Aryan Luthra
CS285 HW1 Report
Sept 14 2020

Table 1: Behavior of similar BC agents in 2 different environments

Environment	Average Return	STD return	Expert Average Return	Expert STD Return	Special Hyperpams
Ant	1498	856	4713	13	Hidden Layer Size = 128 Eval Batch Size = 10000 Agent Train Steps Per Iter = 2000
Humanoid	312	42	10345	21	Hidden Layer Size = 128 Eval Batch Size = 10000 Agent Train Steps Per Iter = 2000

Fig 1: Average BC Return VS Agent Train Steps Per Iter
Caption: As Agent Train Steps Per Iter increases, we see that our average return increases. I chose this to illustrate that the more expert data we have, the better we can fit the experts policy.

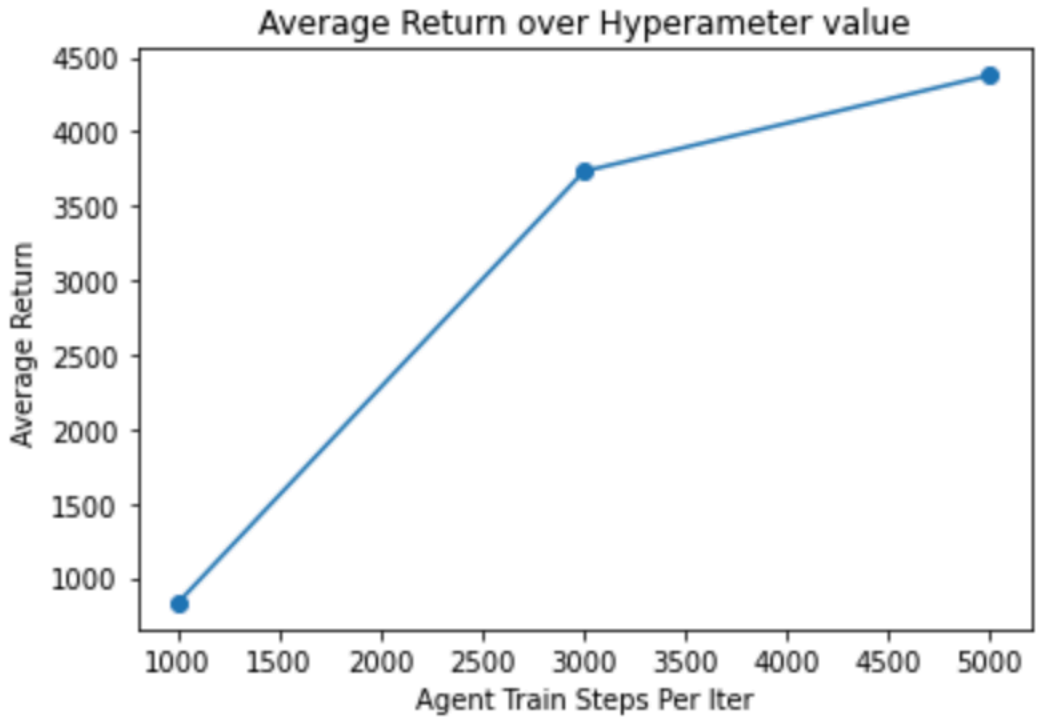

Fig 2: DAgger results over iterations with Ant & Humanoid environment.
Caption: Used same hyperparams as in Table 1, except training_time_steps was reset to default 1000. Ran for 15 DAgger iterations