Assignment 1: Imitation Learning

Andrew ID: `ayanovic`
Collaborators: `kaustabp`

1 Behavioral Cloning

1.1 Part 2

Label	Mean Train	Std Train	Mean Eval	Std Eval	
Ant-v2	4713.65	12.2	688.3	0.0	14.6
Humanoid-v2	10344.51	20.98	280.03	25.05	2.71
Walker2d-v2	5566.85	9.24	459.27	722.48	8.25
Hopper-v2	3772.67	1.95	844.55	319.59	22.39
HalfCheetah-v2	4205.78	83.04	2926.69	159.5	69.59

Table 1: Environment comparison using default parameters over 5 trajectories.

1.2 Part 3

index	Label	Mean Train	Std Train	Mean Eval	Std Eval	
0	Ant-v2	4713.65	12.2	2195.62	1663.23	46.58
1	Humanoid-v2	10344.52	20.98	306.38	58.01	2.96

Table 2: Ant-v2 (above 30%) and Humanoid-v2 (below 30%) comparison using default parameters over 5 trajectories. Note, all parameters at default values, with the exception of the learning rate for better performance (alpha = 1e-2)

1.3 Part 4

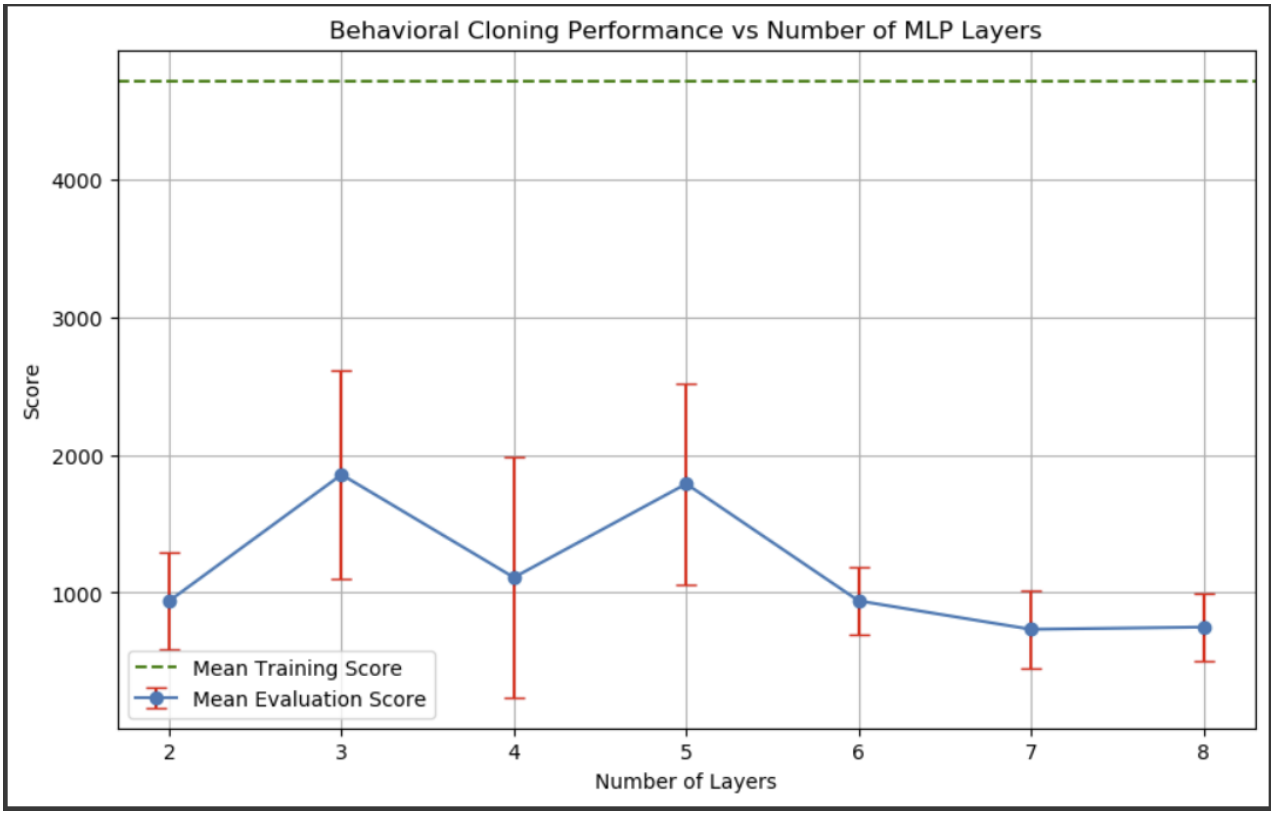

Figure 1: Hyper-parameter performance testing with variation in MLP neural layers. The rest of the parameters held at default values.

2 DAgger

2.1 Part 2

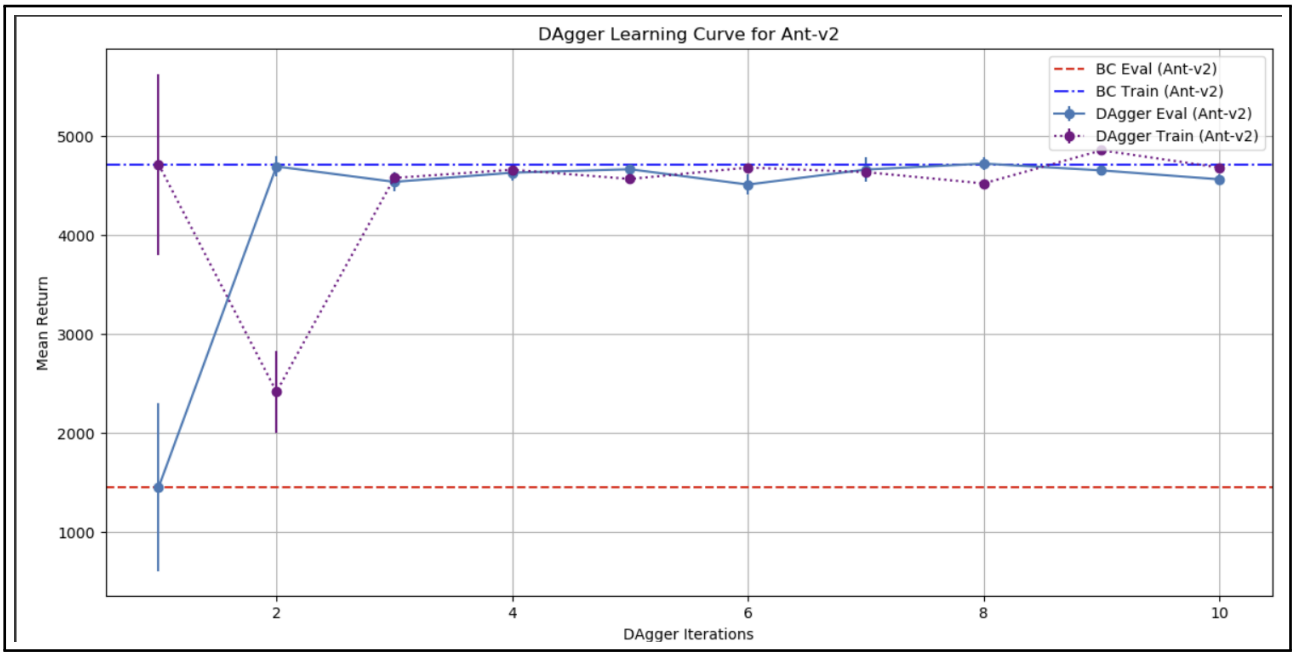

Figure 2: Implementation for Ant-v2 environment. All parameters at default values, with the exception of the learning rate for better performance (alpha = 1e-2)

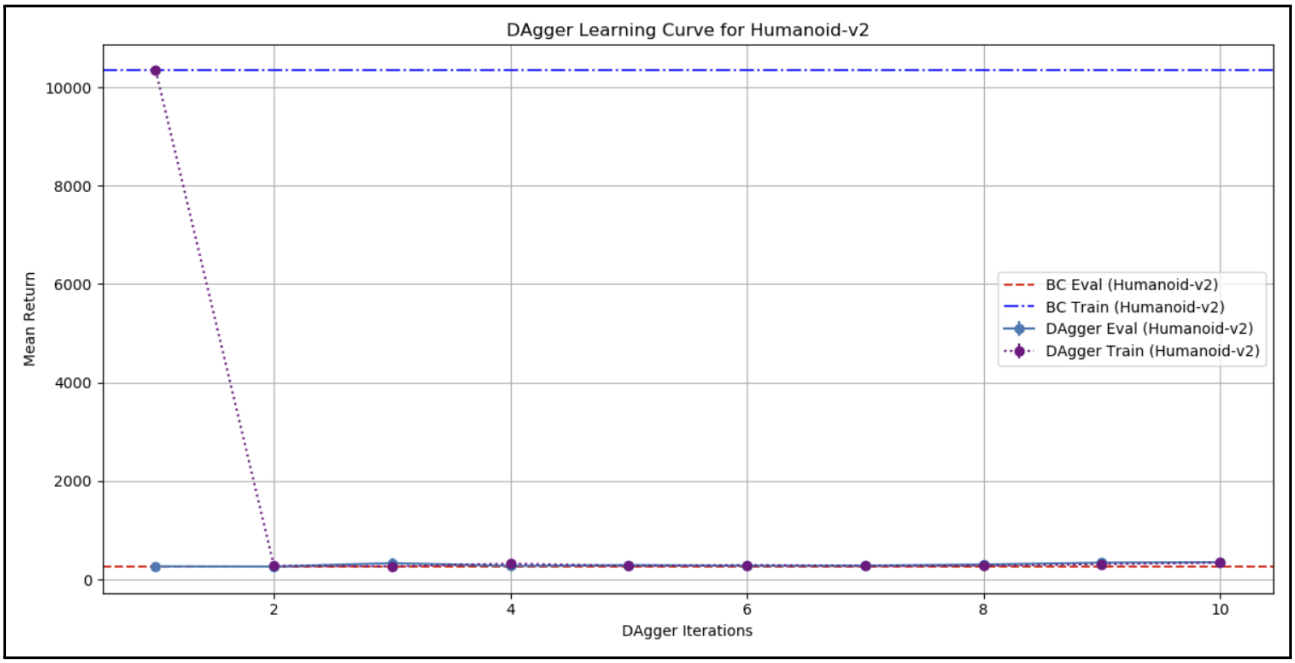

Figure 3: Implementation for Humanoid-v2 environment. All parameters at default values, with the exception of the learning rate for better performance (alpha = 1e-2)