Homework 1

1.2

| Environment	Return Mean	Return Std
Ant	4765.671875	101.028511
Humanoid	250.591323	35.290046

In comparison, the expert achieved an average return of 4713.653320 for Ant and 10344.517578 for Humanoid, with standard deviations of 12 and 21, respectively. Therefore behavioral cloning for Ant produced almost 100% of the average return of the expert, while behavioral cloning only produced about 1% of the average return of the expert.

These results were collected with ep_len = 1000 and eval_batch_size = 5000. All of the other hyperparameters (hidden layer size, number of layers, learning rate, etc.) were the defaults specified in run_hw1.py.

1.3

These results were collected for the Ant environment.

BC Agent Performance vs. Hidden Layer Size
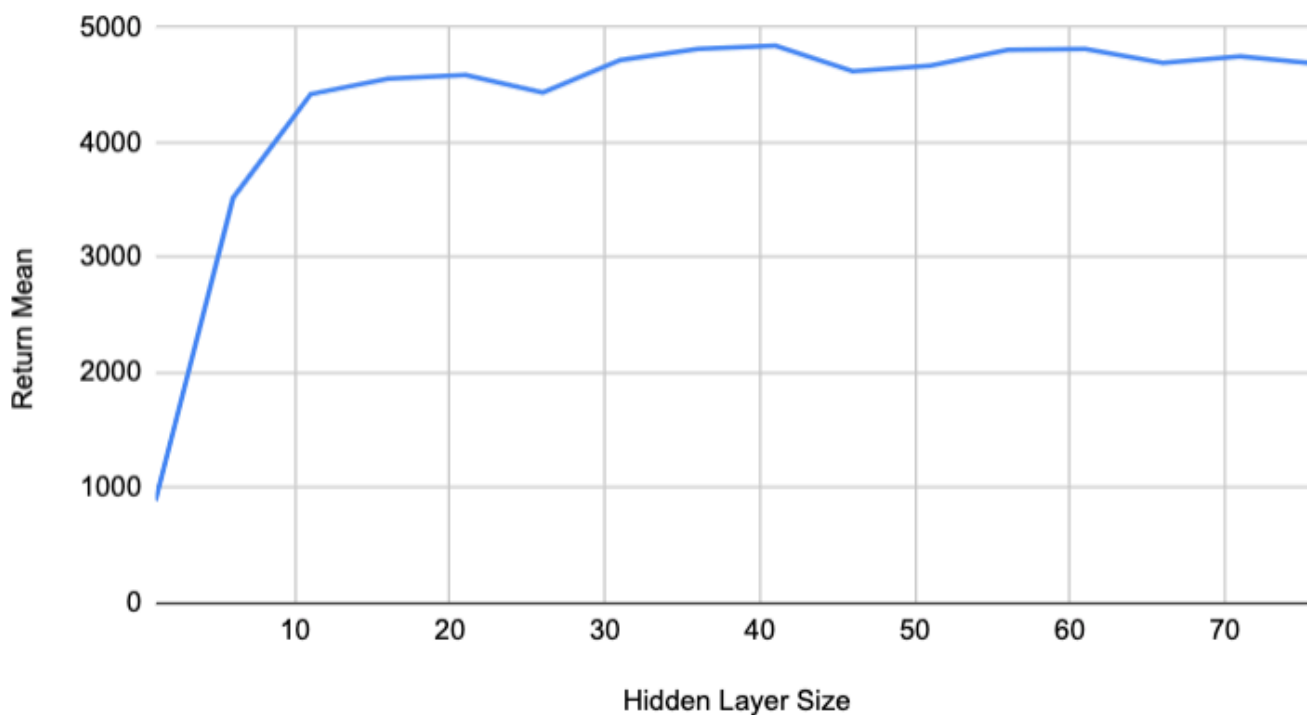
Keeping all other default parameters the same, I decided to perform a sweep across the hyperparameter corresponding to hidden layer size. I chose this parameter because I was curious if the default, 64 was too small, or if by increasing the hidden layer size (thus increasing the total number of parameters of the model) we could increase behavioral cloning performance. I was suprised to see that the hidden layer size need only be greater than 10 in order to achieve competitive results with larger sizes. Increasing the size without increasing the number of training steps did not increase performance, though it also did not hurt performance.

2.2

DAgger Return Mean vs Number of Iterations for Ant Environment
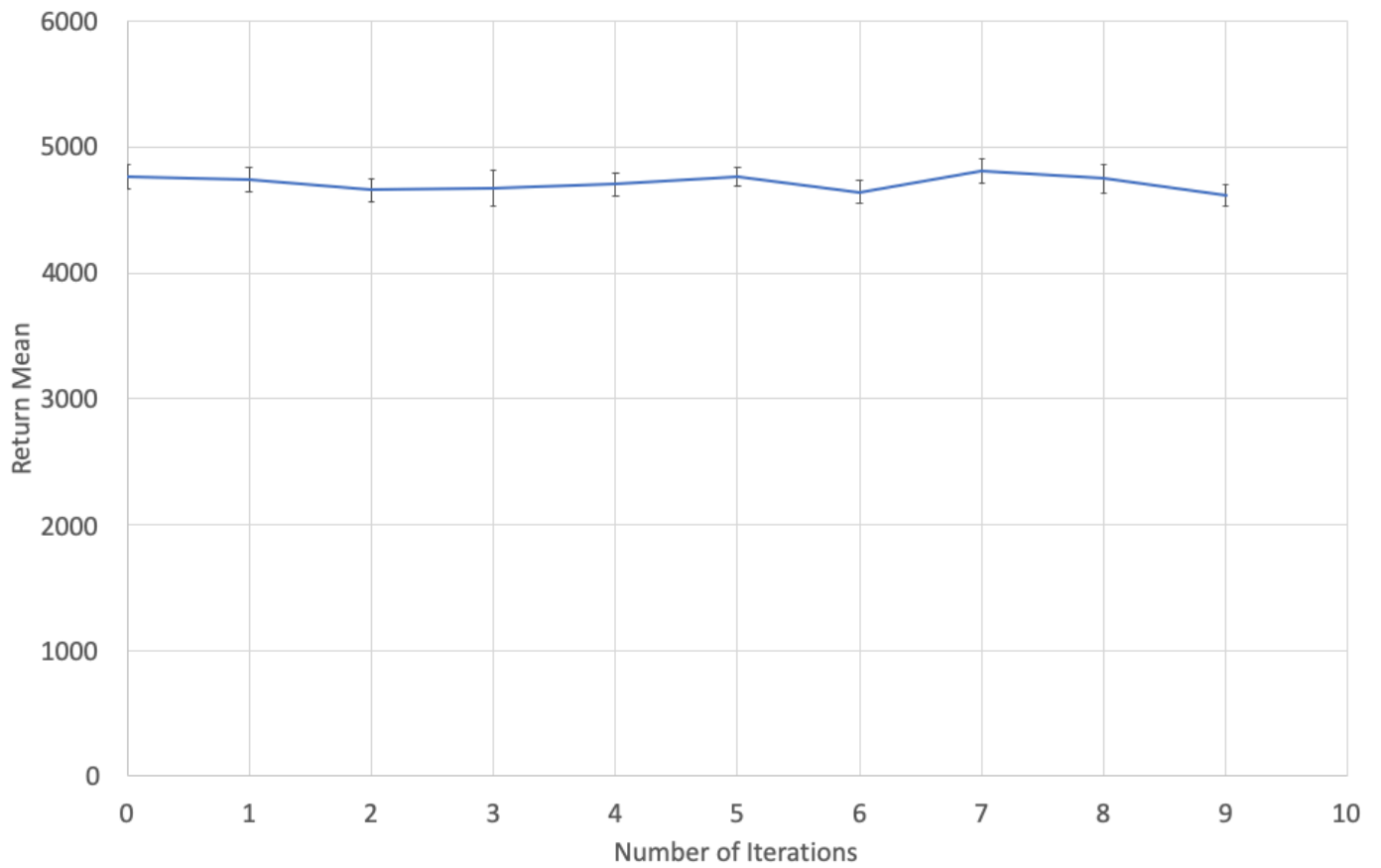
DAgger Return Mean vs Number of Iterations for Humanoid Environment
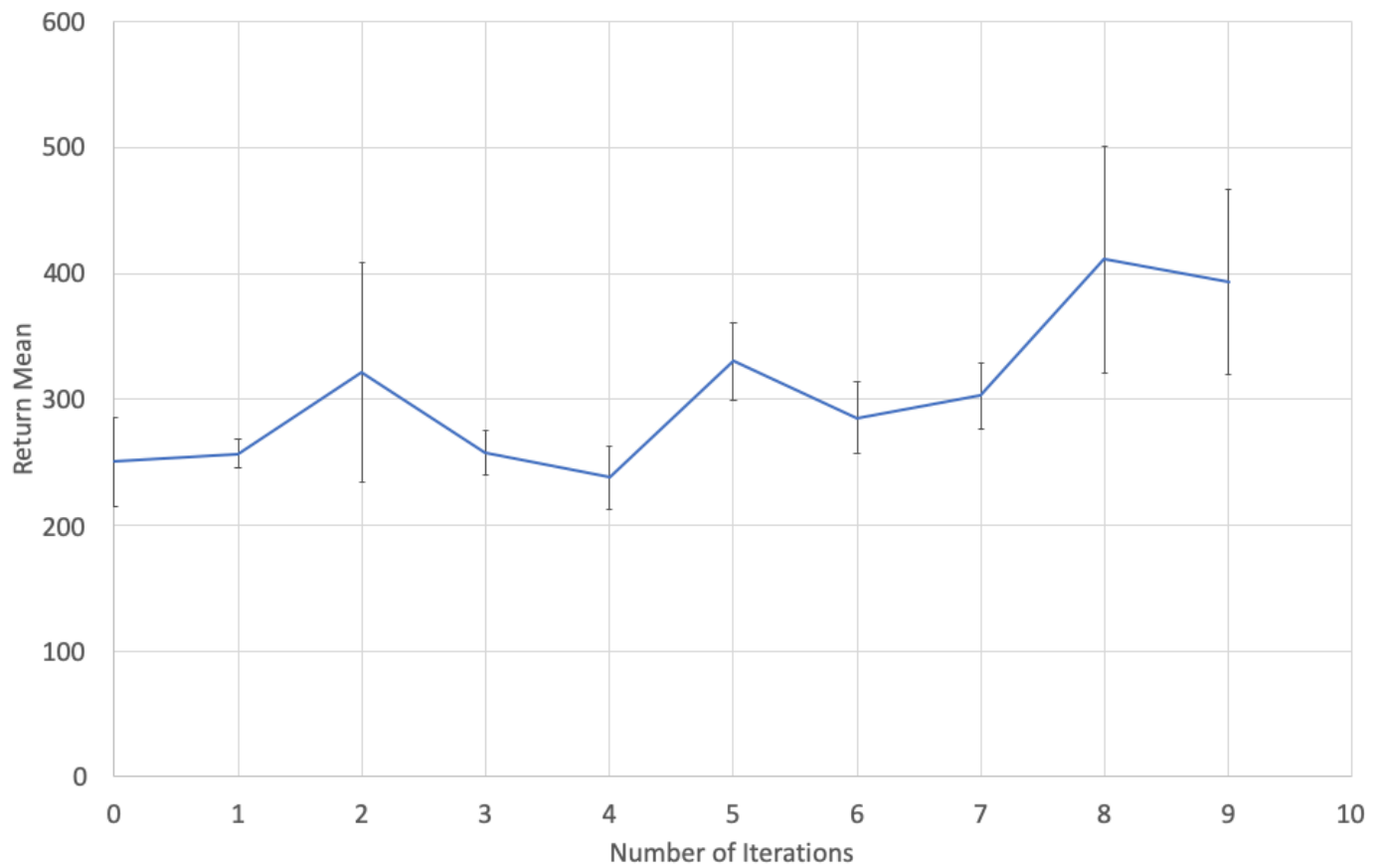
For both of these experiments, I used ep_len = 1000 and eval_batch_size = 5000. All of the other default parameters were the same.