

## CS 294-112 – Homework#1

## Section 2

2.

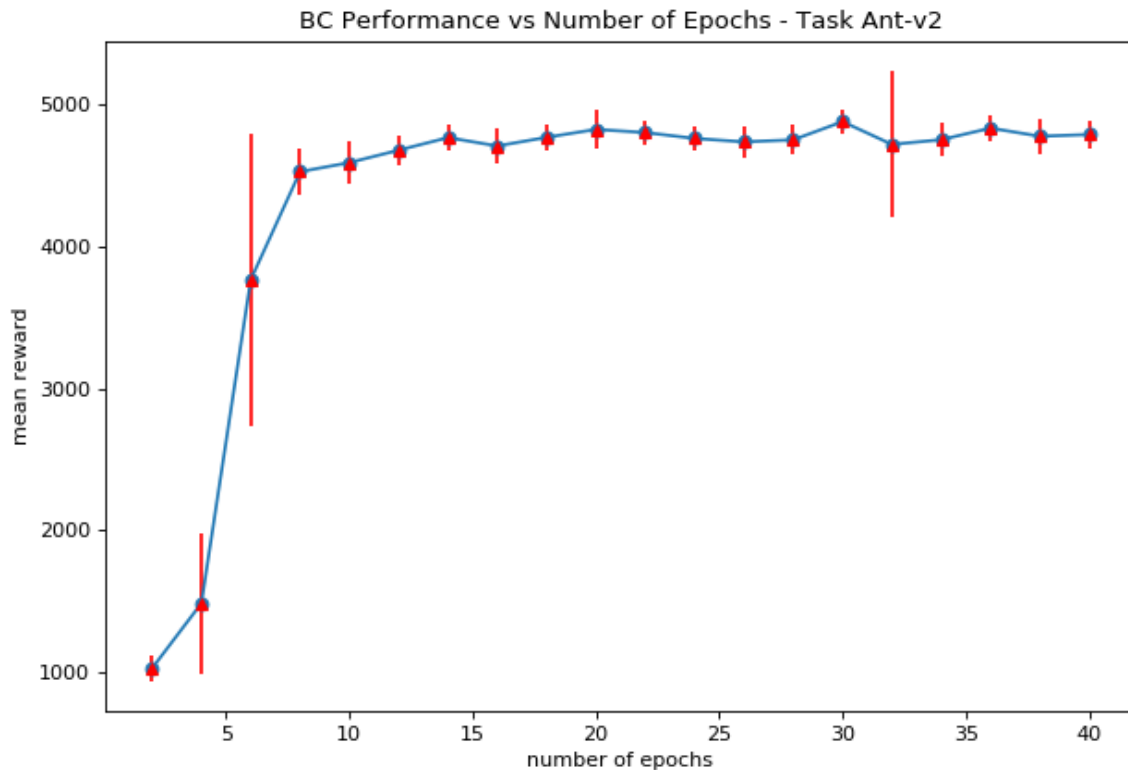
**Table 1.** Comparison between Expert policy and Behavioral Cloning (BC) policy. BC model is the fully connected neural network with two hidden layers, each of which has size (64,). Activation: ReLU.

The BC policy works well for the task Ant, HalfCheetah and Walker, but not for the remaining.

Rollouts: 20. Epochs: 100

		Ant	HalfCheetah	Hopper	Humanoid	Reacher	Walker
Mean	<b>Expert</b>	4819.59	4158.68	3780.22	10413.67	-3.71	5504.33
	<b>BC</b>	4803.70	4104.19	1894.39	1450.24	-6.89	5605.95
STD	<b>Expert</b>	99.31	52.34	3.91	46.43	1.38	106.12
	<b>BC</b>	120.25	86.86	503.84	971.65	3.87	37.93

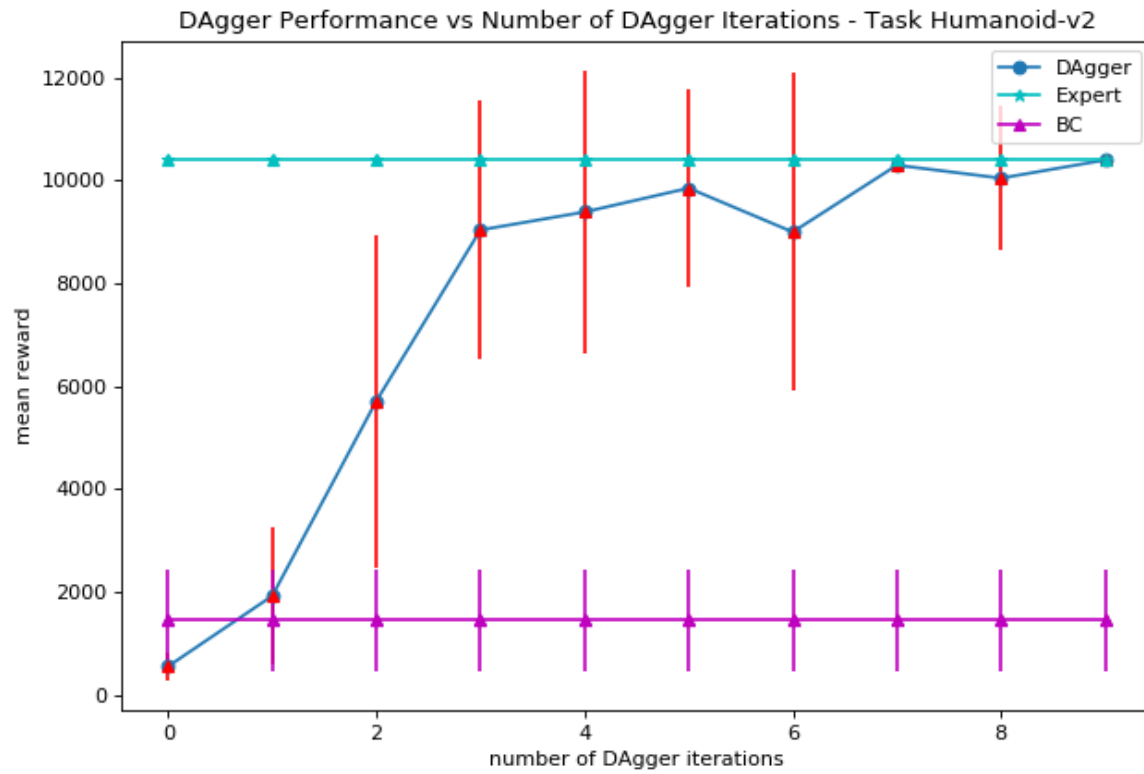
3.



**Figure 1.** Task: Ant. Hyperparameter: number of training epochs. Rationale: The performance is usually sensitive to the number of training epochs, because the initial weights ( $W$ ,  $b$ ) are generated randomly, so the more training epochs are, the more the model learn (the weights are adjusted).

### Section 3

2.



**Figure 2.** Comparison of Dagger, Expert and BC. Task: Humanoids. DAgger iterations: 10. DAgger Epochs: 50. BC Epochs: 100. Rollouts: 20. The architecture of DAgger and BC are the same as in Section 2, I.e. 2 hidden layers of size (64,) and activation ReLU. Expert and BC are not the function of Dagger Iterations so they are straight horizontal lines.

## Section 4

**Table 2.** Comparison between BC policy and BC new policy 1. The BC model doesn't change. The BC new has the same architecture with the BC but the activation is 'tanh' instead of ReLU  
The BC new policy has worse performance than the BC.

		<b>Ant</b>	<b>Humanoid</b>
Mean	<b>BC</b>	4803.70	1450.24
	<b>BC new 1</b>	4680.05	495.83
STD	<b>BC</b>	120.25	971.65
	<b>BC new 1</b>	440.61	208.15

**Table 3.** Comparison between BC policy and BC new policy 2. The BC model doesn't change. The BC new has one more hidden layers and bigger size, i.e. 3 hidden layers of size (128,) and the activation is "ReLU" as the same as the BC  
The BC new policy 2 has better performance than the BC, but the result is not consistent. Sometimes it is worse.

		<b>Ant</b>	<b>Humanoid</b>
Mean	<b>BC</b>	4803.70	1450.24
	<b>BC new 2</b>	4754.67	2473.32
STD	<b>BC</b>	120.25	971.65
	<b>BC new 2</b>	81.24	408.15