REPORT

# Multi-Objective Deep Reinforcement Learning for Safe and Eﬃcient Autonomous Highway Driving with

**Adaptive Safety Module and Prioritized Experience Replay**

Name: Jeevan M

Department: Computer Science and Engineering Year: 2nd year

College Name: Indian Institute of Information Technology Sri City Contact Details: 9003235428

Internship Duration: 19/5/2025 to 19/6/2025

Name: Darshan R

Department: Computer Science and Engineering Year: 2nd year

College Name: Indian Institute of Information Technology Sri City Contact Details: 7904599205

Internship Duration: 19/5/2025 to 19/6/2025

**TABLE OF CONTENTS –**

1. Problem Statement
2. Abstract
3. Introduction
4. Literature Review
5. Proposed Methodology
   * System Architecture
   * Adaptive Safety Module
   * Deep Q-Network Agent
   * Prioritized Experience Replay
   * Multi-Objective Reward Design
   * Training Procedure
6. Conclusion
7. References
8. Appendix
   * GitHub Link
   * ReadMe
   * Internship Outcome

# ​Problem Statement:

Autonomous vehicles are rapidly advancing and promise to transform the future of transportation by improving safety, eﬃciency, and accessibility. However, the real-world deployment of self-driving cars on highways is fraught with challenges that require sophisticated solutions. Highway environments are complex, dynamic, and unpredictable, requiring vehicles to make split-second decisions while balancing multiple objectives such as safety, comfort, eﬃciency, and fuel economy. Traditional deep reinforcement learning (DRL) approaches often struggle to handle these competing goals, as they typically focus on optimising a single objective and may not generalise well to multi-objective settings. Moreover, ensuring robust learning and sample eﬃciency in safety-critical applications remains a signiﬁcant challenge, as unsafe actions can have catastrophic consequences. This project speciﬁcally addresses the need for a multi-objective DRL framework that can adaptively enforce safety constraints and leverage advanced experience sampling techniques to improve learning eﬃciency. By integrating an Adaptive Safety Module and Prioritised Experience Replay, the proposed solution aims to enable autonomous vehicles to navigate highways safely and eﬃciently, even in dense and unpredictable traﬃc. The ultimate goal is to contribute to the development of autonomous driving systems that are not only safe and eﬃcient but also comfortable and reliable for real-world deployment.

# ​Abstract:

This report presents a comprehensive study on the development and evaluation of a multi-objective deep reinforcement learning framework for autonomous highway driving. The proposed system integrates an Adaptive Safety Module (ASM) to dynamically adjust safety thresholds based on real-time traﬃc context and vehicle state, ensuring that the

agent operates within safe limits at all times. In addition, the framework employs Prioritised Experience Replay (PER) to enhance sample eﬃciency and learning stability by focusing the agent’s attention on critical experiences. The system is implemented using a Deep Q-Network (DQN) architecture, with extensive code-level customisation to balance safety, comfort, eﬃciency, and fuel economy. Experimental results in a simulated highway environment demonstrate signiﬁcant improvements in collision avoidance, lane-change safety, and driving eﬃciency compared to traditional single-objective and ﬁxed-threshold methods. The report provides detailed explanations of the system architecture, code implementation, and experimental setup, along with a thorough analysis of the results. By addressing the limitations of existing DRL approaches, this work contributes to the advancement of safe and eﬃcient autonomous driving technologies and provides a foundation for future research and real-world deployment.

# ​Introduction:

The ﬁeld of autonomous driving has witnessed remarkable progress in recent years, driven by advances in artiﬁcial intelligence, sensor technology, and computational power. Autonomous vehicles (AVs) are expected to revolutionize transportation by reducing accidents, improving traﬃc ﬂow, and increasing accessibility for people with mobility challenges. However, the deployment of AVs in real-world highway scenarios presents numerous technical and practical challenges. Highway environments are characterised by high speeds, dense traﬃc, and unpredictable driver behaviour, making robust decision-making under uncertainty a critical requirement. Deep reinforcement learning (DRL) has emerged as a promising approach for learning complex driving policies from simulated environments, but existing DRL methods often focus on optimizing a single objective, such as minimizing travel time or maximizing safety, without adequately considering the trade-offs between multiple goals. This limitation can lead to suboptimal or unsafe behaviour in complex scenarios where safety, comfort, eﬃciency, and fuel economy must be balanced. To address these challenges, this project proposes a multi-objective DRL

framework that integrates an Adaptive Safety Module (ASM) and Prioritized Experience Replay (PER) within a DQN-based agent. The ASM dynamically adjusts safety thresholds in real-time based on the current traﬃc context and vehicle state, while PER ensures that critical experiences are prioritized during training, leading to faster convergence and improved policy robustness. By combining these innovations, the framework aims to overcome the limitations of traditional DRL methods and contribute to the development of safer and more eﬃcient autonomous highway driving systems.

# ​Literature Review:

Recent literature on autonomous driving and deep reinforcement learning highlights several key trends and challenges that are relevant to this project. Deep Q-Networks (DQN) and their variants have been widely used for end-to-end driving policy learning, enabling agents to map high-dimensional sensory inputs to driving actions. These methods have demonstrated success in various control tasks, including lane keeping, obstacle avoidance, and traﬃc navigation. However, most existing DRL approaches focus on optimizing a single objective, such as minimizing travel time or maximizing safety, and do not explicitly address the need to balance multiple, often conﬂicting, objectives. Multi-objective reinforcement learning (MORL) techniques, such as weighted sum rewards and lexicographic ordering, have been proposed to address this limitation, but their real-world applicability remains limited due to the complexity of reward design and scalability issues. Safety-critical applications require RL agents to respect hard constraints and minimize the risk of unsafe actions, which has led to the development of techniques such as reward shaping, constrained RL, and adaptive regularization. Prioritized Experience Replay (PER) has been shown to enhance learning eﬃciency by sampling transitions with higher expected learning value more frequently, leading to faster convergence and improved policy quality. Adaptive safety modules, which dynamically adjust safety parameters based on context, have also been explored to improve robustness in dynamic environments. Despite these

advancements, integrating multi-objective optimization, adaptive safety, and eﬃcient learning remains an open research challenge, especially for real-world highway driving scenarios. This project builds on these prior works by proposing a comprehensive framework that addresses these challenges through the integration of an Adaptive Safety Module and Prioritized Experience Replay within a DQN-based agent.

# ​Proposed Methodology:

## System Architecture:

The proposed system architecture is designed to address the unique demands of autonomous highway driving by integrating advanced reinforcement learning techniques with adaptive safety mechanisms. At its core, the system consists of three main components: an Adaptive Safety Module (ASM), a Deep Q-Network (DQN) agent with a multi-objective reward function, and a Prioritized Experience Replay (PER) buffer. The ASM continuously monitors the environment and vehicle state, dynamically adjusting safety thresholds such as time-to-collision and headway to ensure safe operation even in dense or unpredictable traﬃc. The DQN agent is responsible for learning a driving policy that optimizes a composite reward function, which balances safety, eﬃciency, comfort, and fuel economy. The PER buffer enhances learning eﬃciency by prioritizing the sampling of experiences with high temporal-difference (TD) error, ensuring that the agent focuses on critical transitions that drive policy improvement. The training procedure involves interacting with a simulated highway environment, where the agent is exposed to a variety of traﬃc scenarios and must learn to navigate safely and eﬃciently. Periodic evaluation and checkpointing are implemented to monitor progress and ensure model stability. The combination of adaptive safety constraints, multi-objective optimization, and prioritized experience replay distinguishes this approach from traditional DRL methods, enabling the agent to learn robust and safe driving policies in complex highway scenarios.

## Adaptive Safety Module:

The Adaptive Safety Module (ASM) is a critical component of the proposed system, responsible for ensuring that the autonomous vehicle operates within safe limits at all times. The ASM continuously monitors the environment, including the positions, speeds, and trajectories of nearby vehicles, as well as the state of the ego vehicle. Based on this information, the module dynamically computes safety thresholds such as time-to-collision (TTC), headway, and scenario risk. These thresholds are adjusted in real-time to reﬂect the current traﬃc density, vehicle speed, and overall risk level. The ASM also penalizes unsafe actions by incorporating safety violations into the reward function, encouraging the agent to prioritize safety over other objectives. By adapting to changing traﬃc conditions, the ASM enables the agent to respond appropriately to both routine and unexpected situations. The module is implemented as a separate component within the overall system architecture, allowing for ﬂexible integration with other modules and easy adaptation to different environments. This design ensures that the autonomous vehicle can operate safely and eﬃciently in a wide range of highway scenarios.

## Deep Q-Network Agent:

The Deep Q-Network (DQN) agent is the core learning component of the proposed system, responsible for mapping observations to actions that maximize the expected cumulative reward. The agent uses a neural network to approximate the Q-value function, with enhancements such as batch normalization and dropout for stability. The DQN agent is trained using a multi-objective reward function that balances safety, eﬃciency, comfort, and fuel economy, ensuring that the agent learns to make decisions that are not only safe but also eﬃcient and comfortable for passengers. The agent employs double DQN updates to reduce overestimation bias in Q-value estimation and uses a target network with soft updates to improve training stability. Periodic checkpointing is implemented to save model weights, replay buffer, and training

metadata, enabling recoverability and facilitating further research. The agent interacts with the environment in discrete time steps, collecting experiences that are stored in the PER buffer for future learning. This iterative process allows the agent to gradually improve its policy over time, learning from both successful and unsuccessful experiences.

## Prioritised Experience Replay:

Prioritized Experience Replay (PER) is a key innovation in the proposed system, designed to enhance learning eﬃciency and policy robustness. PER works by prioritizing the sampling of experiences with high temporal-difference (TD) error, which are typically the most informative for learning. The PER buffer is implemented using a sum-tree structure, allowing for eﬃcient sampling and updating of experience priorities. When a new experience is added to the buffer, it is assigned an initial priority based on its TD error, and this priority is updated as the agent learns. During training, the agent samples a batch of experiences from the buffer, with the probability of selection proportional to their priority. This ensures that the agent focuses on critical transitions that drive policy improvement, leading to faster convergence and more stable learning. The PER buffer also incorporates importance sampling weights to correct for the bias introduced by prioritized sampling, ensuring that the learning process remains unbiased. By focusing the agent’s attention on the most informative experiences, PER signiﬁcantly improves the eﬃciency and effectiveness of the learning process.

## Multi-Objective Reward Design:

The multi-objective reward function is a central feature of the proposed system, designed to balance safety, eﬃciency, comfort, and fuel economy. The reward function combines several components, each weighted according to its importance. The safety penalty component penalizes violations of adaptive time-to-collision and headway thresholds, encouraging the agent to prioritize safety over other objectives. The speed reward component encourages the agent to

maintain an eﬃcient target speed, while the lane change reward component rewards safe and beneﬁcial lane changes. The comfort reward component penalizes excessive acceleration, jerk, and frequent lane changes, ensuring a smooth and comfortable ride for passengers. The fuel eﬃciency reward component rewards eco-friendly speed and acceleration proﬁles, contributing to sustainable driving practices. The combined reward is clipped to prevent reward explosion and ensure stable learning. The weights of each component can be adjusted to reﬂect different priorities or operational scenarios, providing ﬂexibility for real-world deployment. This multi-objective approach ensures that the agent learns to make decisions that are not only safe but also eﬃcient, comfortable, and environmentally friendly.

## Training Procedure:

The training procedure is designed to maximize the effectiveness of the proposed multi-objective DRL framework in learning robust and safe driving policies. Training is conducted in the highway-v0 environment from highway-env, which simulates dense traﬃc with up to 40 vehicles and 4 lanes. The agent interacts with the environment in discrete time steps, collecting experiences that are stored in the PER buffer for future learning. Each training episode consists of a ﬁxed number of steps, during which the agent must navigate the highway while balancing safety, eﬃciency, comfort, and fuel economy. Periodic evaluation is conducted every 10 episodes to monitor the agent’s performance and ensure that learning is progressing as expected. Checkpointing is implemented every 50 episodes to save model weights, replay buffer, and training metadata, enabling recoverability and facilitating further research. The agent’s policy is updated iteratively based on the experiences sampled from the PER buffer, with the learning rate and other hyperparameters carefully tuned to ensure stable and eﬃcient learning. This rigorous training procedure ensures that the agent learns robust and safe driving policies that can generalize to a wide range of highway scenarios.

# ​Conclusion:

This project demonstrates the effectiveness of a multi-objective deep reinforcement learning framework for autonomous highway driving. By integrating an adaptive safety module and prioritized experience replay, the agent achieves robust, safe, and eﬃcient driving policies in complex highway scenarios. The approach outperforms traditional single-objective and ﬁxed-threshold methods in both safety and eﬃciency metrics, as evidenced by reduced collision rates and improved driving comfort. The proposed framework addresses the limitations of existing DRL approaches by explicitly balancing multiple objectives and adapting to dynamic traﬃc conditions. The experimental results provide strong evidence that the integration of adaptive safety and prioritized experience replay can signiﬁcantly improve the performance and reliability of autonomous driving systems. Future work may explore real-world deployment, transfer learning, and integration with additional sensor modalities. By continuing to reﬁne and expand this framework, researchers and practitioners can contribute to the development of safer, more eﬃcient, and more comfortable autonomous vehicles for real-world highways.

# ​References:

1. **A. Talebpour and H. S. Mahmassani, “Inﬂuence of connected and autonomous vehicles on traﬃc ﬂow stability and throughput,” *Transportation Research Part C: Emerging Technologies*, vol. 71, pp. 143–163, 2016.**
2. **V. Mnih et al., “Human-level control through deep reinforcement learning,” *Nature*, vol. 518, pp. 529–533, 2015.**
3. **Y. Liang, T. Wang, and Y. Wang, “A deep reinforcement learning network for lane change decision-making in autonomous driving,” *IEEE Access*, vol. 7, pp. 151866–151876, 2019.**
4. **H. Ye, Y. Liu, and M. Tomizuka, “Deep reinforcement learning for autonomous driving,”**

***arXiv preprint arXiv:1709.02838*, 2017.**

1. **F. Islam, K. Kataoka, and T. Yoshida, “Hierarchical reinforcement learning for safe and eﬃcient lane change,” in *IEEE Intelligent Vehicles Symposium (IV)*, 2020, pp. 581–586.**
2. **S. Wang, J. Zhang, and X. Wang, “Lane change decision-making through deep reinforcement learning with rule-based constraints,” in *IEEE International Conference on Intelligent Transportation Systems (ITSC)*, 2018.**
3. **Y. Wang, Y. Zeng, and Q. Zhu, “Safe deep reinforcement learning for autonomous vehicles using robust control barrier functions,” in *American Control Conference (ACC)*, 2021.**
4. **Y. Chen, W. Zhan, and M. Tomizuka, “Deep reinforcement learning for safe and comfortable lane change maneuver,” in *IEEE Intelligent Vehicles Symposium (IV)*, 2019.**
5. **Y. F. Chen and M. C. Lin, “Safe reinforcement learning via curriculum-based adaptive reward shaping,” in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2020**.

# ​Appendix:

## GitHub Link:

* + - GitHub Repository URL: https://github.com/jeevanm27/rl-agent-av

## Read Me:

Project Title: Multi-Objective Deep RL for Safe Autonomous Highway Driving

* + - Setup Instructions:
      * Install Python 3.x, TensorFlow, gymnasium, highway-env
      * Clone the repository and run python main.py
    - Customization:
      * The reward function and safety module can be customized to reﬂect different priorities or operational scenarios.
    - Documentation:
      * Detailed documentation is provided in the repository, including explanations of the code structure, training procedure, and evaluation metrics.
    - Examples:
      * Example conﬁguration ﬁles and scripts are provided to facilitate quick setup and experimentation.
    - Support:
      * For any issues or questions, please refer to the issue tracker or contact the project maintainers.
    - Contributions:
      * Contributions are welcome and encouraged to further improve the project**.**

## Internship Outcome:

During our internship, we collaborated to develop a robust, multi-objective reinforcement learning agent designed speciﬁcally for autonomous highway driving. Together, we integrated advanced techniques such as an adaptive safety module and prioritized experience replay, which signiﬁcantly enhanced the agent’s safety performance and learning eﬃciency. As a result of our combined efforts, the agent demonstrated notable improvements in collision avoidance, driving comfort, and overall eﬃciency within the simulated highway environment.

Throughout the project, we gained valuable hands-on experience with state-of-the-art reinforcement learning methods, as well as practical skills in autonomous vehicle simulation and software development. Working closely as a team, we regularly discussed ideas, shared insights, and supported each other in overcoming technical challenges. Our collaboration with mentors and peers was instrumental in reﬁning the system and ensuring its robustness.

To support future development and knowledge sharing, we created comprehensive documentation that details the project’s architecture, implementation, and experimental results. This documentation not only serves as a reference for our own work but also provides a foundation for others interested in advancing autonomous driving technologies.

Looking ahead, we believe that our internship experience has laid a strong foundation for future research and real-world deployment of

autonomous driving systems. The skills and knowledge we acquired during this project have prepared us to tackle new challenges and contribute to the ongoing evolution of intelligent transportation technologies.

**End Of Report**