Learning rate: 0.0025, batch size: 32, target update: 100

ε-greedy, ε = 0.01