

General Sir John Kotelawala Defense University

Department of Electrical, Electronics & Telecommunication Engineering

Machine Learning

ET 4103

Assignment – 04

Reinforcement Learning

Index No : D/ENG/22/0120/ET

Name : M. A. E. Wijesuriya

Intake : 39

Submission Date : 20/06/2025

**Q1. Utilize the given Jupyter notebook[1] for Reinforcement Learning. Comment on the code and the output of the program, explaining utilized Machine Learning concepts where necessary**

The following code is a python program that demonstrates Reinforcement Learning. It explains the basic concepts utilized in Reinforcement Learning, and provides code snippets that can be run in the Jupyter notebook provided to witness these concepts put into practice in real time.

The basic principle of Reinforcement Learning is explained: An agent interacts with the environment, which responds with a reward or cost, which enables the agent to learn behavior that leads to greater rewards and lower costs. The problem is characterized as a Markov Decision Process, which has:

* **S** ­– A finite set of states that the agent can inhabit
* **A** – A finite set of actions that the agent can take in each state
* ***R*:S×A→[*Rmin*,*Rmax*]⊂R –** The bounded reward or cost function that gives the agent reinforcement
* ***P*:S×A→Δ(S) –** Transition probabilities of the agent moving to the next state
* ***γ –*** A discount factor. The closer to one it is, the less it encourages the agent to reach rewards as quickly as possible.

The **Policy** of an agent is the strategy or set of rules it utilizes in order to decide what actions it should take at each state in the environment. The optimal policy gives us the value function ***V*∗(*s*)=max*a*∈A[*R*(*s*,*a*)+*γ*∑*s*′∈S*P*(*s*,*a*)(*s*′)*V*∗(*s*′)]**