Assignment 1: Imitation Learning

Andrew ID: sghatge
Collaborators: None

1 Behavioral Cloning (65 pt)

1.1 Part 2 (10 pt)

Table 1: Report your result in this table.

Metric/Env	Ant-v2	Humanoid-v2	Walker2d-v2	Hopper-v2	HalfCheetah-v2
Mean	4713.65	10344.51	5566.84	3772.67	4205.77
Std.	12.19	20.98	9.23	1.94	83

1.2 Part 3 (35 pt)

Table 2: Fill your results in this table.

Env	Ant-v2			Hopper-v2		
Metric	Mean	Std.	Ep_Len	Mean	Std.	Ep_Len
Expert	4713.65	12.19	1000	3772.67	1.94	1000
BC	1462.35	1206.26	941	740	194.62	201

To achieve 30% accuracy for Ant-v2, I changed eval_batch_size to be 5000 and num_agent_train_steps_per_iter to be 2000. I compared its performance with Hopper-v2 agent because their average costs for the expert policy is in similar range. Average Episode length is 1000 for both the experts, but Hopper returns a very low value of evaluation average return because its average ep_len is around 200 whereas for the ant agent it is around 900.

1.3   Part 4 (20 pt)

I chose learning rate as the hyperparameter to be tuned to observe its impact on the performance of the behavioral cloning agent, which in this case is the Ant-v2 environment. Increasing the learning rate allowed for quicker convergence, and we can see the average values are getting more closer to the expert average values (as seen in the table above), as we increase the learning rate.
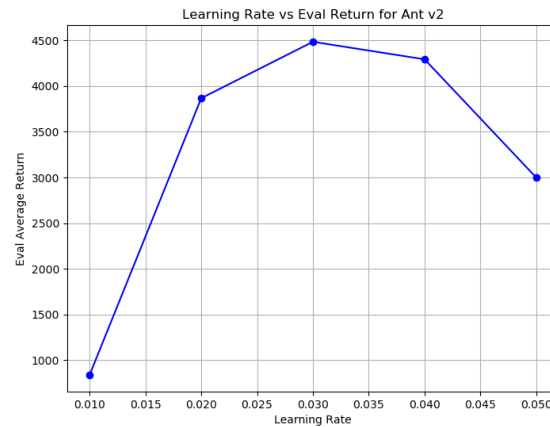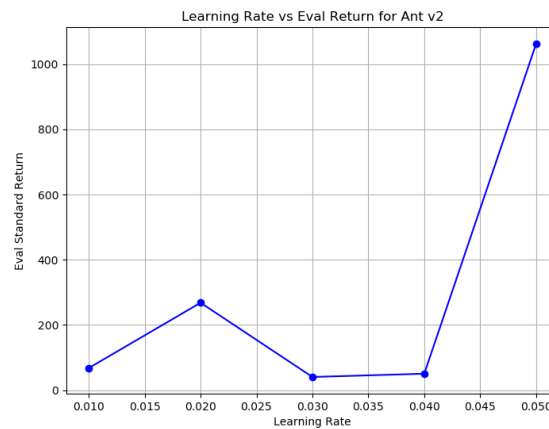

Figure 1: BC agent's performance varies with the value of changing learning rate parameter in Ant environment.

2   DAgger (35 pt)

2.1   Part 2 (35 pt)

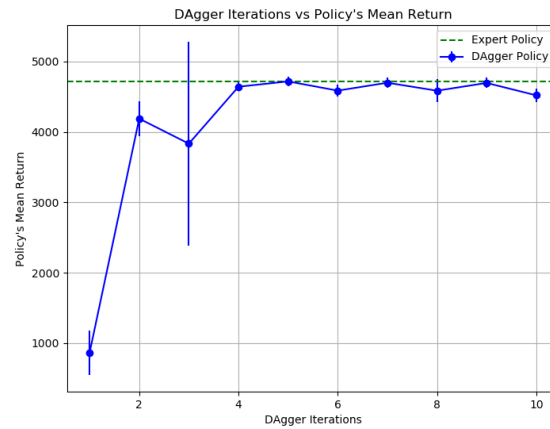BC agents performance using DAgger algorithm by using the default hyperparameters for more than 2 trajectories.
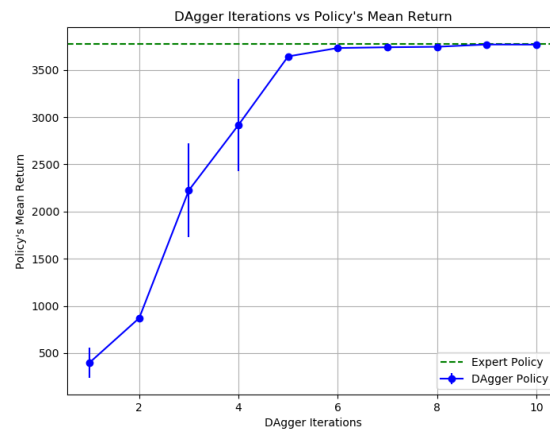

Figure 2: BC agent: Ant-v2

Figure 3: BC agent: Hopper-v2