A close-up of a logo

Description automatically generated

**COMP4082 – Autonomous Robotics System**

**Coursework**

**Task 1**

**Authors**

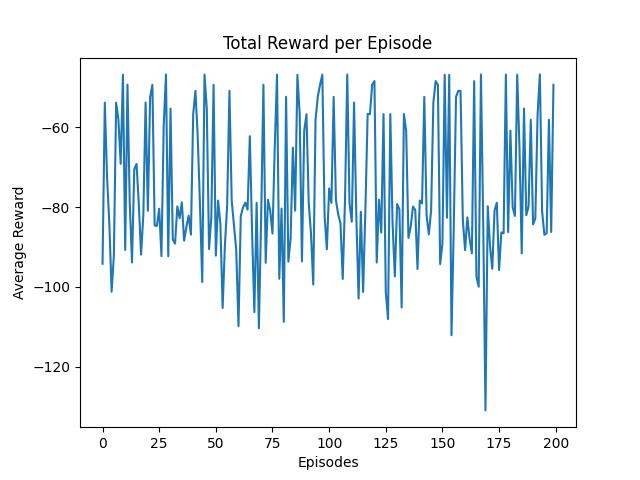
Samien Shaheed **(20291244)**

# **OpenAI Gymnasium Classic Controls**

## Mountain Car

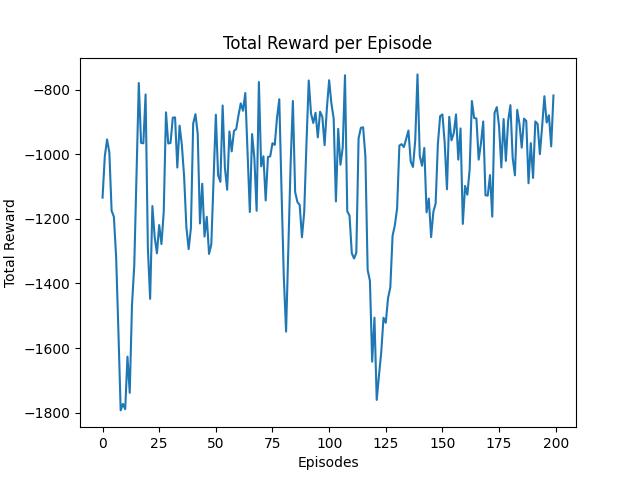
Experimentations on the “MountainCar-v0” environment were conducted during initial trial sessions. The environment has a discrete action space, where the primary strategy included pushing the car right if the velocity was positive and vice versa.

The agent was rewarded -1 for each timestep, not at the goal, but an additional energy stored value was added to the reward function for experimentation purposes. The function calculated the total energy in the system by adding potential and kinetic energy.



## Pendulum

Similarly, an experiment was carried out on the “Pendulum-v1” environment, which contained a continuous action space. For a continuous action space, the action space needs to be discretized for the Q-learning algorithm. This concept had then been the basis for Task 2.



# **Research Gaps & Questions**

Several research papers have investigated an effective technique to implement environments with continuous action spaces on Q-learning algorithms. One such included proposing a novel Q-learning with a tabular discretization to help an agent learn for the Mountain Car environment.[1]

Similarly, there have been experiments on Adaptive Discretization Methods for Reinforcement Learning to be applied to large and potentially continuous action spaces.[2]

Furthermore, there has been a study of how the discretization methods affect the learning of Q-learning for continuous action spaces.[3] This study included using different discretization methods using Project Malmo, an AI experimentation platform built on Minecraft. They had witnessed that a larger discrete action space does not necessarily lead to better results. However, this finding was not further explored.

My experiment includes studying the behaviour of Q-learning algorithms using different granularities of discretization on continuous action space. This experiment involves observing the learning process, trade-offs between exploration and exploitation, and stability.

I’ve chosen the above papers as they lay a good foundation for contemporary research on this concept. It explains the limitations of Q-learning when applied to continuous action spaces and looks into the effectiveness of the discretization technique. These works give me a good framework to continue experimentation on the behaviour of different granularities for discretization.

**Papers**

[1] S. Teja Chavali, C. Tej Kandavalli, T. M. Sugash, and J. Amudha, “Modelling a Reinforcement Learning Agent For Mountain Car Problem Using Q - Learning With Tabular Discretization,” *MysuruCon 2022 - 2022 IEEE 2nd Mysore Sub Sect. Int. Conf.*, pp. 1–5, 2022, doi: 10.1109/MysuruCon55714.2022.9972352.

[2] S. R. Sinclair, T. Wang, G. Jain, S. Banerjee, and C. L. Yu, “Adaptive discretization for model-based reinforcement learning,” *Adv. Neural Inf. Process. Syst.*, vol. 2020-December, pp. 1–50, 2020.

[3] C. Science, “Comparing Discretization Methods for Applying Q-learning in Continuous State-Action Space,” 2017.