1 Behavioral Cloning

2. Please see Table 1 for the results. The run log files for the data are
   - `p_1_2-BC_Ant-v2`
   - `p_1_2_BC_Hopper-v2`

Table 1: Results for Ant-v2 and Hopper-v2; Training data points: 2000; Evaluation data points: > 50000; Batch size: 100; Number of iterations: 10000; Hidden layers: 2; Layer size: 64;

	Ant-v2	**Hopper-v2**
Expert Policy	4713.65	3772.67
Behavioral Cloning	4605.49 ± 538.01	314.12 ± 70.24

3. Figure 1 for the results.

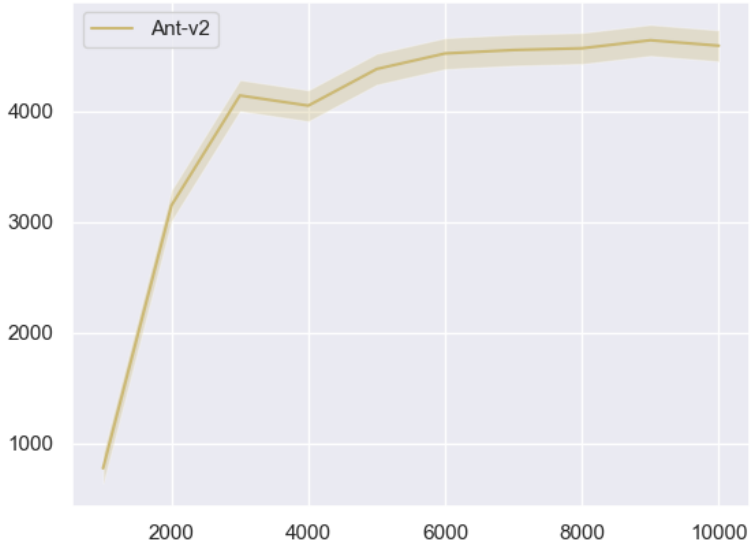

Figure 1: The figure shows the performance with respect to training iterations. For each number of iterations, 10 different seeds are used. Mean average returns are reported. Shades are the standard deviations of average returns. Note that the standard deviations shown here are not of the returns of a single run, but of the average returns of multiple runs.

2 Dagger

2. Please see Figure 2 for the results.

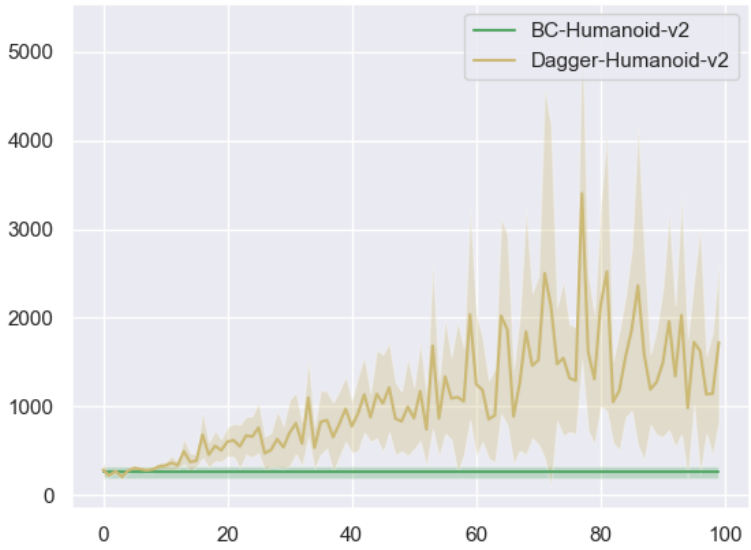

Figure 2: The figure shows the performance of Behavior Cloning and Dagger for Humanoid-v2. Most conditions are the same for section 1. For fair comparison, I set the number of steps per iteration to 50000 for Behavior Cloning, and the number of iterations and the number of steps per iterations to 100 and 500 for Dagger such that both algorithms go through the same number of datapoints during training.