

Assignment 1: Imitation Learning

Andrew ID: yulong1

Collaborators: N/A

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	4713.653	10344.518	5566.846	3772.670	4205.778
Std.	12.197	20.981	9.238	1.948	83.039

1.2 Part 3 (35 pt)

Table 2: Fill your results in this table.

Env	Ant-v2		Humanoid-v2	
Metric	Mean	Std.	Mean	Std.
Expert	4713.653	12.197	10344.518	20.981
BC	1729.545	767.105	281.326	15.861

Table 3: BC agent performance with 3 layers MLP of width 256, trained for 1000 steps.

1.3 Part 4 (20 pt)

The graph shows how the BC agent's performance varies with the MLP width. We can see that the performance increases as the width increases initially, which makes sense because a wider network has more parameters to learn from the data. However, after a certain point, the performance starts to decrease. This is because the network becomes too large and starts to overfit to the data.

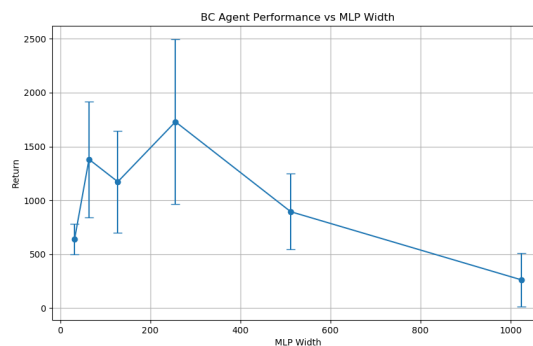


Figure 1: BC agent's performance varies with the value of MLP width in Ant-v2 environment.

2 DAgger (35 pt)

2.1 Part 2 (35 pt)

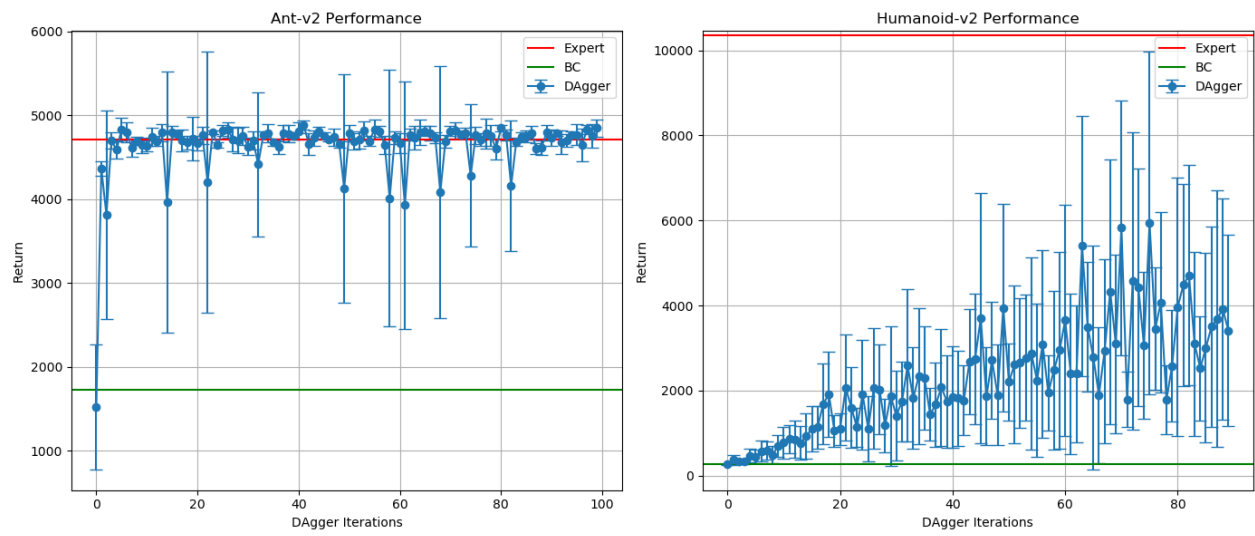


Figure 2: BC with DAgger performance with 2 layers MLP of width 64, trained for 1000 steps per iterations.