Assignment 1: Imitation Learning

Andrew ID: kkabeer
Collaborators: ssaha3, hravisan, asenathi

1 Behavioral Cloning

1.1 Part 2

	Ant	Humanoid	Walker2d	Hopper	HalfCheetah
Average Return of Training Data	4713.65	10344.51	5566.84	3772.67	4205.77
Standard Deviation of Training Data	12.19	20.98	9.23	1.94	83.03

1.2 Part 3

	Ant	Humanoid	Walker2d	Hopper	HalfCheetah
Average Return of Evaluation Data	1502.93	314.30	342.79	805.38	2120.04
Standard Deviation of Evaluation Data	713.67	33.77	259.76	254.06	1193.84
% Performance of Expert	31.88%	3.04%	6.16%	21.35%	50.41%

I am choosing the Ant environment, which has a performance of 31.88% of the expert, and the Hopper environment, which has a performance of 21.35% of the expert Parameters used: Network size - 2 hidden layers of size 64 each Amount of data - Training data batch size: 5000 and Evaluation data batch size: 5000

1.3 Part 4

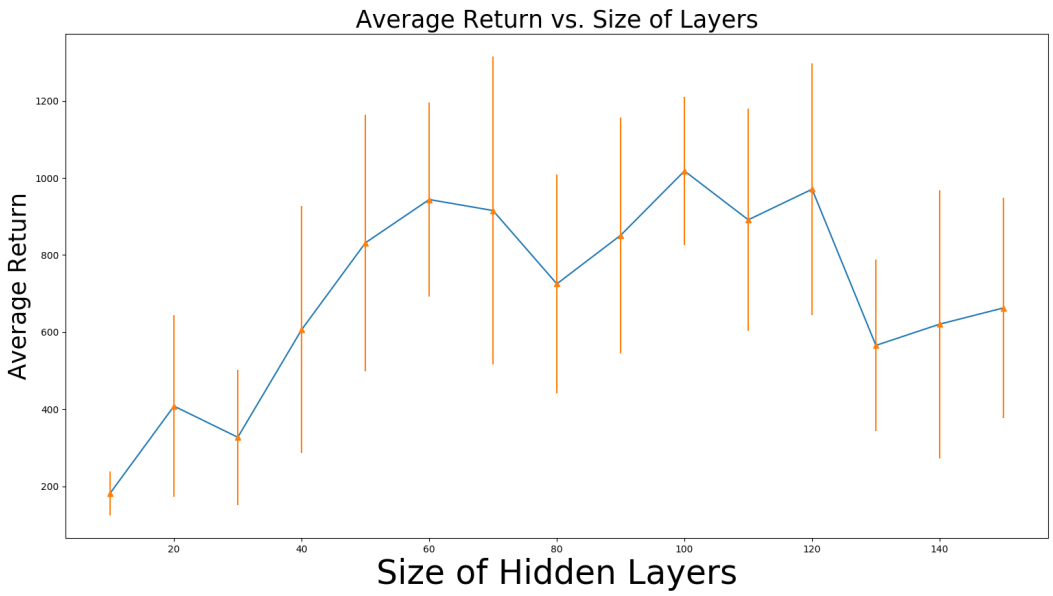

Figure 1: Plot of the average return of the evaluation data vs the size of the hidden layer. I chose this because I felt this parameter would have the most effect on the training of the neural net.

2 DAgger

2.1 Part 2

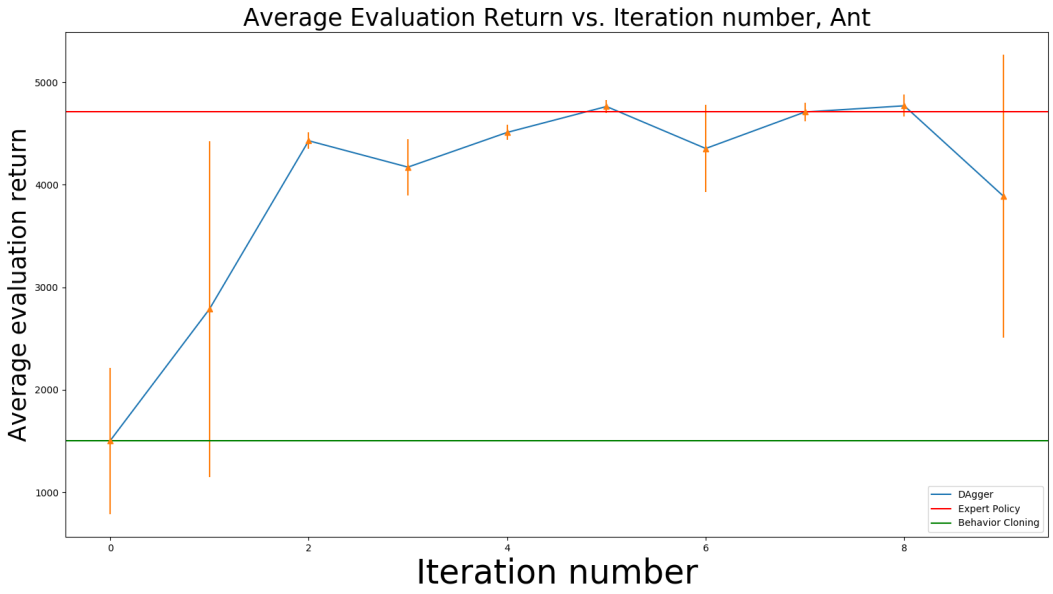

Figure 2: Plot of the average return of the evaluation data vs the iteration number. I chose this because I felt this parameter would have the most effect on the training of the neural net.

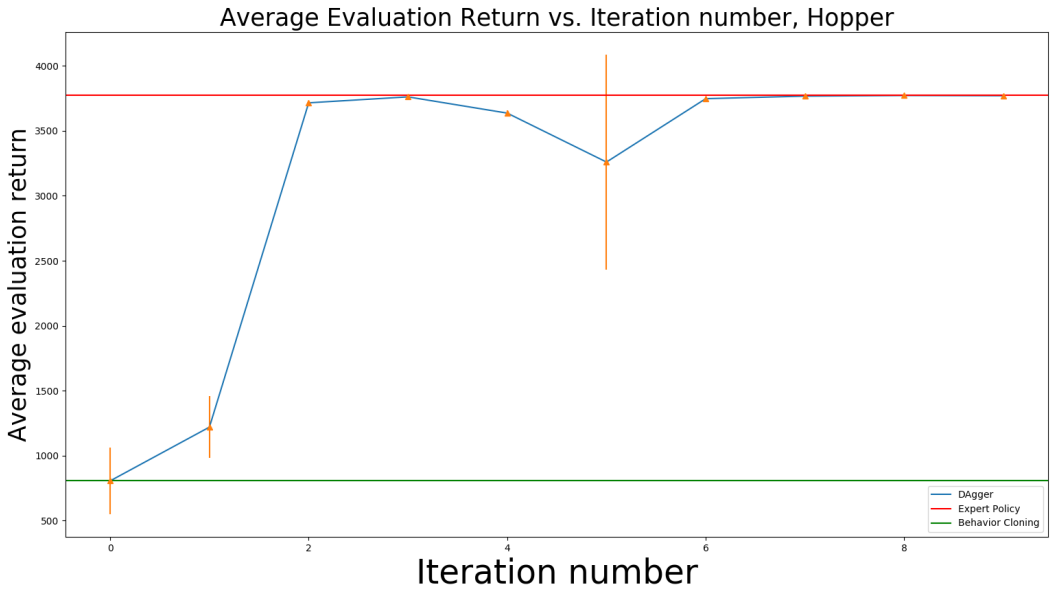

Figure 3: Plot of the average return of the evaluation data vs the iteration number. I chose this because I felt this parameter would have the most effect on the training of the neural net.

The parameters used for both the Hopper and the Ant are the following: Number of iterations: 10 Evaluation Batch size: 5000 Batch Size: 5000 Network size: 2 hidden layers of size 64 each