Deep Reinforcement Learning Assignment 1
Imitation Learning

Master IASD - Théophane Vallaeys, Sihan Xie

January 24, 2023

1 Behavioral Cloning

1.1 Results on tasks

Task	Expert return	Average return (eval)	σ_{return}
Ant	4713	4795 **(102%)**	77
HalfCheetah	4205	4021 **(96%)**	95
Hopper	3772	1088 **(29%)**	292
Walker2d	5566	907 **(16%)**	359

Table 1: Performance of trained policies on different tasks
Training and evaluation settings : 20000 training steps for one iteration, batch size and eval batch size of 10000, episode length of 1000, and a network with 5 hidden layers of width 256, trained with lr=0.001

1.2 Experiment with hyperparameters

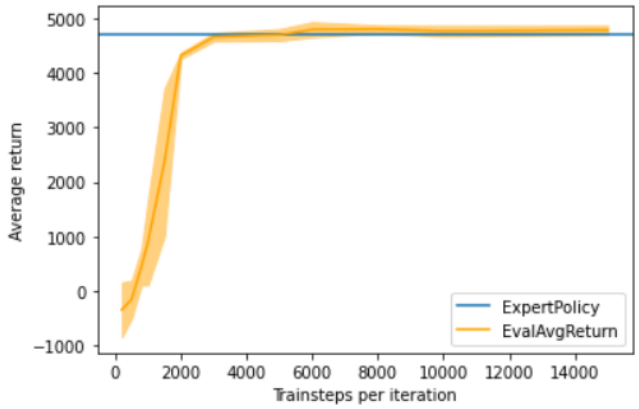

Figure 1: Performance of behavioral cloning when applying different amounts of training steps (episode length of 1000, batch size and eval batch size of 5000, and a network with 5 hidden layers of width 256, trained with lr=0.001)

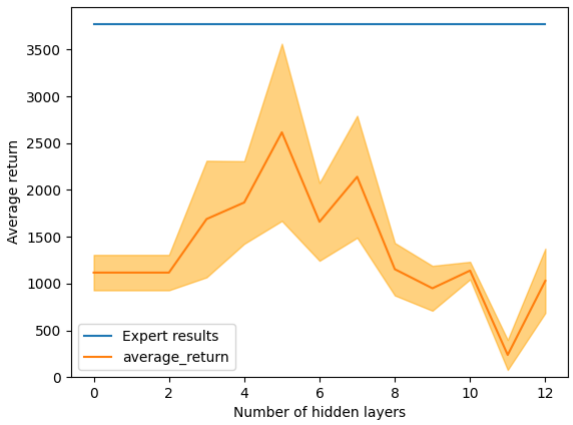

Figure 2: Performance of behavioral cloning when using different number of hidden layers (each of width 128). Trained with episode length of 1000, lr=0.01 and a batch size of 10000. The eval batch size is 50000.

2 DAgger

2.1 Results on tasks

Task	Expert return	Average return (eval)	σ_{return}
Ant	4713	4824 **(102%)**	92
HalfCheetah	4205	3998 **(95%)**	69
Hopper	3772	3720 **(99%)**	2.6
Walker2d	5566	5380 **(97%)**	44

Table 2: Performance of trained policies on different tasks
Training and evaluation settings : 2000 training steps for 10 iteration, batch size and eval batch size of 10000, episode length of 1000, and a network with 5 hidden layers of width 256, trained with lr=0.001

2.2 Curve for Ant and Walker2d tasks

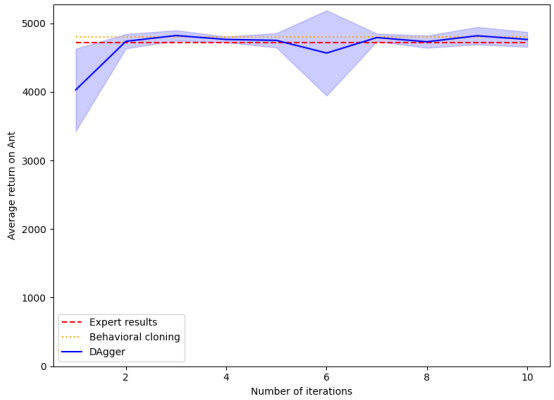

Figure 3: Performance of DAgger ont the Ant task, compared to the expert policy and Behavioral cloning

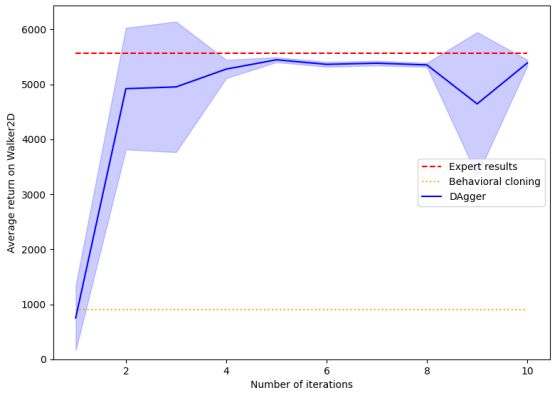

Figure 4: Performance of DAgger ont the Walker2D task, compared to the expert policy and Behavioral cloning