

# Homework 1

Max Fu

September 2020

## 1 Behavioral Cloning

2 Model / Parameter Comparison

	Ant-v2	Walker-v2
Expert	4713.65	5566.85
Best - 2000	$4090.54 \pm 1079.29$	$2908.09 \pm 2292.15$
Baseline - 500	$792.78 \pm 97.84$	$187.45 \pm 1.30$

Table 1: The experiments shares the same model architecture (2 hidden layers, each with dimension equals 64), with epoch length = 1000, evaluation batch size = 5000, and 10000 sets of training data collected during each iteration. The only variable is the number of gradient steps for training policy. The best experiment uses 2000 gradient steps while the baseline experiment uses 500 gradient steps.

## 2 DAgger