Sri Lanka Institute of Information Technology

**Deep Learning – SE4050**

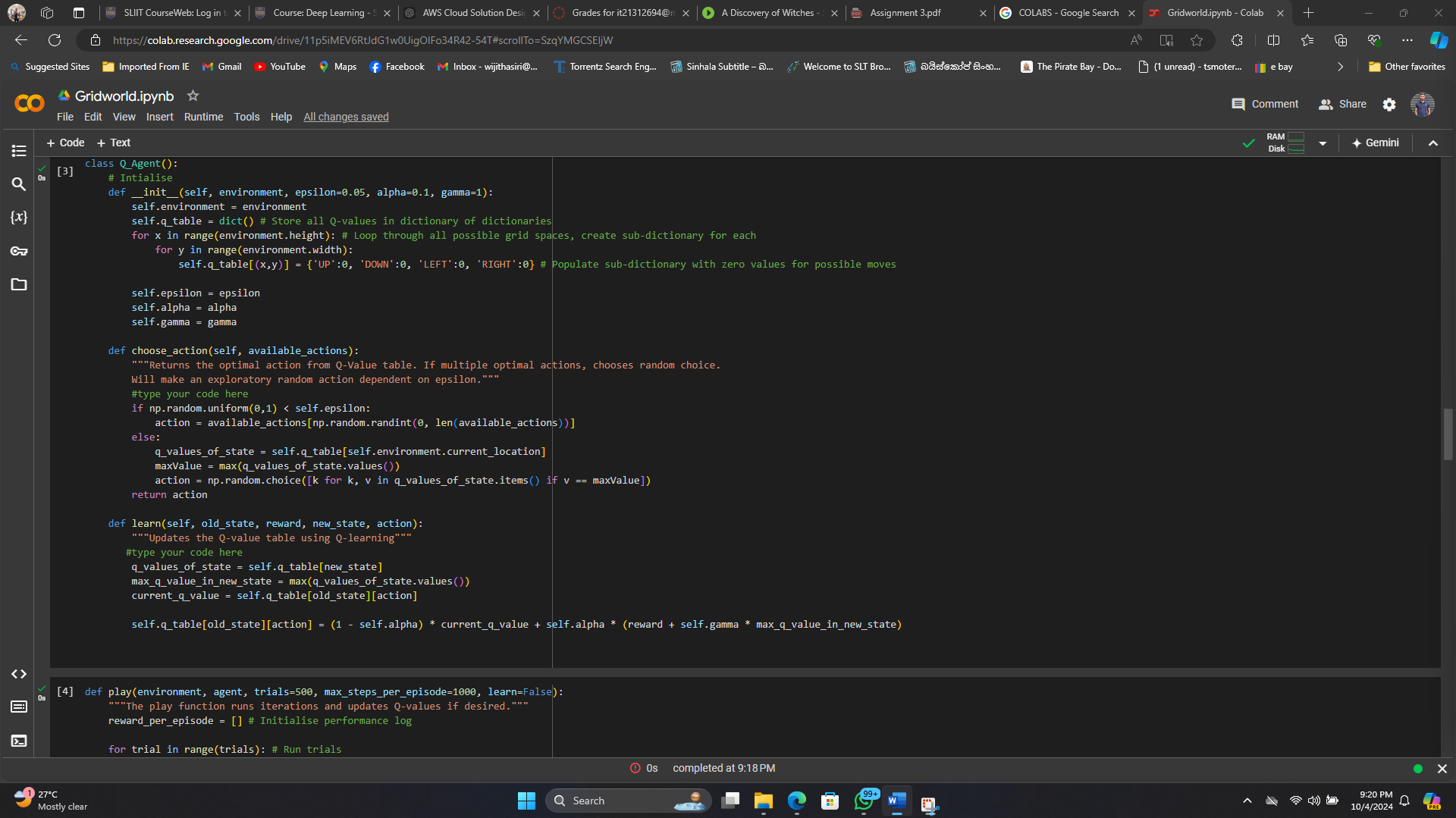
Lab 08 – 2024, Year 4 Semester 1

A picture containing text, clipart, vector graphics

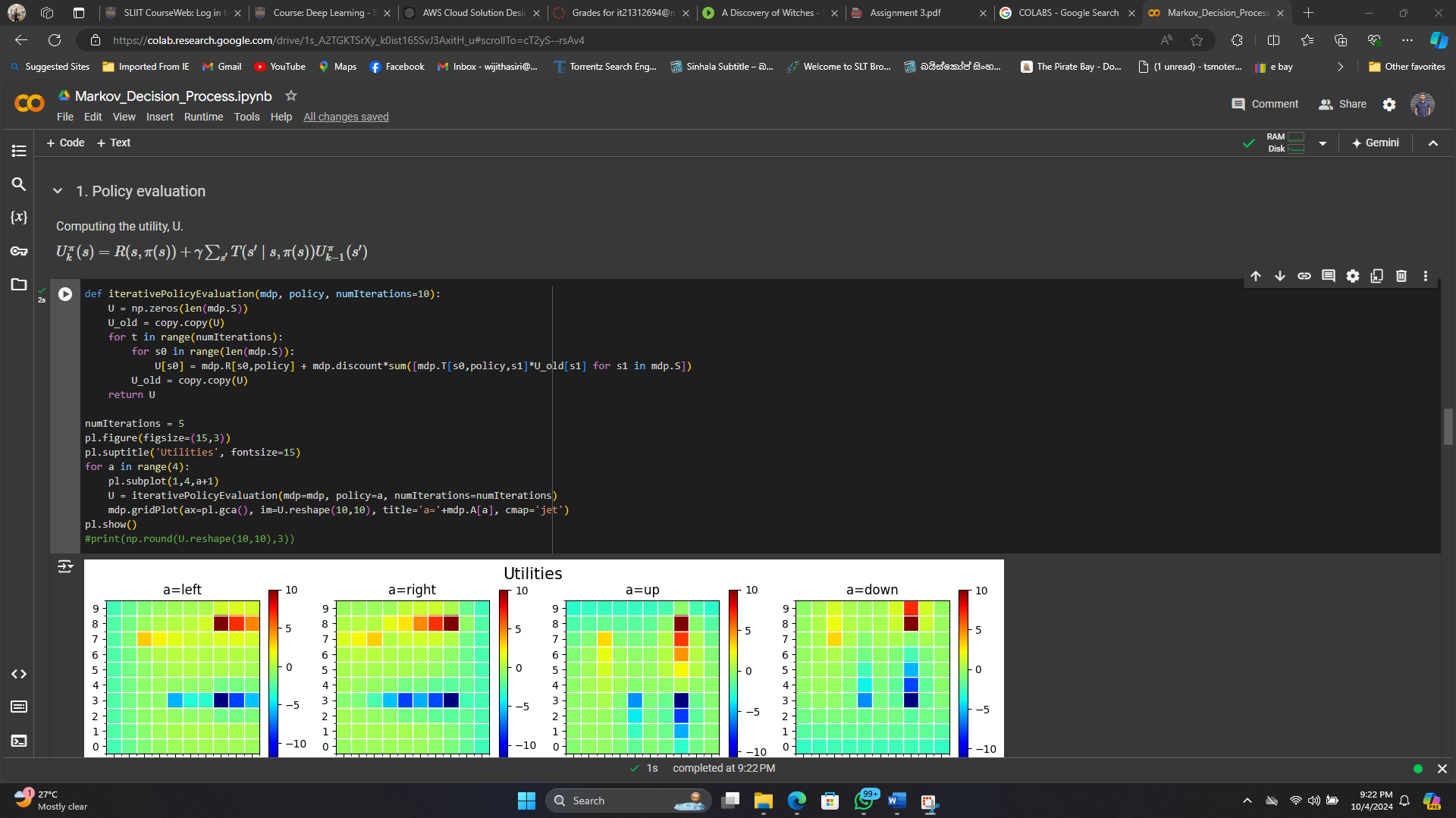
Description automatically generated

IT21312694-P.K.T.SANDEEPTHA

Task 01

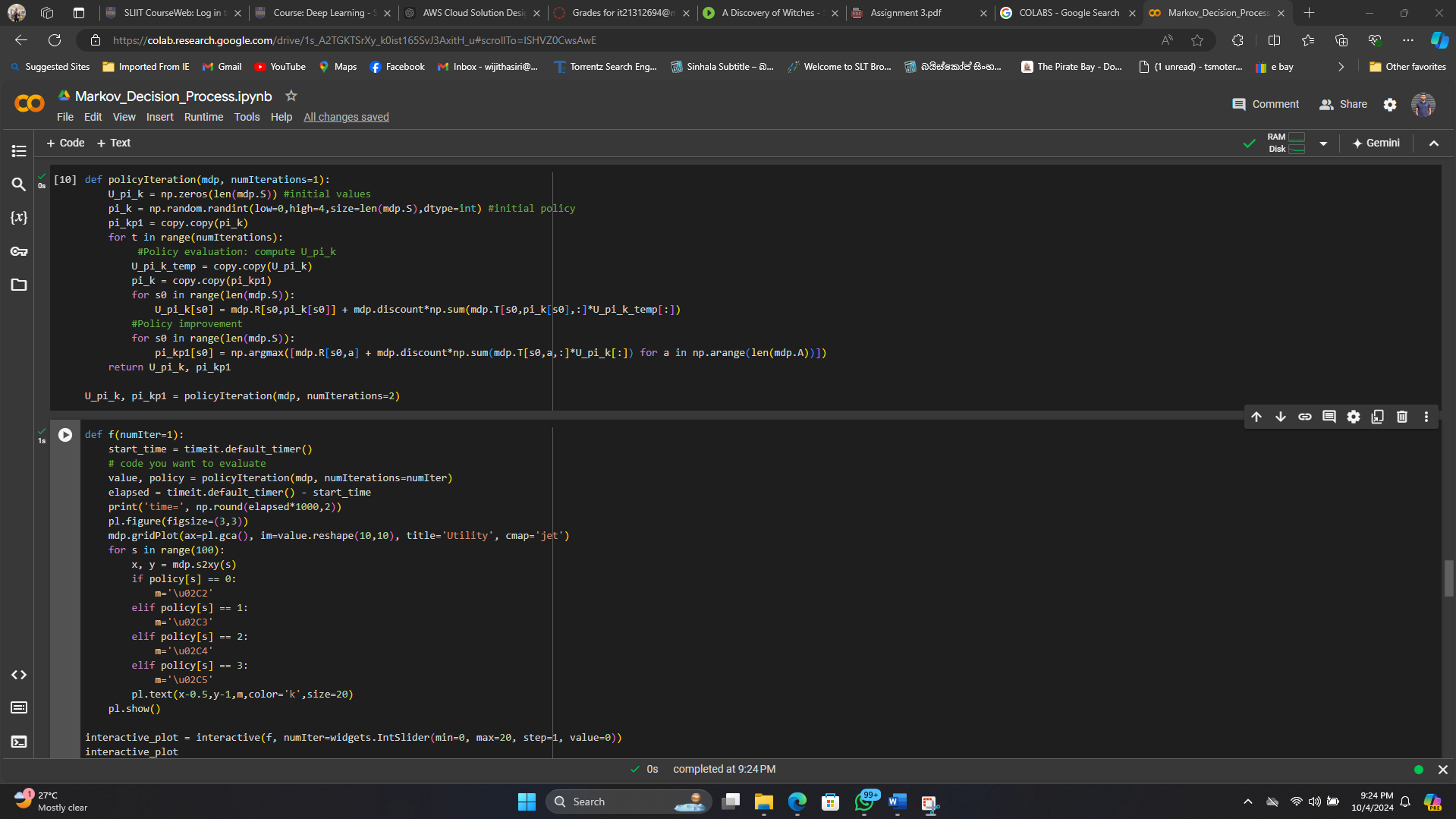
Gridworld

Markov\_Decision\_Process



A screenshot of a computer

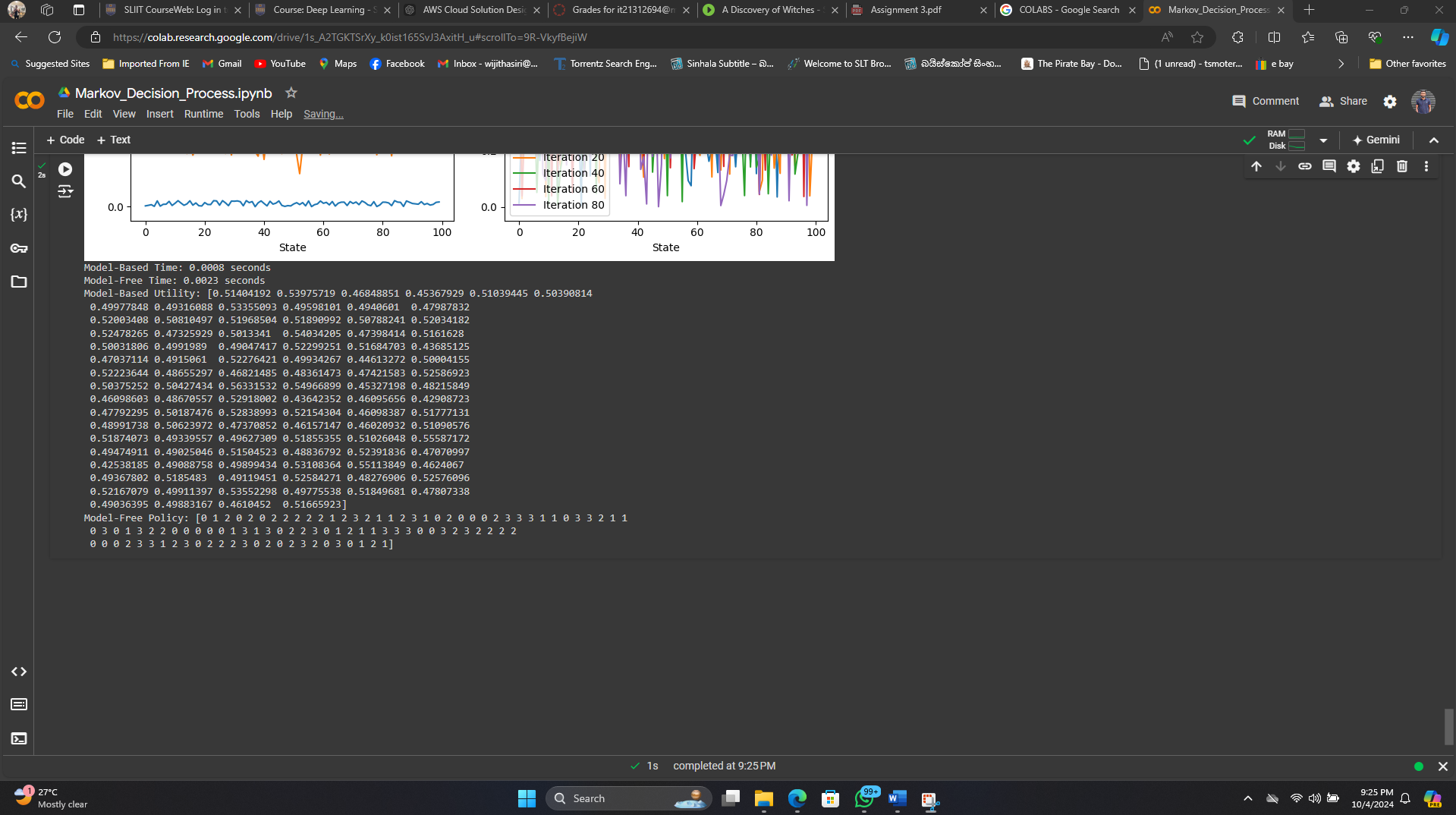
Description automatically generated



Task 02

**Model-Based algorithms** rely on having or building a model of the environment, which allows the agent to simulate future outcomes and plan effectively.

**Model-Free algorithms** learn directly from interactions with the environment, without needing an explicit model of the environment's dynamics, but typically require more exploration and data to converge to optimal policies.



A screenshot of a computer

Description automatically generated

Task 03

**Observations:**

* In this task, we extended the Markov Decision Process (MDP) with Q-Learning to develop a Deep Q-Learning (DQN) model.
* The Q-values are approximated using a neural network instead of a traditional lookup table.
* We implemented an epsilon-greedy strategy to balance exploration and exploitation, and tested different epsilon values (0.1, 0.5, 0.9) to analyze their impact on the performance of the agent.
* The epsilon-greedy strategy plays a critical role in ensuring that the agent explores the environment sufficiently while still exploiting the knowledge it has gained so far.
* As we decrease the epsilon value, the agent becomes more exploitative, while larger epsilon values promote more exploration.
* The results show that moderate epsilon values (e.g., 0.5) tend to strike a balance between exploration and exploitation, leading to more consistent learning and faster convergence.

**The following plot illustrates the performance of DQN with different epsilon values over 50 episodes:**

