

Policy Gradient running average rewards (1000 episode window)

