Training P-H-D (σ=0.6)

REINFORCE
DQN

training returns

episode