1.   Discuss the algorithm you implemented:

DQN is an extension of Q learning, which is a typical deep reinforcement learning method. In DQN, a Q function expresses all action values under all states, and it is approximated using a convolutional neural network. Using the approximated Q function, an optimal policy can be derived. In DQN, a target network, which calculates a target value and is updated by the Q function at regular intervals, is introduced to stabilize the learning process. A less frequent updates of the target network would result in a more stable learning process.

2.   What is the main improvement over the vanilla DQN?

One of the problems of the DQN algorithm is that it overestimates the true rewards; the Q-values think the agent is going to obtain a higher return than what it will obtain in reality. To fix this, Double DQN algorithm uses a simple trick: decoupling the action selection from the action evaluation. First, the main neural network θ decides which one is the best next action a' among all the available next actions, and then the target neural network evaluates this action to know its Q-value.

3.   Show and discuss your results after applying the two algorithms implementation on the environment. Plots should include epsilon decay and the reward per episode.
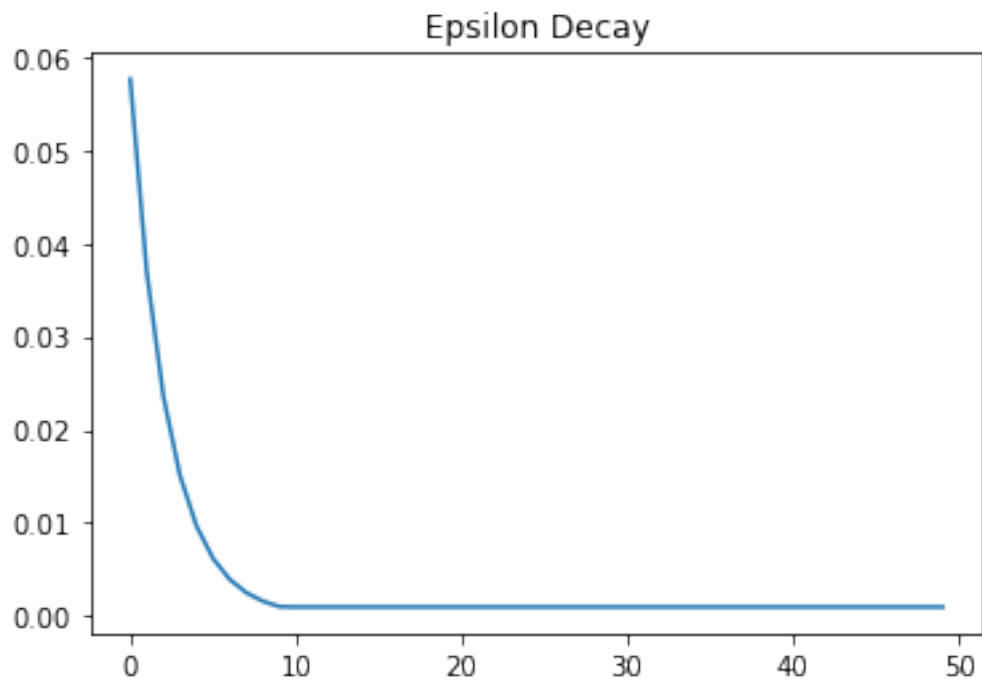
Deep Q Network (Epsilon Decay and Reward Per Episode)

Total Reward / Episode

Epsilon Decay

Double Deep Q Network (Epsilon Decay and Reward Per Episode)


Total Reward / Episode
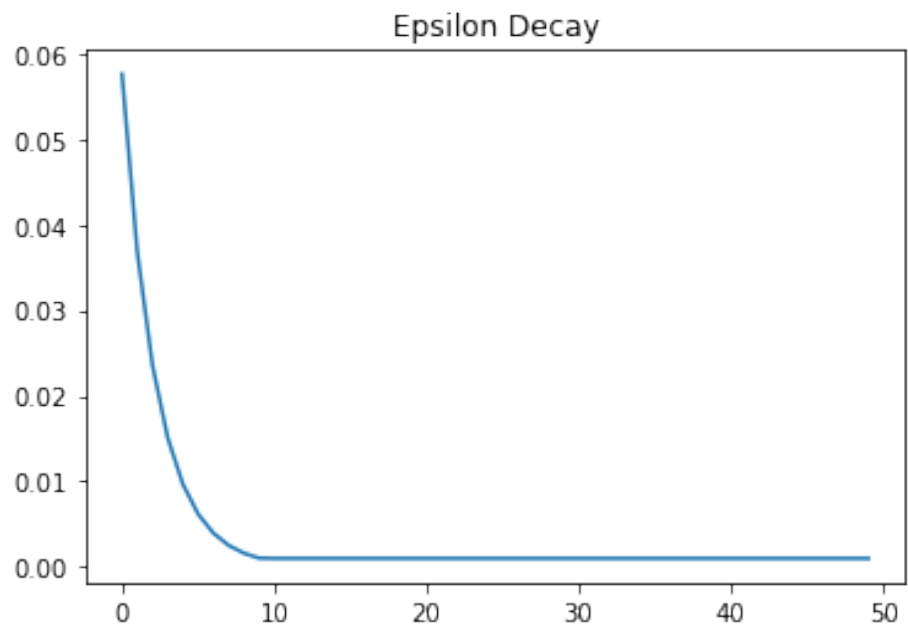
Epsilon Decay

5. Compare the performance of both algorithms (DQN & Improved version of DQN) on the same environments (e.g. show one graph with two reward dynamics) and provide your interpretation of the results.
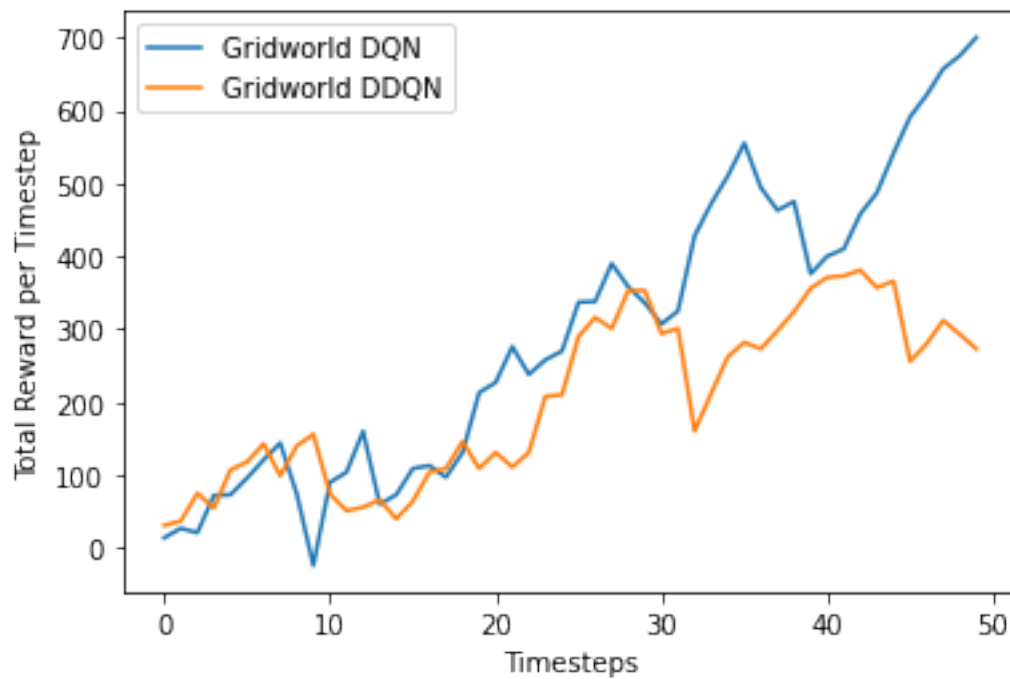
Interpretation: Both Algorithms, Deep Q-Network and Double Deep Q-Network are significantly less different in terms of performance for 50 episodes.

6. Provide your interpretation of the results. E.g. how the same algorithm behaves on different environments, or how various algorithms behave on the same environment.
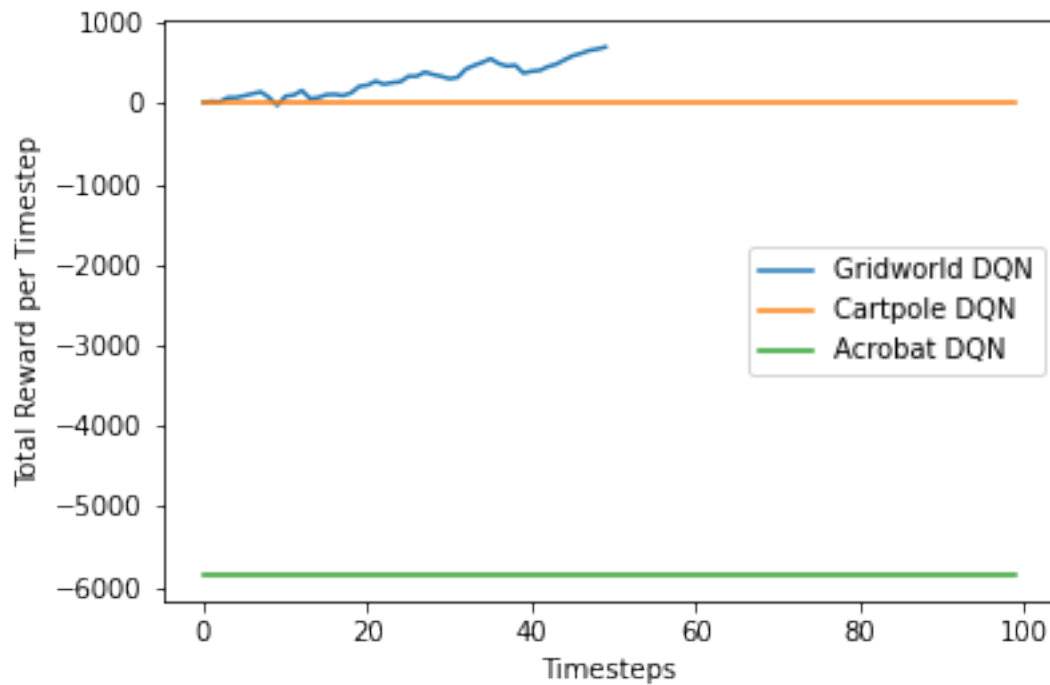
Interpretation: Deep Q-Network performed better in Gridworld compared to Cartpole and Acrobat.