

CS294–Fall 2018 — Homework 1 Solutions

Jiajian Lu, SID 3033084290

2. Behavioral Cloning

1. Experts of task on Humanoid and Hopper were used to generate roll-outs. Behavioral cloning was implemented using neural net.
2. We can see that BC agent can achieve comparable performance in Hopper task while it did bad in Humanoid task.

Table 1: Comparison between experts and BC agents

Agents	Mean of Return	Standard Deviation of Return
Humanoid Expert	10378.958833425057	59.262322171635454
Humanoid Cloning Agent	2449.1045018483537	1567.489887659889
Hopper Expert	3779.322590705575	3.8369100710996893
Hopper Cloning Agent	2080.669053080116	123.02304312553584

Cloning Agents have 4 hidden layers, 128 neurons on each layer and they are trained for 200 training epochs using all the expert data.

3. For the Hopper task, the number of training epochs was used as a hyperparameter. Training the entire data once is one epoch and we can use the entire data many times to train our network to have a better accuracy. Below is the figure of epochs vs mean of return.



Figure 1: As training epoch increases, the performance of behavior cloning agent increases but the agent might overfit after 200 epochs

3. DAgger

1. Implemented DAgger according to the algorithm.
2. DAgger improved a lot for Hopper task.

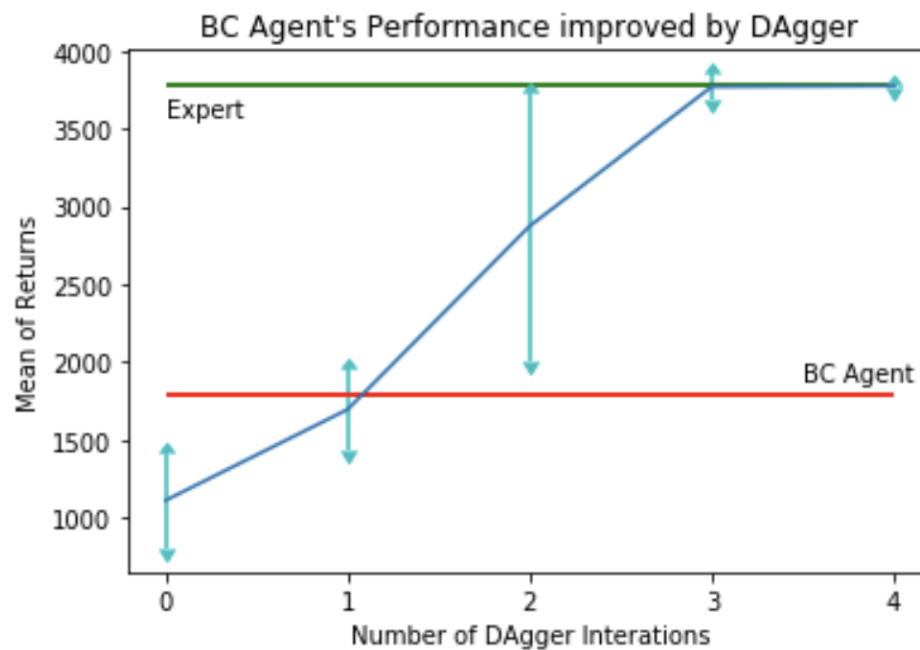


Figure 2: DAgger for 4 layers neural net with 20 training epochs