HW1
Siyao Jia, SID: 3032531126

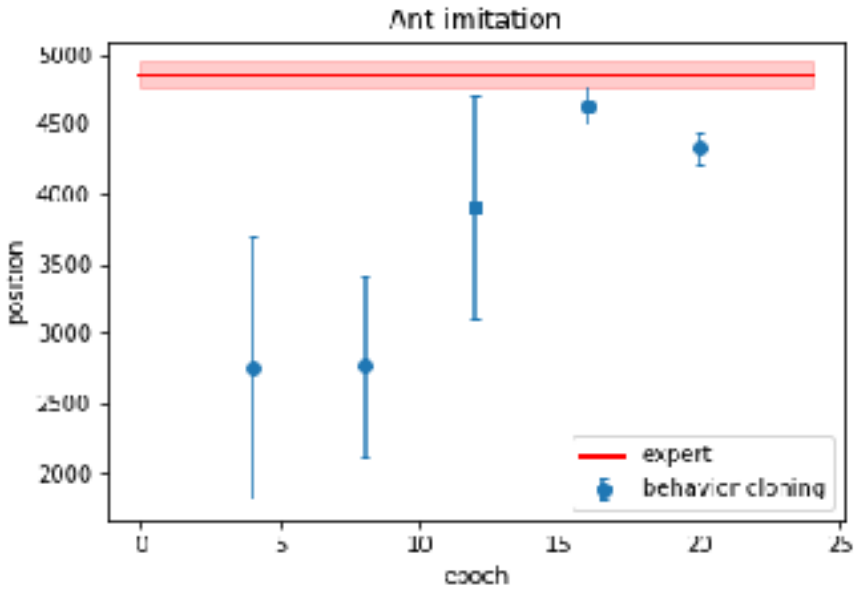
Section 2: Behavioral Cloning

Section 2.2:

task	Mean	std
Ant-expert	4812	75
Ant-behavior cloning	4303	102
Hopper-expert	3778	4
Hopper-Behavior cloning	350	182

I run the expert model 20 times to generate 20k expert training data. I use a neural network with 4 fully connected hidden layers, 128 neurons per layer, with “relu” activation. Totally I have 20k input data, in which 25% are used as validation data. I repeated training process 10 iterations. Here, Ant is an example where behavior cloning achieves comparable performance to the expert, while the Hopper is not.

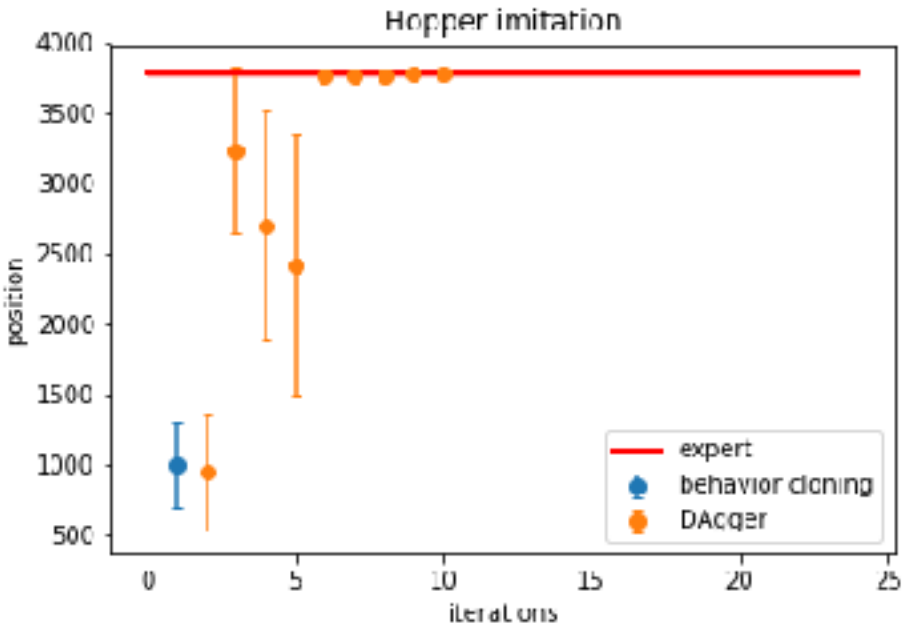
Section 2.3:



This plot shows how the how the BC agent’s performance varies with the number of epochs for Ant task. We can see the Behavior learned position becomes more accurate (approaches the expert position) and prices (smaller standard deviation) as we use more epochs. This is because the more epochs we trained the NN, the better our model is.

Section 3: DAgger

Section 3.2:



Similar as before: I run the expert model 20 times to generate 20k expert training data. I use a neural network with 4 fully connected hidden layers, 128 neurons per layer, with “relu” activation. Totally I have 20k input data, in which 25% are used as validation data. One difference is that I repeated training process 16 iterations here.

This plot shows how the how the DAgger agent’s performance varies with the number of iterations for Hopper task. We can see the DAgger learned position becomes more accurate (approaches the expert position) and prices (smaller standard deviation) as we use more iterations. After 5 iterations, the Dagger performance agrees with the expert within 1 sigma.