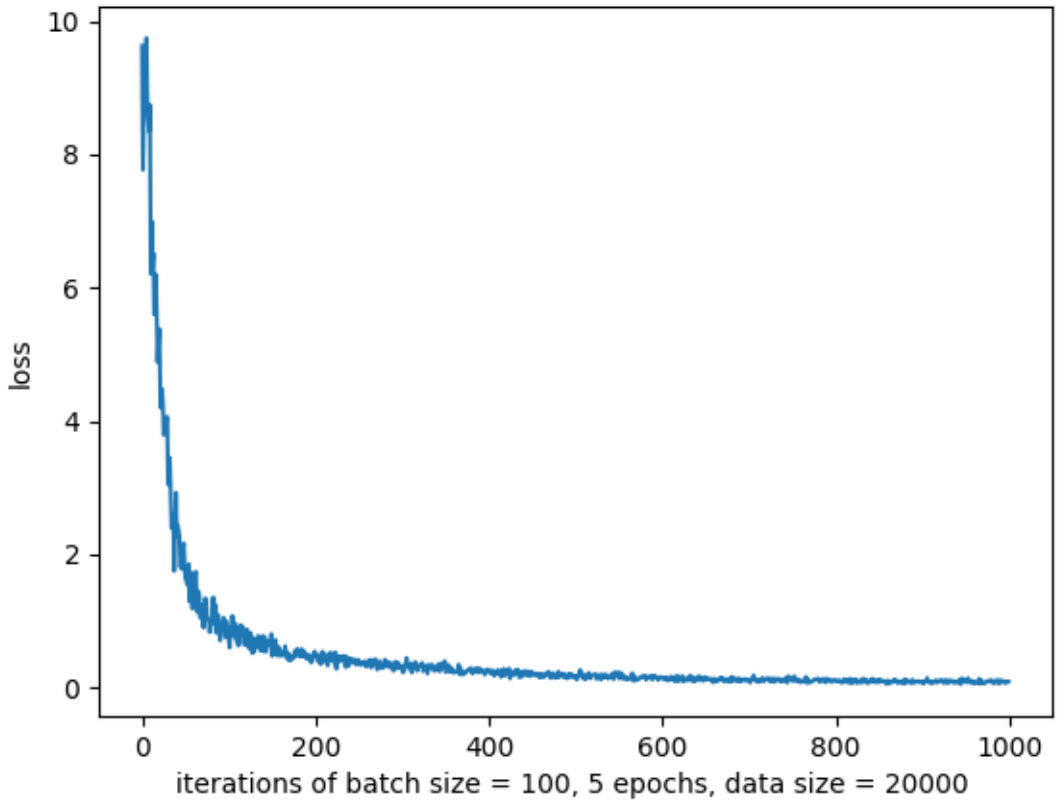
Deep RL Assignment 1 Warmup Amin Ghafari Zeydabadi 25911193

Please check HW_README.md for instructions on how to run the codes

Section 2. Warmup

The mean loss function vs. iterations



Section 3. Behavioral Cloning

1. Same network and iteration numbers for each case:

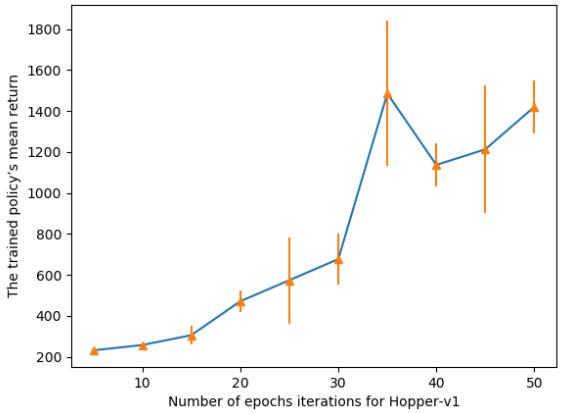
- A neural network with 3 fully connected hidden layer with 40, 40, and 20 nodes.
- Data size = 20000
- Number of epochs = 20
- Number of rollouts to find mean and std = 20

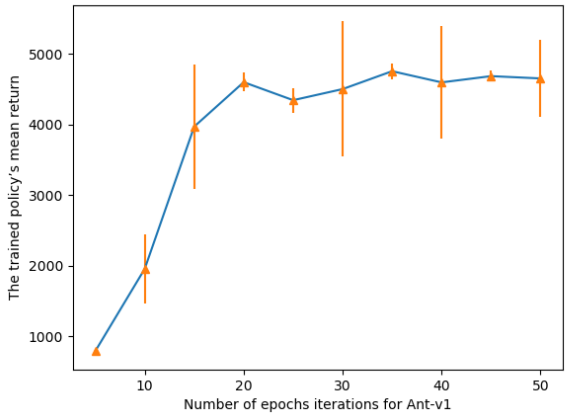
	Expert-mean	Expert-std	Trained-mean	Trained-std	
Hopper-v1	3778.45	3.72	445.81	216.96	Not a good result

Ant-v1	4758.84	398.61	4425.80	239.27	Comparable result

2. Hyperparameter: the number of training epochs,

Here we see how the number of training epochs affect the “returns” from the trained policy for Hopper-v1 and Ant-v:
- A neural network with 3 fully connected hidden layer with 40, 40, and 20 nodes.
- Data size = 20000
- epoch step = 5
- Number of rollouts to find mean and std = 20
- Number of epoch iteration = 10





Section 4. DAgger

Hopper-v1
- A neural network with 3 fully connected hidden layer with 40, 40, and 20 nodes.
- Data size = 20000
- Number of epochs for each training = 20
- Number of rollouts to find mean and std = 20
- Number of Dagger iterations = 20