CS285– Fall 2020 — Homework 1Solutions

Tomas Santos, SID 3032632501

1. Behavioral Cloning

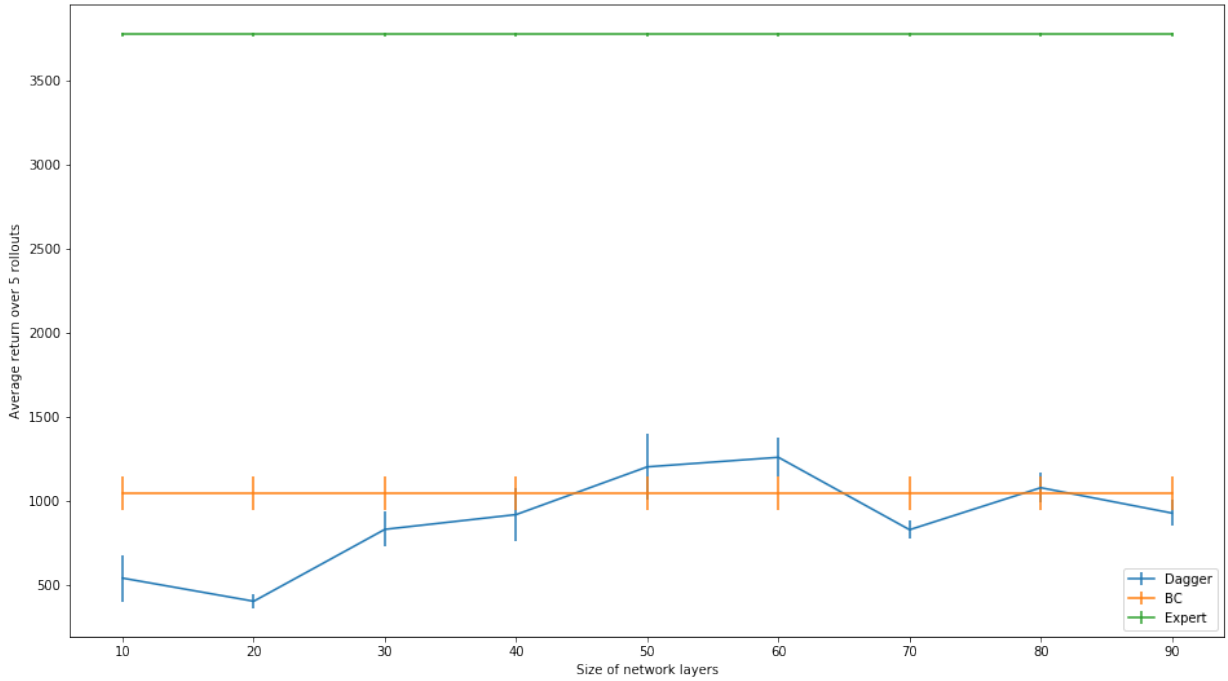
2) Run behavioral cloning (BC) and report results on two tasks: the Ant environment, where where a behavioral cloning agent should achieve at least 30% of the performance of the expert, and one environment of your choosing where it does not.

Task	Eval_AverageReturn	Eval_StdReturn	Train_AverageReturn	$\frac{\text{Eval\_AverageReturn}}{\text{Train\_AverageReturn}}$
Ant-BC	4719.7041015625	99.9489974975586	4713.6533203125	1.001
Hopper-BC	1046.32666015625	306.4642639160156	3772.67041015625	.2777

Parameters: ep_len=1000, eval_batch_size=5000, batch_size=1000, train_batch_size=100, n_layers=2, size=64, LR=5e-3

Table 1: Ant Environment

3) Experiment with one set of hyperparameters that affects the performance of the behavioral cloning

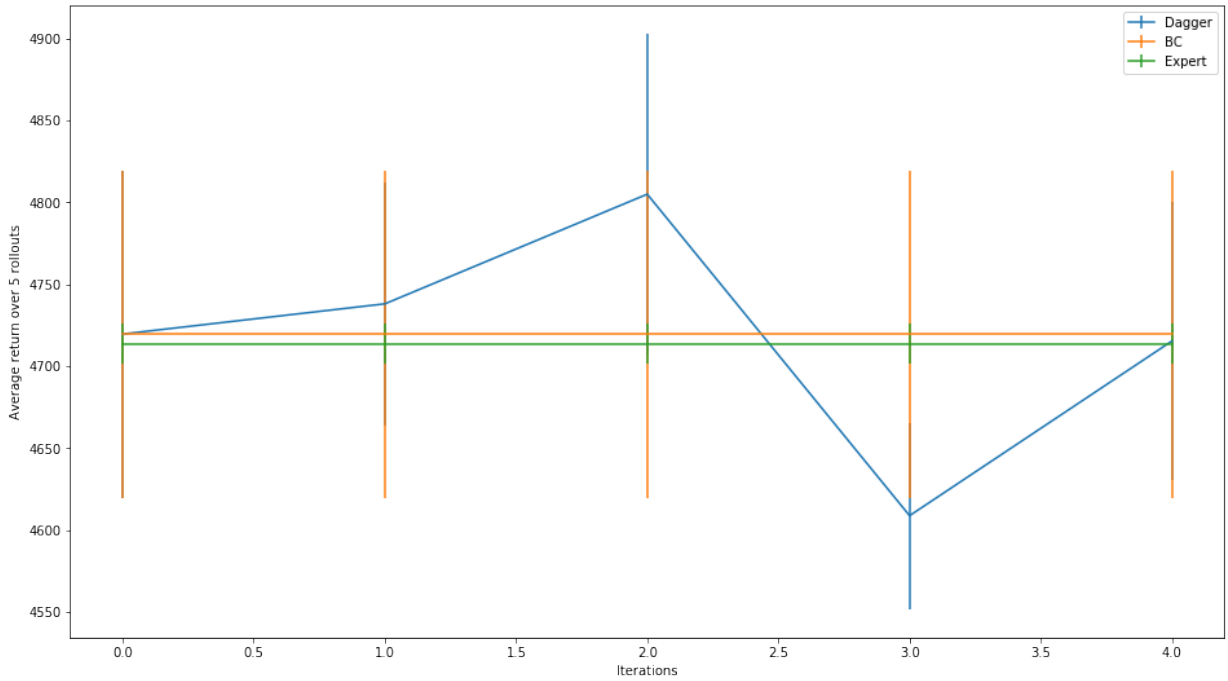


Parameters: ep_len=1000, eval_batch_size=5000, batch_size=1000, train_batch_size=100, n_layers=2, size=64, LR=5e-3

Figure 1: Hopper-v2 with modified MLP network sizes showcasing the ability of an MLP to generalize the hopper environment in different sizes

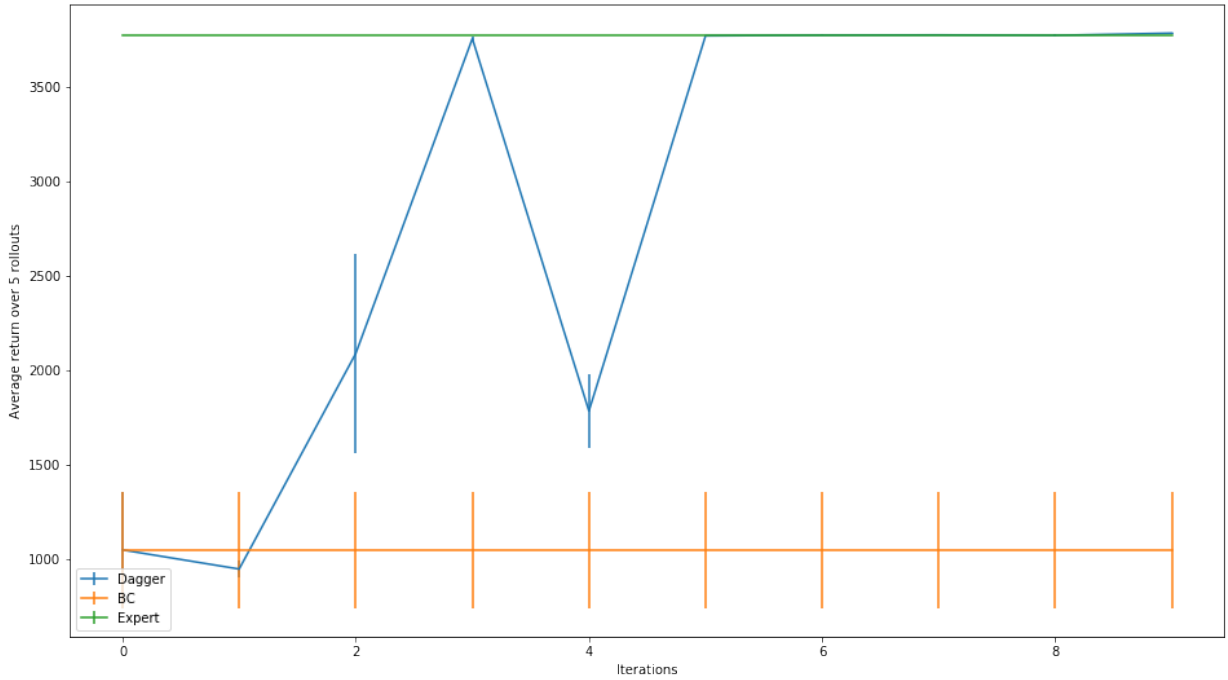
2. Dagger

Figure 2: Dagger Learning Plot for Ant-v2



Parameters: ep_len=1000, eval_batch_size=5000, batch_size=1000, train_batch_size=100, n_layers=2, size=64, LR=5e-3

Figure 3: Dagger Learning Plot for Hopper-v2



Parameters: ep_len=1000, eval_batch_size=5000, batch_size=1000, train_batch_size=100, n_layers=2, size=64, LR=5e-3