

$\epsilon$ -greedy agent policy with gamma = 0.95

