Instruction: Write the corresponding solutions to the following questions, document every important line of your statements for both classes of questions, and Turn in the git clone address in the Google Classroom Platform.

Goal: You are given two exercises. The goal of the two exercises is to test the understanding and implementation of reinforcement learning algorithms in the context of grid world environments.

It is an individual assignment.
Code will be checked for plagiarism.
Due date is Jan 18, 2025

Question 1:

Implement the value iteration algorithm to find the optimal policy for a fully observable environment represented as an MDP. Consider a grid world environment where the agent can move in four directions: up, down, left, and right. The grid world is represented as a 5x5 grid, and each cell has a reward value associated with it. The agent receives a reward of +10 for reaching the goal state and a reward of -1 for every other step. The discount factor (gamma) is set to 0.9. Write the code to perform value iteration and print the optimal policy for each state after convergence.

Question 2:

Implement a basic Q-learning algorithm for a simple grid world environment. The grid world consists of a 5x5 grid where an agent can move in four directions: up, down, left, and right. The agent receives a reward of +10 for reaching the goal state and a reward of -1 for every other step. The agent's Q-table should be initialized with zeros, and the learning rate (alpha) and discount factor (gamma) should be set as 0.5 and 0.9, respectively. Write the code to perform Q-learning for a fixed number of episodes, and print the learned Q-values after training.