Adam Kozdrowicz

Deep RL Assignment 1: Imitation Learning

Section 2.2

Architecture of the good behavioral cloning policy: 128 Neuron Dense Layer -> LeakyRelu with alpha of 0.1 -> Logits Layer with 17 outputs.

Architecture of the poor behavioral cloning policy: 64 Neuron Dense Layer -> LeakyRelu with alpha of 0.1 -> Batch Normalization Layer -> 64 Neuron Dense Layer -> LeakyRelu with alpha of 0.1 -> Batch Normalization Layer -> Logits Layer with 17 outputs.

Both networks were trained with mean square error, Adam optimizers, and batch sizes of 32.

Code can be ran using:

```
python hw1_solutions.py hw1.2.2.1 rollout_data/Humanoid-v2.pkl Humanoid-v2
python hw1_solutions.py hw1.2.2.2 rollout_data/Humanoid-v2.pkl Humanoid-v2
```

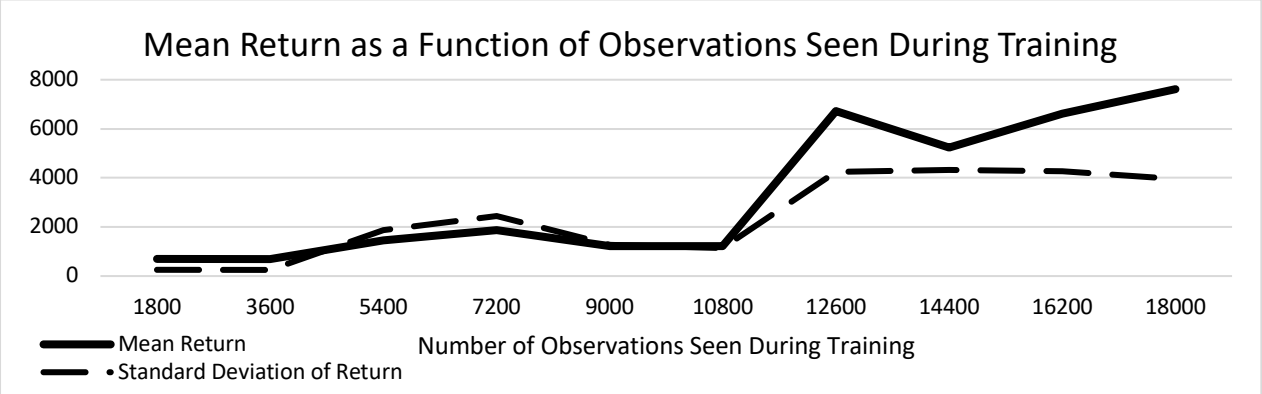
			Rollout																			
	Mean Return	Standard Deviation of Return	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Expert Policy	10404	42	10374	10306	10425	10382	10400	10485	10411	10383	10440	10394	10482	10371	10423	10357	10418	10435	10406	10354	10449	10386
Good Behavioral Cloning	9353	2963	791	10483	10439	10513	10500	10429	10371	10505	10489	418	10487	10438	10384	10395	10414	10448	10498	10498	8077	10485
Poor Behavioral Cloning	4423	3731	1448	10360	4527	6087	301	5164	457	1368	5775	1309	3727	9218	10265	385	3841	2107	10345	10358	745	680

Section 2.3

The greater the varying amount of observations the behavioral cloning algorithm sees, the higher the returns. This indicates that a diverse amount of training data makes the algorithm perform better in the testing environment.

Code can be ran using:

```
python hw1_solutions.py hw1.2.3 rollout_data/Humanoid-v2.pkl Humanoid-v2
```



Section 3.2

The DAgger algorithm was unsuccessful in generating performance that beat out the best behavioral cloning algorithm. The implementation of the DAgger algorithm is correct, and the issue lays in the fact that the underlying model is not training long enough on the newly concatenated datasets. This causes the model to start losing its accuracy over time.

Code can be ran using:

```
python hw1_solutions.py hw1.3 rollout_data/Humanoid-v2.pkl Humanoid-v2 --
expert_policy_file=experts/Humanoid-v2.pkl
```