COMP9444 assignment 3 report

student name: HANG ZHANG

student id: z5153042

1. ## implementation detail

   GAMMA: 0.99

   Initial epsilon: 0.9

   final epsilon: 0.01

   epsilon decay: 200

   batch size: 200

   hidden nodes size: 100

   replay buffer size: 10000

2. ## Batching and experience replay

   The basic idea is that by storing an agent's experiences, and then randomly drawing batches of them to train the network. For our DQN, we will build a simple class that handles storing and retrieving memories.

3. ## Early stopping

   In order to accelerate the process and to prevent overfitting. This is done by recording rewards in each episode and stopping further training when the last 10 episodes got reward not less than 195. We consider this result is "well-learned".

References:

https://medium.com/@awjuliani/simple-reinforcement-learning-with-tensorflow-part-4-deep-q-networks-and-beyond-8438a3e2b8df

https://medium.com/init27-labs/understanding-q-learning-the-cliff-walking-problem-80198921abbc