**Deep RL Assignment 1: Imitation Learning**

**Kourosh Hakhamaneshi**

2. Behavioral Cloning

2.2)

The following settings were used during training for all tasks:

NN\_model: 2 hidden layers with 128 and 64 neurons each, respectively.

Learning rate: Decaying learning rate function: **0.005\*(0.5)^(global\_step/10000) with staircase=True.**

For training+validation set **20 rollouts** of the expert policy has been executed (20000 observation, action pairs) and **2000 epochs** with a **batch size of 256** were the settings for training.

Validation set size was 0.1 times the size of the training set.

Surprisingly the bc agent has a comparable result to the expert on all tasks. Here is a table of  comparison for 20 rollouts of each task.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  |  | Ant-v2 | HalfCheetah-v2 | Hopper-v2 | Humanoid-v2 | Reacher-v2 | Walker2d-v2 |
| Expert | avg | 4867 | 4162 | 3778 | 10391 | -4.59 | 5521 |
|  | std | 114.8 | 99 | 3.6 | 57 | 1.66 | 70.6 |
| BC | avg | 4801 | 4127.1 | 3778 | 10399 | -4 | 5508 |
|  | std | 111 | 80.14 | 2.7 | 49.26 | 1.74 | 80 |