

PPO: Performance with tuned hyperparameters and exponential learning rate annealing

