

Assignment 1: Imitation Learning

Andrew ID: yxiu2
Collaborators: yxiu2

1 Behavioral Cloning

1.1 Part 2

Environment	Mean	Standard Deviation
Ant-v2	4713.65	12.19
Humanoid-v2	10344.52	20.98
Walker2d-v2	5566.85	9.24
Hopper-v2	3772.67	1.95
HalfCheetah-v2	4205.78	83.04

Table 1: Mean and Standard deviation of return over five trajectories of the expert data.

1.2 Part 3

Environment	Eval_AverageReturn	Eval_StdReturn	Initial_DataCollection_AverageReturn	Performances
Ant-v2	1512.86	1049.34	4713.65	32.09%
Humanoid-v2	267.94	69.69	10344.51	2.59%

Table 2: Mean, Standard deviation and performances over trajectories of the expert data in environment Ant and Humanoid. The behavioral cloning agent in the Ant environment could achieve 32.09% of the performance of the expert, while and in the Humanoid environment, the performance only achieves 2.59%. Both evaluations are conducted on five rollouts, where the eval_batch_size is 5000 and ep_len is 1000. For a fair comparison, both training and evaluations are run with the same hyperparameters such as network size = 64, number of training iterations = 1000, and n layers = 2.

1.3 Part 4

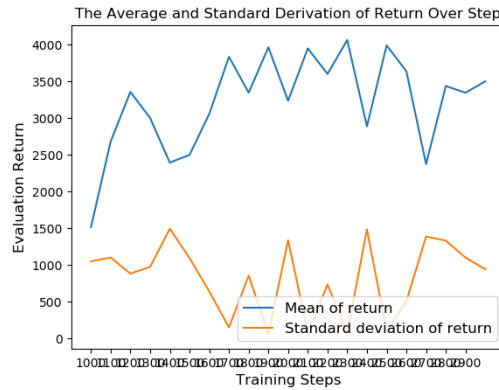


Figure 1: The Figure shows Ant-v2 BC agent's performances with different training steps from 1000 to 3000. The hyperparameters are set with same network size 64 and n layer = 2, the eval_batch_size is 5000 and ep_len is 1000. The reason of modifying training steps is, we believe the network would have better performance as the network are trained with more steps to fit the data. The empirical performance seems to increase after 1000 steps as the average return goes higher, but still with large oscillations.

2 DAgger

2.1 Part 2

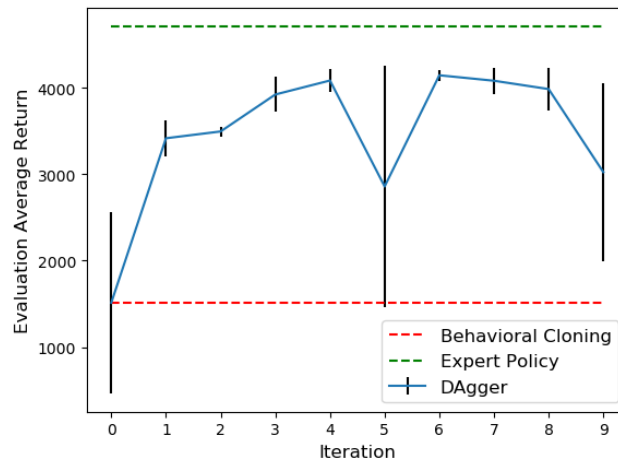


Figure 2: The Figure shows Ant-v2 agent's performance for DAgger, BC, and expert policy in Ant environment. The hyperparameters are set with same network size 64 and n layer = 2, training_steps = 1000, the eval_batch_size is 5000 and ep_len is 1000, which collects five rollouts.

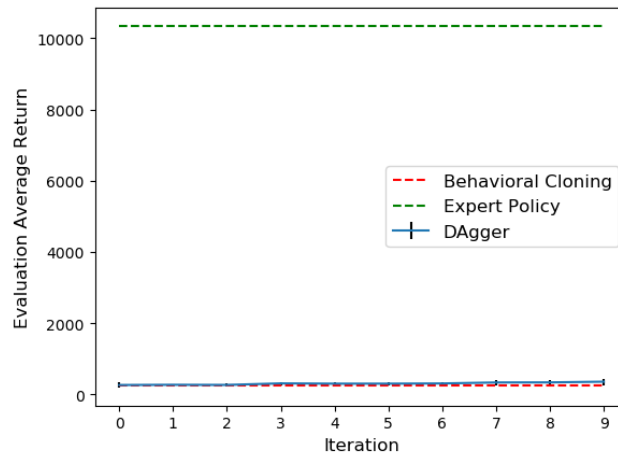


Figure 3: The Figure shows Humanoid-v2 agent's performance for DAgger, BC, and expert policy in Humanoid environment. The hyperparameters are set with same network size 64 and n layer = 2, the eval_batch_size is 5000 and ep_len is 1000, which collects five rollouts.