Project 1 – Report

The solution implements a Deep Q Learning approach (DQN) with a neural network with two hidden layers (fully connected) with 256 neurons each.
The algorithm uses a Replay Buffer which includes [state, action, next state, reward, done] from 10k experiences and for each training step 64 are randomly picked.
The training process uses an epsilon greedy strategy with a rate decay of 0.995 starting from 1.0 to 0.01 for every episode.

The training progress for 2500 episodes shows the following performance. It shows, that the final performance is already reached after ~1000 steps.