## A logo of a wolf Description automatically generated with low confidenceCSE -5820-MACHINE LEARNING

## Project Report: Amazon AWS Deep racer

**Objective:**

The primary objective of AWS Deep Racer is to provide a platform for developing and testing reinforcement learning (RL) algorithms for autonomous driving applications. It is designed to simulate real-world autonomous driving scenarios and allow users to train their RL models in a virtual environment before deploying them in the real world.

**Motivation:**

AWS Deep Racer serves as a platform for training and testing RL algorithms in a competitive setting, such as the AWS Deep Racer League, where developers can compete against each other by training their models to race autonomously around a virtual track. The objective of the competition is to develop the fastest and most efficient racing algorithm using RL techniques.

**I have used two different reward functions on two different models. In the first model I have used the Empire City Training – Counterclockwise track and in the second model I have used Expedition Loop – Counterclockwise track. In the following I am providing details of both of my models.**

**Methods and Parameters(model 1):**

Briefly I am providing the details of my environment simulation.

Name: Group5-finalmodel-v2

Race type: Time trial

Track: Empire City Training – Counterclockwise

Sensor: Camera

Action space: Continuous

Speed: [ 1.5 : 2 ] m/s

Steering angle:  [ -30 : 30 ] °

Reinforcement learning algorithm: PPO

The goal of my reward function is to incentivize an agent to learn a policy that achieves a particular task. In the case of a racing game, the goal is to navigate the track as quickly and efficiently as possible. Therefore, the reward function should assign higher values to actions that bring the agent closer to this goal and lower values to actions that hinder progress.

**Training and Evaluation:**

 The average completion percentage during the training is around 53 percent which went up to 100 per cent during the evaluation period.

A picture containing text, diagram, line, plot

Description automatically generated**A screenshot of a computer screen

Description automatically generated with low confidence**

The track completion during the evaluation has received 100% track completion with four laps per evaluation. The lowest track completion time I have got from this model was 12 seconds without any offtrack records. From my understanding, we can say that the reward function has some tough penalties that make the agent focus on completion more than the time.I have used an “All wheels on Track” checkpoint, which is very harsh for the agent. The track has curves where the agent could go faster if we allowed it to stay on track with two wheels. We also used a “Distance from the centre” checkpoint, making the agent stay near the centerline. These two checks were giving the agent more rewards than the speed threshold check. So, the agent decided to stay on track while compromising the speed threshold points. I have also included reward for maintaining direction , maintain steering direction and maintain throttle while streeing.

**Methods and Parameters(model-1I):**

Briefly I am providing the details of my environment simulation.

Name: Group5-finalmodel-v0

Race type: Time trial

Track: Expedition Loop– Counterclockwise

Sensor: Camera

Action space: Continuous

Speed: [ 1.1 : 2 ] m/s

Steering angle:  [ -30 : 30 ] °

Reinforcement learning algorithm: SAC

The goal of this reward function is to calculate a reward value based on several factors, including the car's position on the track, its speed, and whether all wheels are on the track.

**Description -**

This is a reward function for a DeepRacer model that takes in the params dictionary, which contains information about the current state of the car on the track. The function calculates a reward value based on several factors, including the car's position on the track, its speed, and whether or not all wheels are on the track.

**Training and Evaluation:**

 The average completion percentage during the training is below 25 percent, which didn’t perform well even during evaluation period.

A picture containing text, screenshot, plot, diagram

Description automatically generated**A picture containing text, screenshot, font, number

Description automatically generated**

The track completion during the evaluation has received 100% track completion with 3 laps per evaluation. The lowest track completion time I have got from this model was 58 seconds numerous numbers of offtrack records(46).

**Conclusion:**

From my models and analysis, I could understand that carefully constructing a reward function will be helpful in developing an agent that can go faster and more efficiently. In my models I have not played with hyperparameters due to exceeding free trial limit , but I think hyperparameters would play a major factor in good model creation. Overall, this has been a good learning curve in Reinforcement Learning and had fun while creating these models and trying to build a model which will win the competition.