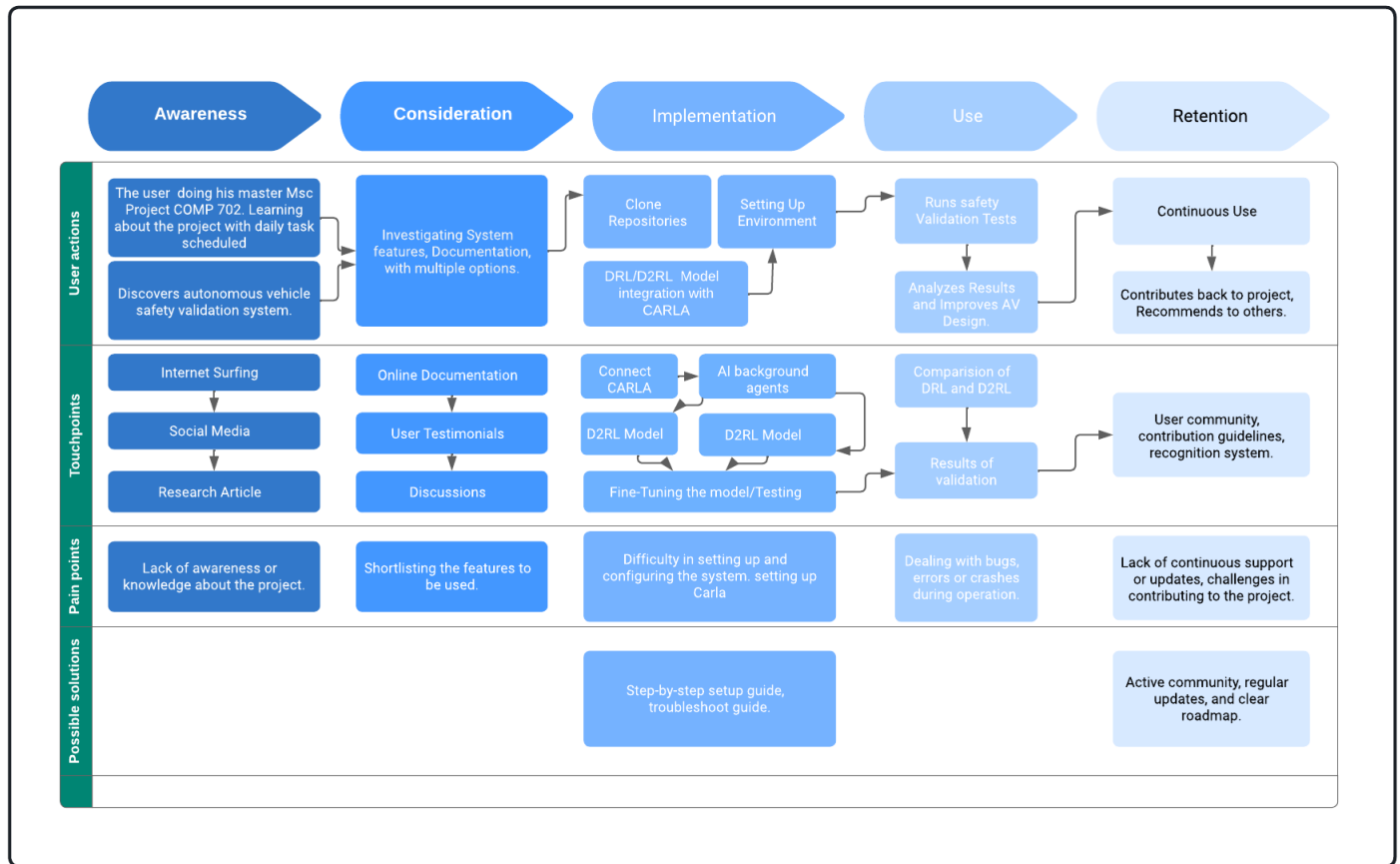
```
Initialize actor and critic networks
for each episode do
    Initialize state
    for each step do
        Select action from actor network given current state
        Execute action in the environment
        Observe reward and new state
        Compute advantage and target value for critic
        Update actor and critic networks using advantage and target value
    end for
end for
```

EXPERIMENT DESIGN:
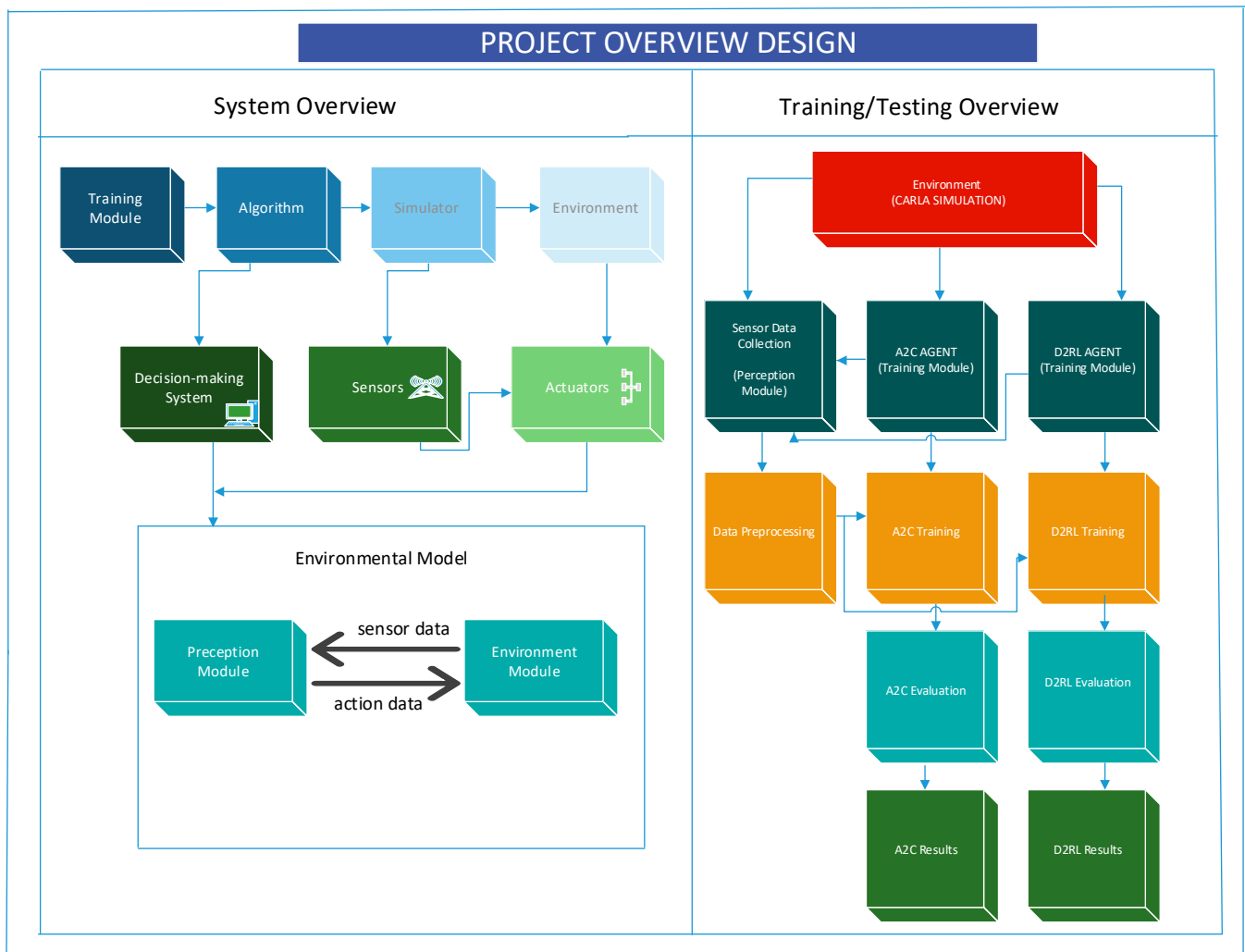
To make sure that our model works and is reliable, we plan to run several tests in which the vehicle moves through different pre-set situations in the CARLA simulator. The vehicle's performance will be judged by how well it follows traffic rules, how often it gets into accidents, and how quickly and safely it gets to its goal generally. These results will be analyzed statistically to learn more about how well the model works and where it could be improved.
In short, this project is a good mix of theoretical machine-learning concepts and real-world use. By using the A2C algorithm well in the CARLA simulator, we hope to make a big difference in the field of safety evaluation for driverless vehicles.

5-STAGE USER JOURNEY MAP:

	Awareness	Consideration	Implementation	Use	Retention
User actions	The user doing his master Msc Project COMP 702. Learning about the project with daily task scheduled / Discovers autonomous vehicle safety validation system.	Investigating System features, Documentation, with multiple options.	Clone Repositories / Setting Up Environment / DRL/D2RL Model integration with CARLA	Runs safety Validation Tests / Analyzes Results and Improves AV Design.	Continuous Use / Contributes back to project, Recommends to others.
Touchpoints	Internet Surfing / Social Media / Research Article	Online Documentation / User Testimonials / Discussions	Connect CARLA / AI background agents / D2RL Model / D2RL Model / Fine-Tuning the model/Testing	Comparision of DRL and D2RL / Results of validation	User community, contribution guidelines, recognition system.
Pain points	Lack of awareness or knowledge about the project.	Shortlisting the features to be used.	Difficulty in setting up and configuring the system. setting up Carla	Dealing with bugs, errors or crashes during operation.	Lack of continuous support or updates, challenges in contributing to the project.
Possible solutions			Step-by-step setup guide, troubleshoot guide.		Active community, regular updates, and clear roadmap.

USER INTERFACE MOCKUP

For our safety validation system for autonomous vehicles, we foresee a straightforward and intuitive user interface. This interface will predominantly serve two functions:

1.Provide a visual representation of the simulation environment in the CARLA simulator, including the autonomous vehicle, other traffic participants, and the road network.

2.Display real-time data and metrics, including the current state of the autonomous vehicle (e.g., its speed and location), its actions, and the reward it receives at each time step. It may also display performance metrics, such as the total reward earned during an episode, the number of collisions, and other safety or productivity indicators.

It is anticipated that the UI will be incorporated into the CARLA simulator, which already provides some UI capabilities. However, we may need to modify it to suit our needs.

MOCK-UP OF THE USER INTERFACE:

Our proposed user interface consists of two principal panels:
Left Pane: Simulation Perspective

This is a real-time 3D visualization of the simulation. It displays the autonomous vehicle, other vehicles, pedestrians, and road infrastructure. The highlighted autonomous vehicle facilitates monitoring. Right Panel: Data and Metrics

This panel displays data and metrics in real-time. It could consist of:

1. Current State: Information regarding the current state of the autonomous vehicle (e.g., position, speed, steering angle).
2. Actions: The actions performed by the vehicle at each time step (e.g., acceleration, deceleration, and turning).
3. Reward: The compensation received by the vehicle at each time phase.
4. Performance Metrics: Overall performance metrics, including the total reward accumulated over the course of an episode, the number of collisions, and the average pace, among others.

The user spends time becoming acquainted with the system. They review the project documentation, the source code, the results of any published validation testing, and compare it to other alternatives. They may also conduct preliminary experiments with the pre-trained models or post queries on the community forum or discussion board for the project.

DATA SOURCES:

Our initiative focuses predominantly on the development of a reinforcement learning model for validating the safety of autonomous vehicles and does not use any external datasets for training the model.

Nevertheless, we utilize the open-source CARLA simulator to create and evaluate our model in realistic urban environments. CARLA includes a plethora of assets, such as diverse vehicle classes, pedestrians, city configurations, and weather conditions. Even though these are not traditional "datasets," they constitute the environment in which our model operates and thus influence its learning process.

We chose CARLA because of its open-source nature, realism, and wealth of built-in assets. By utilizing an open-source simulator, we ensure that others in the research community can access and reproduce our project.
As the CARLA simulator does not contain any personally identifiable information, our endeavor has no confidentiality or data privacy concerns.

TESTING:

Multiple components of our project necessitate thorough testing, and our testing strategy incorporates unit tests and integration tests to ensure that the system functions accurately as a whole.

1. Each component, including the A2C model, the CARLA simulator environment, the traffic management system, etc., will be subjected to independent unit testing. In the case of the A2C model, for example, we will ensure that the model can accurately update its parameters based on a set of experiences and output plausible actions based on the input state.

2. Integration Testing: After each component has passed unit testing, we will undertake integration tests to ensure that all components function as intended when combined. This entails ensuring that the model can interact correctly with the CARLA simulator and that the traffic management system can control non-player vehicles precisely based on the model's actions.

3. Simulation Testing: We will conduct a series of simulations in the CARLA environment to evaluate the ability of our autonomous vehicle to navigate a variety of scenarios, including varying traffic conditions and weather conditions. The objective of these evaluations is to evaluate the overall system's robustness and

efficacy.

4. Stress Testing: Additionally, we will stress test our system by exposing it to the most difficult circumstances, such as heavy traffic, poor weather, and aggressive non-player vehicle behavior. This test will help us comprehend our system's limitations and identify areas for improvement.

5. Performance Testing: During simulations, we will monitor the system's resource utilization to identify any performance bottlenecks. This test is designed to ensure that the system functions efficiently and does not deplete resources unnecessarily.

Due to the simulation-based nature of our project, we do not intend to use external beta evaluators to adhere to ethical standards. All testing will be performed internally utilizing predefined simulation scenarios within the CARLA simulator. This strategy guarantees that our testing process poses no risk to people, property, or the environment, as all testing is performed in a virtual environment.

EVALUATION:

Our system for validating the safety of autonomous vehicles will be judged based on the following criteria:

1. Safe Navigation: The project's primary objective is to ensure the autonomous vehicle's safe navigation in a variety of traffic conditions. During simulations, the success of a system will be determined by the number of traffic violations and accidents that occur; a successful system will have a low incidence of these incidents.

2. Effective Learning: Another crucial success factor is the AI model's capacity to learn and adapt to new traffic scenarios and unforeseen occurrences. This will be determined by the number of episodes required for the model to navigate without committing traffic violations or causing accidents.

3. Robustness: The system's robustness will be evaluated based on the model's ability to manage a variety of scenarios, including various weather conditions, illumination conditions, and traffic densities. This will be evaluated based on the efficacy of the model in these various scenarios.

4. Real-time Performance: An autonomous driving system must have the ability to operate and make decisions in real-time. Consequently, system latency and processing time will be assessed.

Multiple members of our team will be responsible for evaluating various system components during the evaluation procedure. We will use automated testing tools to objectively evaluate the system's efficacy based on the criteria.

Due to the simulation-based nature of the project, no external beta evaluators will be used, ensuring that the evaluation portion of the project is conducted ethically. All information obtained during the evaluation process will be anonymized and used exclusively for system enhancement.

We anticipate that the evaluation will disclose the current implementation's strengths and weaknesses, providing us with a clear path for further development and fine-tuning. A successful outcome would be a system that can navigate a variety of traffic scenarios within the simulation environment safely and efficiently.

ETHICAL CONSIDERATIONS:

In our efforts to create an AI-based safety validation system for autonomous vehicles, we are committed to upholding the highest ethical standards throughout the duration of the project. We have read and comprehended the supplied ethical guidelines and are committed to adhering to them.

Key ethical considerations for our project include:

1. Data Use and Privacy: Our initiative is simulation-based and does not utilize real-world data or personal information. Any data generated throughout the duration of the project, including performance metrics and

simulation results, will be anonymized, and stored securely to maintain confidentiality.

2. Testing and Evaluation: All tests and evaluations are conducted in a virtual environment with no human evaluators involved. To ensure objectivity and impartiality, the evaluation of the system is solely based on predefined criteria and performance metrics.

3. Potential Misuse: The software we are developing is intended for simulating the safety trials of autonomous vehicles. We are aware of the potential hazards associated with misusing the software, such as using it to create dangerous situations on actual roads, and we have taken measures to prevent such misuse, such as providing thorough documentation and warnings about misuse.

4. Safety: Ensuring the safety of prospective future users of autonomous vehicles is one of the primary ethical considerations in our endeavor. The system is being developed with a strong emphasis on preventing any user-harming situations.

In conclusion, we commit to maintaining transparency, respecting privacy, and prioritizing safety while remaining cognizant of potential ethical issues that may arise during the development and implementation of our project.

PROJECT PLAN:

MSc Project - COMP702
Project Start Date : 05/06/2022 (Week 1)
Project Start Date : 23/09/2022 (Week 16)

Task	Status	Start Date	End Date	Time Period (in Weeks)							
				Week 1 & 2	Week 3 & 4	Week 5	Week 6	Week 7-12	Week 13	Week 14 &15	Week 16
Background reading and literature review.	DONE	05.06.2023	18.06.2023								
Project specification and design	DONE	19.06.2023	02.07.2023								
Finalize the development spec and design	DONE	03.07.2023	09.07.2023								
Specification and Design report and presentation slides	MILESTONE	14.07.2023	14.07.2023			Milestone 1 - 14/07/2023					
Implentation &Testing of project simulation design	PROGRESS	10.07.2023	16.07.2023								
Continues Implementation And Testing Carla Simulation	TO DO	17.07.2023	20.08.2023								
Finalize software implementation and testing.	TO DO	21.08.2023	27.08.2023								
Final presentation slides deadline	MILESTONE	02.09.2023	02.09.2023			Milestone 2 - 02/09/2023					
Start with the dissertation write-up.	TO DO	28.08.2023	17.09.2023								
Finalize the write-up of the dissertation.	TO DO	18.09.2023	22.09.2023								
Dissertation deadline	MILESTONE	23.09.2023	23.09.2023			Milestone 3 - 23/09/2023					

RISKS AND CONTINGENCY PLANS:

Risks	Contingencies	Likelihood	Impact
Hardware failure	Regular backups of code and data. Immediate replacement of failed hardware.	Low	High
Software failure	Use of version control systems and regular testing.	Medium	High
Running out of time	Proper time management, task prioritization and efficient use of resources.	High	Very High
Programming problems	Regular code review and debugging, seeking help when needed.	Medium	High
Illness	Make use of downtime to rest and recover and adjust project timeline as needed.	Low	Medium
Data loss	Regular backups of all work, use of reliable storage solutions.	Low	Very High

Dependence on external tools/APIs	Have backup tools/APIs in place, familiarize with more than one tool.	Medium	Medium
Internet connectivity issues	Have a backup internet source, perform tasks that don't require internet during outage.	Low	Medium
Difficulties with the A2C Algorithm	Study the algorithm extensively before implementation. If issues persist, consult with peers, advisors, or online communities.	Medium	Medium
Unexpected Results from AI Training	Set checkpoints during training, allowing to revert to previous states. Run preliminary tests with different parameters to choose the most promising configuration.	High	Medium

REFERENCES:

1. A. Ferdowsi, U. Challita, W. Saad and N. B. Mandayam, "Robust Deep Reinforcement Learning for Security and Safety in Autonomous Vehicle Systems," 2018 21st International Conference on Intelligent Transportation Systems (ITSC), Maui, HI, USA, 2018, pp. 307-312, Doi: 10.1109/ITSC.2018.8569635.

2. Feng, S., Sun, H., Yan, X. et al. Dense reinforcement learning for safety validation of autonomous vehicles. Nature 615, 620–627 (2023). https://doi.org/10.1038/s41586-023-05732-2

3. Dosovitskiy, A., Ros, G., Codevilla, F., Lopez, A. & Koltun, V.. (2017). CARLA: An Open Urban Driving Simulator. <i>Proceedings of the 1st Annual Conference on Robot Learning</i>, in <i>Proceedings of Machine Learning Research</i> 78:1-16 Available from https://proceedings.mlr.press/v78/dosovitskiy17a.html.

4. S. Kumar, L. Zhao, and X. Fernando, "Multi-Agent Deep Reinforcement Learning-Empowered Channel Allocation in Vehicular Networks," in IEEE Transactions on Vehicular Technology, vol. 71, no. 2, pp. 1726-1736, Feb. 2022, Doi: 10.1109/TVT.2021.3134272.

5. You, Yurong & Pan, Xinlei & Wang, Ziyan & Lu, Cewu. (2017). Virtual to Real Reinforcement Learning for Autonomous Driving.

6. Fischer, Peter & Plaß, Bastian & Krauß, Thomas & Kurz, Franz & Runge, Hartmut. (2017). Validation of HD maps for autonomous driving.

7. LeCun, Yann & Bengio, Y. & Hinton, Geoffrey. (2015). Deep Learning. Nature. 521. 436-44. 10.1038/nature14539.

8. Feng, Shuo & Feng, Yiheng & Yan, Xintao & Shen, Shengyin & Xu, Shaobing & Liu, Henry. (2020). Safety assessment of highly automated driving systems in test tracks: A new framework. Accident Analysis & Prevention. 144. 105664. 10.1016/j.aap.2020.105664.

9. Kalra, Nidhi & Paddock, Susan. (2016). Driving to safety: How many miles of driving would it take to demonstrate autonomous vehicle reliability?. Transportation Research Part A: Policy and Practice. 94. 182-193. 10.1016/j.tra.2016.09.010.

10. Sutton, Richard & Barto, Andrew. (1998). Reinforcement Learning: An Introduction. IEEE transactions on neural networks / a publication of the IEEE Neural Networks Council. 9. 1054. 10.1109/TNN.1998.712192.