|  |
| --- |
| python cs285/scripts/run\_hw1.py \  --expert\_policy\_file cs285/policies/experts/Ant.pkl \  --env\_name Ant-v4 --exp\_name bc\_ant --n\_iter 1 \  --expert\_data cs285/expert\_data/expert\_data\_Ant-v4.pkl \  --video\_log\_freq -1 |

Collecting data for eval...

steps = 500

steps = 1000

steps = 1000 ; max\_path\_length = 1000

rollout is done

Eval\_AverageReturn : 4727.23095703125

Eval\_StdReturn : 0.0

Eval\_MaxReturn : 4727.23095703125

Eval\_MinReturn : 4727.23095703125

Eval\_AverageEpLen : 1000.0

Train\_AverageReturn : 4713.6533203125

Train\_StdReturn : 12.196533203125

Train\_MaxReturn : 4725.849609375

Train\_MinReturn : 4701.45654296875

Train\_AverageEpLen : 1000.0

Train\_EnvstepsSoFar : 0

TimeSinceStart : 0.4597129821777344

Training Loss : 0.002416986972093582

Initial\_DataCollection\_AverageReturn : 4713.6533203125

|  |
| --- |
| python cs285/scripts/run\_hw1.py \  --expert\_policy\_file cs285/policies/experts/Walker2d.pkl \  --env\_name Walker2d-v4 --exp\_name bc\_walker2d --n\_iter 1 \  --expert\_data cs285/expert\_data/expert\_data\_Walker2d-v4.pkl \  --video\_log\_freq -1 |

Eval\_AverageReturn : 250.1002655029297

Eval\_StdReturn : 93.46504974365234

Eval\_MaxReturn : 370.69097900390625

Eval\_MinReturn : 6.183746337890625

Eval\_AverageEpLen : 113.33333333333333

Train\_AverageReturn : 5566.845703125

Train\_StdReturn : 9.237548828125

Train\_MaxReturn : 5576.08349609375

Train\_MinReturn : 5557.6083984375

Train\_AverageEpLen : 1000.0

Train\_EnvstepsSoFar : 0

TimeSinceStart : 0.3455338478088379

Training Loss : 0.027531182393431664

Initial\_DataCollection\_AverageReturn : 5566.845703125

|  |
| --- |
| python cs285/scripts/run\_hw1.py \  --expert\_policy\_file cs285/policies/experts/Walker2d.pkl \  --env\_name Walker2d-v4 --exp\_name bc\_walker2d --n\_iter 1 \  --expert\_data cs285/expert\_data/expert\_data\_Walker2d-v4.pkl \  --video\_log\_freq -1 \  --n\_layers 4 --batch\_size 2000 |

Eval\_AverageReturn : 111.82000732421875

Eval\_StdReturn : 276.063232421875

Eval\_MaxReturn : 1157.0645751953125

Eval\_MinReturn : 0.1564505398273468

Eval\_AverageEpLen : 45.291666666666664

Train\_AverageReturn : 5566.845703125

Train\_StdReturn : 9.237548828125

Train\_MaxReturn : 5576.08349609375

Train\_MinReturn : 5557.6083984375

Train\_AverageEpLen : 1000.0

Train\_EnvstepsSoFar : 0

TimeSinceStart : 0.5693256855010986

Training Loss : 0.016747789457440376

Initial\_DataCollection\_AverageReturn : 5566.845703125

|  |
| --- |
| python cs285/scripts/run\_hw1.py \  --expert\_policy\_file cs285/policies/experts/Ant.pkl \  --env\_name Ant-v4 --exp\_name dagger\_ant --n\_iter 10 \  --do\_dagger --expert\_data cs285/expert\_data/expert\_data\_Ant-v4.pkl \  --video\_log\_freq -1 |

\*\*\*\*\*\*\*\*\*\* Iteration 0 \*\*\*\*\*\*\*\*\*\*\*\*

Training agent using sampled data from replay buffer...

Beginning logging procedure...

Collecting data for eval...

steps = 500

steps = 1000

steps = 1000 ; max\_path\_length = 1000

rollout is done

Eval\_AverageReturn : 4727.23095703125

Eval\_StdReturn : 0.0

Eval\_MaxReturn : 4727.23095703125

Eval\_MinReturn : 4727.23095703125

Eval\_AverageEpLen : 1000.0

Train\_AverageReturn : 4713.6533203125

Train\_StdReturn : 12.196533203125

Train\_MaxReturn : 4725.849609375

Train\_MinReturn : 4701.45654296875

Train\_AverageEpLen : 1000.0

Train\_EnvstepsSoFar : 0

TimeSinceStart : 0.4637007713317871

Training Loss : 0.002416986972093582

Initial\_DataCollection\_AverageReturn : 4713.6533203125

Done logging...

\*\*\*\*\*\*\*\*\*\* Iteration 1 \*\*\*\*\*\*\*\*\*\*\*\*

Collecting data to be used for training...

batch\_size = 1000

sampling 0-th trajectory

steps = 500

steps = 1000

steps = 1000 ; max\_path\_length = 1000

rollout is done

envsteps\_this\_batch = 1000

Relabelling collected observations with labels from an expert policy...

Training agent using sampled data from replay buffer...

Beginning logging procedure...

Collecting data for eval...

steps = 500

steps = 1000

steps = 1000 ; max\_path\_length = 1000

rollout is done

Eval\_AverageReturn : 4763.7607421875

Eval\_StdReturn : 0.0

Eval\_MaxReturn : 4763.7607421875

Eval\_MinReturn : 4763.7607421875

Eval\_AverageEpLen : 1000.0

Train\_AverageReturn : 4585.6728515625

Train\_StdReturn : 0.0

Train\_MaxReturn : 4585.6728515625

Train\_MinReturn : 4585.6728515625

Train\_AverageEpLen : 1000.0

Train\_EnvstepsSoFar : 1000

TimeSinceStart : 1.1147429943084717

Training Loss : 0.0019827978685498238

Initial\_DataCollection\_AverageReturn : 4713.6533203125

Done logging...

\*\*\*\*\*\*\*\*\*\* Iteration 2 \*\*\*\*\*\*\*\*\*\*\*\*

Collecting data to be used for training...

batch\_size = 1000

sampling 0-th trajectory

steps = 500

steps = 1000

steps = 1000 ; max\_path\_length = 1000

rollout is done

envsteps\_this\_batch = 1000

Relabelling collected observations with labels from an expert policy...

Training agent using sampled data from replay buffer...

Beginning logging procedure...

Collecting data for eval...

steps = 500

steps = 1000

steps = 1000 ; max\_path\_length = 1000

rollout is done

Eval\_AverageReturn : 4857.740234375

Eval\_StdReturn : 0.0

Eval\_MaxReturn : 4857.740234375

Eval\_MinReturn : 4857.740234375

Eval\_AverageEpLen : 1000.0

Train\_AverageReturn : 4841.4267578125

Train\_StdReturn : 0.0

Train\_MaxReturn : 4841.4267578125

Train\_MinReturn : 4841.4267578125

Train\_AverageEpLen : 1000.0

Train\_EnvstepsSoFar : 2000

TimeSinceStart : 1.7824878692626953

Training Loss : 0.001708485302515328

Initial\_DataCollection\_AverageReturn : 4713.6533203125

Done logging...

\*\*\*\*\*\*\*\*\*\* Iteration 3 \*\*\*\*\*\*\*\*\*\*\*\*

Collecting data to be used for training...

batch\_size = 1000

sampling 0-th trajectory

steps = 500

steps = 1000

steps = 1000 ; max\_path\_length = 1000

rollout is done

envsteps\_this\_batch = 1000

Relabelling collected observations with labels from an expert policy...

Training agent using sampled data from replay buffer...

Beginning logging procedure...

Collecting data for eval...

steps = 500

steps = 1000

steps = 1000 ; max\_path\_length = 1000

rollout is done

Eval\_AverageReturn : 4638.30078125

Eval\_StdReturn : 0.0

Eval\_MaxReturn : 4638.30078125

Eval\_MinReturn : 4638.30078125

Eval\_AverageEpLen : 1000.0

Train\_AverageReturn : 4831.33203125

Train\_StdReturn : 0.0

Train\_MaxReturn : 4831.33203125

Train\_MinReturn : 4831.33203125

Train\_AverageEpLen : 1000.0

Train\_EnvstepsSoFar : 3000

TimeSinceStart : 2.4557690620422363

Training Loss : 0.0015135176945477724

Initial\_DataCollection\_AverageReturn : 4713.6533203125

Done logging...

\*\*\*\*\*\*\*\*\*\* Iteration 4 \*\*\*\*\*\*\*\*\*\*\*\*

Collecting data to be used for training...

batch\_size = 1000

sampling 0-th trajectory

steps = 351 ; max\_path\_length = 1000

rollout is done

envsteps\_this\_batch = 351

Relabelling collected observations with labels from an expert policy...

Training agent using sampled data from replay buffer...

Beginning logging procedure...

Collecting data for eval...

steps = 500

steps = 1000

steps = 1000 ; max\_path\_length = 1000

rollout is done

Eval\_AverageReturn : 4789.07666015625

Eval\_StdReturn : 0.0

Eval\_MaxReturn : 4789.07666015625

Eval\_MinReturn : 4789.07666015625

Eval\_AverageEpLen : 1000.0

Train\_AverageReturn : 1581.5023193359375

Train\_StdReturn : 0.0

Train\_MaxReturn : 1581.5023193359375

Train\_MinReturn : 1581.5023193359375

Train\_AverageEpLen : 351.0

Train\_EnvstepsSoFar : 3351

TimeSinceStart : 2.999171018600464

Training Loss : 0.0015012231888249516

Initial\_DataCollection\_AverageReturn : 4713.6533203125

Done logging...

\*\*\*\*\*\*\*\*\*\* Iteration 5 \*\*\*\*\*\*\*\*\*\*\*\*

Collecting data to be used for training...

batch\_size = 1000

sampling 0-th trajectory

steps = 500

steps = 1000

steps = 1000 ; max\_path\_length = 1000

rollout is done

envsteps\_this\_batch = 1000

Relabelling collected observations with labels from an expert policy...

Training agent using sampled data from replay buffer...

Beginning logging procedure...

Collecting data for eval...

steps = 500

steps = 1000

steps = 1000 ; max\_path\_length = 1000

rollout is done

Eval\_AverageReturn : 4833.0126953125

Eval\_StdReturn : 0.0

Eval\_MaxReturn : 4833.0126953125

Eval\_MinReturn : 4833.0126953125

Eval\_AverageEpLen : 1000.0

Train\_AverageReturn : 4774.67431640625

Train\_StdReturn : 0.0

Train\_MaxReturn : 4774.67431640625

Train\_MinReturn : 4774.67431640625

Train\_AverageEpLen : 1000.0

Train\_EnvstepsSoFar : 4351

TimeSinceStart : 3.6817688941955566

Training Loss : 0.0012730089947581291

Initial\_DataCollection\_AverageReturn : 4713.6533203125

Done logging...

\*\*\*\*\*\*\*\*\*\* Iteration 6 \*\*\*\*\*\*\*\*\*\*\*\*

Collecting data to be used for training...

batch\_size = 1000

sampling 0-th trajectory

steps = 500

steps = 1000

steps = 1000 ; max\_path\_length = 1000

rollout is done

envsteps\_this\_batch = 1000

Relabelling collected observations with labels from an expert policy...

Training agent using sampled data from replay buffer...

Beginning logging procedure...

Collecting data for eval...

steps = 500

steps = 1000

steps = 1000 ; max\_path\_length = 1000

rollout is done

Eval\_AverageReturn : 4684.7333984375

Eval\_StdReturn : 0.0

Eval\_MaxReturn : 4684.7333984375

Eval\_MinReturn : 4684.7333984375

Eval\_AverageEpLen : 1000.0

Train\_AverageReturn : 4780.5048828125

Train\_StdReturn : 0.0

Train\_MaxReturn : 4780.5048828125

Train\_MinReturn : 4780.5048828125

Train\_AverageEpLen : 1000.0

Train\_EnvstepsSoFar : 5351

TimeSinceStart : 4.3776257038116455

Training Loss : 0.0011357881594449282

Initial\_DataCollection\_AverageReturn : 4713.6533203125

Done logging...

\*\*\*\*\*\*\*\*\*\* Iteration 7 \*\*\*\*\*\*\*\*\*\*\*\*

Collecting data to be used for training...

batch\_size = 1000

sampling 0-th trajectory

steps = 500

steps = 852 ; max\_path\_length = 1000

rollout is done

envsteps\_this\_batch = 852

Relabelling collected observations with labels from an expert policy...

Training agent using sampled data from replay buffer...

Beginning logging procedure...

Collecting data for eval...

steps = 500

steps = 1000

steps = 1000 ; max\_path\_length = 1000

rollout is done

Eval\_AverageReturn : 4687.1630859375

Eval\_StdReturn : 0.0

Eval\_MaxReturn : 4687.1630859375

Eval\_MinReturn : 4687.1630859375

Eval\_AverageEpLen : 1000.0

Train\_AverageReturn : 3911.548095703125

Train\_StdReturn : 0.0

Train\_MaxReturn : 3911.548095703125

Train\_MinReturn : 3911.548095703125

Train\_AverageEpLen : 852.0

Train\_EnvstepsSoFar : 6203

TimeSinceStart : 5.040961980819702

Training Loss : 0.001317733433097601

Initial\_DataCollection\_AverageReturn : 4713.6533203125

Done logging...

\*\*\*\*\*\*\*\*\*\* Iteration 8 \*\*\*\*\*\*\*\*\*\*\*\*

Collecting data to be used for training...

batch\_size = 1000

sampling 0-th trajectory

steps = 500

steps = 1000

steps = 1000 ; max\_path\_length = 1000

rollout is done

envsteps\_this\_batch = 1000

Relabelling collected observations with labels from an expert policy...

Training agent using sampled data from replay buffer...

Beginning logging procedure...

Collecting data for eval...

steps = 500

steps = 1000

steps = 1000 ; max\_path\_length = 1000

rollout is done

Eval\_AverageReturn : 4819.6904296875

Eval\_StdReturn : 0.0

Eval\_MaxReturn : 4819.6904296875

Eval\_MinReturn : 4819.6904296875

Eval\_AverageEpLen : 1000.0

Train\_AverageReturn : 4647.6796875

Train\_StdReturn : 0.0

Train\_MaxReturn : 4647.6796875

Train\_MinReturn : 4647.6796875

Train\_AverageEpLen : 1000.0

Train\_EnvstepsSoFar : 7203

TimeSinceStart : 5.81683087348938

Training Loss : 0.0009454272803850472

Initial\_DataCollection\_AverageReturn : 4713.6533203125

Done logging...

\*\*\*\*\*\*\*\*\*\* Iteration 9 \*\*\*\*\*\*\*\*\*\*\*\*

Collecting data to be used for training...

batch\_size = 1000

sampling 0-th trajectory

steps = 500

steps = 1000

steps = 1000 ; max\_path\_length = 1000

rollout is done

envsteps\_this\_batch = 1000

Relabelling collected observations with labels from an expert policy...

Training agent using sampled data from replay buffer...

Beginning logging procedure...

Collecting data for eval...

steps = 500

steps = 1000

steps = 1000 ; max\_path\_length = 1000

rollout is done

Eval\_AverageReturn : 4570.1015625

Eval\_StdReturn : 0.0

Eval\_MaxReturn : 4570.1015625

Eval\_MinReturn : 4570.1015625

Eval\_AverageEpLen : 1000.0

Train\_AverageReturn : 4728.23828125

Train\_StdReturn : 0.0

Train\_MaxReturn : 4728.23828125

Train\_MinReturn : 4728.23828125

Train\_AverageEpLen : 1000.0

Train\_EnvstepsSoFar : 8203

TimeSinceStart : 6.539216756820679

Training Loss : 0.0012403689324855804

Initial\_DataCollection\_AverageReturn : 4713.6533203125

Done logging...

|  |
| --- |
| python cs285/scripts/run\_hw1.py \  --expert\_policy\_file cs285/policies/experts/Walker2d.pkl \  --env\_name Walker2d-v4 --exp\_name dagger\_walker2d --n\_iter 20 \  --do\_dagger --expert\_data cs285/expert\_data/expert\_data\_Walker2d-v4.pkl \  --video\_log\_freq -1 |

\*\*\*\*\*\*\*\*\*\* Iteration 0 \*\*\*\*\*\*\*\*\*\*\*\*

Training agent using sampled data from replay buffer...

Beginning logging procedure...

Collecting data for eval...

steps = 115 ; max\_path\_length = 1000

rollout is done

steps = 117 ; max\_path\_length = 1000

rollout is done

steps = 22 ; max\_path\_length = 1000

rollout is done

steps = 117 ; max\_path\_length = 1000

rollout is done

steps = 98 ; max\_path\_length = 1000

rollout is done

steps = 138 ; max\_path\_length = 1000

rollout is done

steps = 117 ; max\_path\_length = 1000

rollout is done

steps = 186 ; max\_path\_length = 1000

rollout is done

steps = 110 ; max\_path\_length = 1000

rollout is done

Eval\_AverageReturn : 250.1002655029297

Eval\_StdReturn : 93.46504974365234

Eval\_MaxReturn : 370.69097900390625

Eval\_MinReturn : 6.183746337890625

Eval\_AverageEpLen : 113.33333333333333

Train\_AverageReturn : 5566.845703125

Train\_StdReturn : 9.237548828125

Train\_MaxReturn : 5576.08349609375

Train\_MinReturn : 5557.6083984375

Train\_AverageEpLen : 1000.0

Train\_EnvstepsSoFar : 0

TimeSinceStart : 0.32794713973999023

Training Loss : 0.027531182393431664

Initial\_DataCollection\_AverageReturn : 5566.845703125

Done logging...

\*\*\*\*\*\*\*\*\*\* Iteration 1 \*\*\*\*\*\*\*\*\*\*\*\*

Collecting data to be used for training...

batch\_size = 1000

sampling 0-th trajectory

steps = 140 ; max\_path\_length = 1000

rollout is done

envsteps\_this\_batch = 140

Relabelling collected observations with labels from an expert policy...

Training agent using sampled data from replay buffer...

Beginning logging procedure...

Collecting data for eval...

steps = 18 ; max\_path\_length = 1000

rollout is done

steps = 18 ; max\_path\_length = 1000

rollout is done

steps = 490 ; max\_path\_length = 1000

rollout is done

steps = 20 ; max\_path\_length = 1000

rollout is done

steps = 18 ; max\_path\_length = 1000

rollout is done

steps = 500

steps = 504 ; max\_path\_length = 1000

rollout is done

Eval\_AverageReturn : 813.19384765625

Eval\_StdReturn : 1146.9317626953125

Eval\_MaxReturn : 2490.88037109375

Eval\_MinReturn : 1.9632518291473389

Eval\_AverageEpLen : 178.0

Train\_AverageReturn : 301.0333557128906

Train\_StdReturn : 0.0

Train\_MaxReturn : 301.0333557128906

Train\_MinReturn : 301.0333557128906

Train\_AverageEpLen : 140.0

Train\_EnvstepsSoFar : 140

TimeSinceStart : 0.6783010959625244

Training Loss : 0.019322369247674942

Initial\_DataCollection\_AverageReturn : 5566.845703125

Done logging...

\*\*\*\*\*\*\*\*\*\* Iteration 2 \*\*\*\*\*\*\*\*\*\*\*\*

Collecting data to be used for training...

batch\_size = 1000

sampling 0-th trajectory

steps = 20 ; max\_path\_length = 1000

rollout is done

envsteps\_this\_batch = 20

Relabelling collected observations with labels from an expert policy...

Training agent using sampled data from replay buffer...

Beginning logging procedure...

Collecting data for eval...

steps = 402 ; max\_path\_length = 1000

rollout is done

steps = 405 ; max\_path\_length = 1000

rollout is done

steps = 500

steps = 550 ; max\_path\_length = 1000

rollout is done

Eval\_AverageReturn : 2026.6265869140625

Eval\_StdReturn : 376.67755126953125

Eval\_MaxReturn : 2559.00048828125

Eval\_MinReturn : 1744.237548828125

Eval\_AverageEpLen : 452.3333333333333

Train\_AverageReturn : 3.9404807090759277

Train\_StdReturn : 0.0

Train\_MaxReturn : 3.9404807090759277

Train\_MinReturn : 3.9404807090759277

Train\_AverageEpLen : 20.0

Train\_EnvstepsSoFar : 160

TimeSinceStart : 1.0216331481933594

Training Loss : 0.01161767914891243

Initial\_DataCollection\_AverageReturn : 5566.845703125

Done logging...

\*\*\*\*\*\*\*\*\*\* Iteration 3 \*\*\*\*\*\*\*\*\*\*\*\*

Collecting data to be used for training...

batch\_size = 1000

sampling 0-th trajectory

steps = 500

steps = 652 ; max\_path\_length = 1000

rollout is done

envsteps\_this\_batch = 652

Relabelling collected observations with labels from an expert policy...

Training agent using sampled data from replay buffer...

Beginning logging procedure...

Collecting data for eval...

steps = 346 ; max\_path\_length = 1000

rollout is done

steps = 500

steps = 1000

steps = 1000 ; max\_path\_length = 1000

rollout is done

Eval\_AverageReturn : 3401.59326171875

Eval\_StdReturn : 1916.3419189453125

Eval\_MaxReturn : 5317.93505859375

Eval\_MinReturn : 1485.251220703125

Eval\_AverageEpLen : 673.0

Train\_AverageReturn : 3216.17724609375

Train\_StdReturn : 0.0

Train\_MaxReturn : 3216.17724609375

Train\_MinReturn : 3216.17724609375

Train\_AverageEpLen : 652.0

Train\_EnvstepsSoFar : 812

TimeSinceStart : 1.4535560607910156

Training Loss : 0.02077975869178772

Initial\_DataCollection\_AverageReturn : 5566.845703125

Done logging...

\*\*\*\*\*\*\*\*\*\* Iteration 4 \*\*\*\*\*\*\*\*\*\*\*\*

Collecting data to be used for training...

batch\_size = 1000

sampling 0-th trajectory

steps = 500

steps = 1000

steps = 1000 ; max\_path\_length = 1000

rollout is done

envsteps\_this\_batch = 1000

Relabelling collected observations with labels from an expert policy...

Training agent using sampled data from replay buffer...

Beginning logging procedure...

Collecting data for eval...

steps = 500

steps = 1000

steps = 1000 ; max\_path\_length = 1000

rollout is done

Eval\_AverageReturn : 5266.5986328125

Eval\_StdReturn : 0.0

Eval\_MaxReturn : 5266.5986328125

Eval\_MinReturn : 5266.5986328125

Eval\_AverageEpLen : 1000.0

Train\_AverageReturn : 5269.1484375

Train\_StdReturn : 0.0

Train\_MaxReturn : 5269.1484375

Train\_MinReturn : 5269.1484375

Train\_AverageEpLen : 1000.0

Train\_EnvstepsSoFar : 1812

TimeSinceStart : 1.8696372509002686

Training Loss : 0.027515370398759842

Initial\_DataCollection\_AverageReturn : 5566.845703125

Done logging...

\*\*\*\*\*\*\*\*\*\* Iteration 5 \*\*\*\*\*\*\*\*\*\*\*\*

Collecting data to be used for training...

batch\_size = 1000

sampling 0-th trajectory

steps = 500

steps = 1000

steps = 1000 ; max\_path\_length = 1000

rollout is done

envsteps\_this\_batch = 1000

Relabelling collected observations with labels from an expert policy...

Training agent using sampled data from replay buffer...

Beginning logging procedure...

Collecting data for eval...

steps = 500

steps = 1000

steps = 1000 ; max\_path\_length = 1000

rollout is done

Eval\_AverageReturn : 5176.466796875

Eval\_StdReturn : 0.0

Eval\_MaxReturn : 5176.466796875

Eval\_MinReturn : 5176.466796875

Eval\_AverageEpLen : 1000.0

Train\_AverageReturn : 5281.7705078125

Train\_StdReturn : 0.0

Train\_MaxReturn : 5281.7705078125

Train\_MinReturn : 5281.7705078125

Train\_AverageEpLen : 1000.0

Train\_EnvstepsSoFar : 2812

TimeSinceStart : 2.29671311378479

Training Loss : 0.018616480752825737

Initial\_DataCollection\_AverageReturn : 5566.845703125

Done logging...

\*\*\*\*\*\*\*\*\*\* Iteration 6 \*\*\*\*\*\*\*\*\*\*\*\*

Collecting data to be used for training...

batch\_size = 1000

sampling 0-th trajectory

steps = 500

steps = 1000

steps = 1000 ; max\_path\_length = 1000

rollout is done

envsteps\_this\_batch = 1000

Relabelling collected observations with labels from an expert policy...

Training agent using sampled data from replay buffer...

Beginning logging procedure...

Collecting data for eval...

steps = 500

steps = 1000

steps = 1000 ; max\_path\_length = 1000

rollout is done

Eval\_AverageReturn : 5333.89453125

Eval\_StdReturn : 0.0

Eval\_MaxReturn : 5333.89453125

Eval\_MinReturn : 5333.89453125

Eval\_AverageEpLen : 1000.0

Train\_AverageReturn : 5214.775390625

Train\_StdReturn : 0.0

Train\_MaxReturn : 5214.775390625

Train\_MinReturn : 5214.775390625

Train\_AverageEpLen : 1000.0

Train\_EnvstepsSoFar : 3812

TimeSinceStart : 2.739464044570923

Training Loss : 0.017856258898973465

Initial\_DataCollection\_AverageReturn : 5566.845703125

Done logging...

\*\*\*\*\*\*\*\*\*\* Iteration 7 \*\*\*\*\*\*\*\*\*\*\*\*

Collecting data to be used for training...

batch\_size = 1000

sampling 0-th trajectory

steps = 500

steps = 1000

steps = 1000 ; max\_path\_length = 1000

rollout is done

envsteps\_this\_batch = 1000

Relabelling collected observations with labels from an expert policy...

Training agent using sampled data from replay buffer...

Beginning logging procedure...

Collecting data for eval...

steps = 500

steps = 1000

steps = 1000 ; max\_path\_length = 1000

rollout is done

Eval\_AverageReturn : 5279.396484375

Eval\_StdReturn : 0.0

Eval\_MaxReturn : 5279.396484375

Eval\_MinReturn : 5279.396484375

Eval\_AverageEpLen : 1000.0

Train\_AverageReturn : 5314.169921875

Train\_StdReturn : 0.0

Train\_MaxReturn : 5314.169921875

Train\_MinReturn : 5314.169921875

Train\_AverageEpLen : 1000.0

Train\_EnvstepsSoFar : 4812

TimeSinceStart : 3.204177141189575

Training Loss : 0.015497862361371517

Initial\_DataCollection\_AverageReturn : 5566.845703125

Done logging...

\*\*\*\*\*\*\*\*\*\* Iteration 8 \*\*\*\*\*\*\*\*\*\*\*\*

Collecting data to be used for training...

batch\_size = 1000

sampling 0-th trajectory

steps = 316 ; max\_path\_length = 1000

rollout is done

envsteps\_this\_batch = 316

Relabelling collected observations with labels from an expert policy...

Training agent using sampled data from replay buffer...

Beginning logging procedure...

Collecting data for eval...

steps = 500

steps = 1000

steps = 1000 ; max\_path\_length = 1000

rollout is done

Eval\_AverageReturn : 5411.7890625

Eval\_StdReturn : 0.0

Eval\_MaxReturn : 5411.7890625

Eval\_MinReturn : 5411.7890625

Eval\_AverageEpLen : 1000.0

Train\_AverageReturn : 1261.9261474609375

Train\_StdReturn : 0.0

Train\_MaxReturn : 1261.9261474609375

Train\_MinReturn : 1261.9261474609375

Train\_AverageEpLen : 316.0

Train\_EnvstepsSoFar : 5128

TimeSinceStart : 3.5946431159973145

Training Loss : 0.015677472576498985

Initial\_DataCollection\_AverageReturn : 5566.845703125

Done logging...

\*\*\*\*\*\*\*\*\*\* Iteration 9 \*\*\*\*\*\*\*\*\*\*\*\*

Collecting data to be used for training...

batch\_size = 1000

sampling 0-th trajectory

steps = 500

steps = 1000

steps = 1000 ; max\_path\_length = 1000

rollout is done

envsteps\_this\_batch = 1000

Relabelling collected observations with labels from an expert policy...

Training agent using sampled data from replay buffer...

Beginning logging procedure...

Collecting data for eval...

steps = 500

steps = 1000

steps = 1000 ; max\_path\_length = 1000

rollout is done

Eval\_AverageReturn : 5376.9560546875

Eval\_StdReturn : 0.0

Eval\_MaxReturn : 5376.9560546875

Eval\_MinReturn : 5376.9560546875

Eval\_AverageEpLen : 1000.0

Train\_AverageReturn : 5449.419921875

Train\_StdReturn : 0.0

Train\_MaxReturn : 5449.419921875

Train\_MinReturn : 5449.419921875

Train\_AverageEpLen : 1000.0

Train\_EnvstepsSoFar : 6128

TimeSinceStart : 4.048988103866577

Training Loss : 0.01107844989746809

Initial\_DataCollection\_AverageReturn : 5566.845703125

Done logging...

\*\*\*\*\*\*\*\*\*\* Iteration 10 \*\*\*\*\*\*\*\*\*\*\*\*

Collecting data to be used for training...

batch\_size = 1000

sampling 0-th trajectory

steps = 500

steps = 1000

steps = 1000 ; max\_path\_length = 1000

rollout is done

envsteps\_this\_batch = 1000

Relabelling collected observations with labels from an expert policy...

Training agent using sampled data from replay buffer...

Beginning logging procedure...

Collecting data for eval...

steps = 500

steps = 1000

steps = 1000 ; max\_path\_length = 1000

rollout is done

Eval\_AverageReturn : 5327.3818359375

Eval\_StdReturn : 0.0

Eval\_MaxReturn : 5327.3818359375

Eval\_MinReturn : 5327.3818359375

Eval\_AverageEpLen : 1000.0

Train\_AverageReturn : 5360.94970703125

Train\_StdReturn : 0.0

Train\_MaxReturn : 5360.94970703125

Train\_MinReturn : 5360.94970703125

Train\_AverageEpLen : 1000.0

Train\_EnvstepsSoFar : 7128

TimeSinceStart : 4.523435115814209

Training Loss : 0.012112093158066273

Initial\_DataCollection\_AverageReturn : 5566.845703125

Done logging...

\*\*\*\*\*\*\*\*\*\* Iteration 11 \*\*\*\*\*\*\*\*\*\*\*\*

Collecting data to be used for training...

batch\_size = 1000

sampling 0-th trajectory

steps = 500

steps = 1000

steps = 1000 ; max\_path\_length = 1000

rollout is done

envsteps\_this\_batch = 1000

Relabelling collected observations with labels from an expert policy...

Training agent using sampled data from replay buffer...

Beginning logging procedure...

Collecting data for eval...

steps = 500

steps = 1000

steps = 1000 ; max\_path\_length = 1000

rollout is done

Eval\_AverageReturn : 5225.982421875

Eval\_StdReturn : 0.0

Eval\_MaxReturn : 5225.982421875

Eval\_MinReturn : 5225.982421875

Eval\_AverageEpLen : 1000.0

Train\_AverageReturn : 5326.40185546875

Train\_StdReturn : 0.0

Train\_MaxReturn : 5326.40185546875

Train\_MinReturn : 5326.40185546875

Train\_AverageEpLen : 1000.0

Train\_EnvstepsSoFar : 8128

TimeSinceStart : 5.01539421081543

Training Loss : 0.01474730484187603

Initial\_DataCollection\_AverageReturn : 5566.845703125

Done logging...

\*\*\*\*\*\*\*\*\*\* Iteration 12 \*\*\*\*\*\*\*\*\*\*\*\*

Collecting data to be used for training...

batch\_size = 1000

sampling 0-th trajectory

steps = 500

steps = 1000

steps = 1000 ; max\_path\_length = 1000

rollout is done

envsteps\_this\_batch = 1000

Relabelling collected observations with labels from an expert policy...

Training agent using sampled data from replay buffer...

Beginning logging procedure...

Collecting data for eval...

steps = 500

steps = 1000

steps = 1000 ; max\_path\_length = 1000

rollout is done

Eval\_AverageReturn : 5347.69970703125

Eval\_StdReturn : 0.0

Eval\_MaxReturn : 5347.69970703125

Eval\_MinReturn : 5347.69970703125

Eval\_AverageEpLen : 1000.0

Train\_AverageReturn : 5417.8759765625

Train\_StdReturn : 0.0

Train\_MaxReturn : 5417.8759765625

Train\_MinReturn : 5417.8759765625

Train\_AverageEpLen : 1000.0

Train\_EnvstepsSoFar : 9128

TimeSinceStart : 5.519887208938599

Training Loss : 0.011027525179088116

Initial\_DataCollection\_AverageReturn : 5566.845703125

Done logging...

\*\*\*\*\*\*\*\*\*\* Iteration 13 \*\*\*\*\*\*\*\*\*\*\*\*

Collecting data to be used for training...

batch\_size = 1000

sampling 0-th trajectory

steps = 500

steps = 1000

steps = 1000 ; max\_path\_length = 1000

rollout is done

envsteps\_this\_batch = 1000

Relabelling collected observations with labels from an expert policy...

Training agent using sampled data from replay buffer...

Beginning logging procedure...

Collecting data for eval...

steps = 500

steps = 1000

steps = 1000 ; max\_path\_length = 1000

rollout is done

Eval\_AverageReturn : 5402.6865234375

Eval\_StdReturn : 0.0

Eval\_MaxReturn : 5402.6865234375

Eval\_MinReturn : 5402.6865234375

Eval\_AverageEpLen : 1000.0

Train\_AverageReturn : 5395.5537109375

Train\_StdReturn : 0.0

Train\_MaxReturn : 5395.5537109375

Train\_MinReturn : 5395.5537109375

Train\_AverageEpLen : 1000.0

Train\_EnvstepsSoFar : 10128

TimeSinceStart : 6.028657913208008

Training Loss : 0.008486851118505001

Initial\_DataCollection\_AverageReturn : 5566.845703125

Done logging...

\*\*\*\*\*\*\*\*\*\* Iteration 14 \*\*\*\*\*\*\*\*\*\*\*\*

Collecting data to be used for training...

batch\_size = 1000

sampling 0-th trajectory

steps = 500

steps = 1000

steps = 1000 ; max\_path\_length = 1000

rollout is done

envsteps\_this\_batch = 1000

Relabelling collected observations with labels from an expert policy...

Training agent using sampled data from replay buffer...

Beginning logging procedure...

Collecting data for eval...

steps = 500

steps = 658 ; max\_path\_length = 1000

rollout is done

steps = 500

steps = 1000

steps = 1000 ; max\_path\_length = 1000

rollout is done

Eval\_AverageReturn : 4143.978515625

Eval\_StdReturn : 1073.1016845703125

Eval\_MaxReturn : 5217.080078125

Eval\_MinReturn : 3070.876708984375

Eval\_AverageEpLen : 829.0

Train\_AverageReturn : 5412.77783203125

Train\_StdReturn : 0.0

Train\_MaxReturn : 5412.77783203125

Train\_MinReturn : 5412.77783203125

Train\_AverageEpLen : 1000.0

Train\_EnvstepsSoFar : 11128

TimeSinceStart : 6.610745191574097

Training Loss : 0.011892340146005154

Initial\_DataCollection\_AverageReturn : 5566.845703125

Done logging...

\*\*\*\*\*\*\*\*\*\* Iteration 15 \*\*\*\*\*\*\*\*\*\*\*\*

Collecting data to be used for training...

batch\_size = 1000

sampling 0-th trajectory

steps = 500

steps = 1000

steps = 1000 ; max\_path\_length = 1000

rollout is done

envsteps\_this\_batch = 1000

Relabelling collected observations with labels from an expert policy...

Training agent using sampled data from replay buffer...

Beginning logging procedure...

Collecting data for eval...

steps = 500

steps = 1000

steps = 1000 ; max\_path\_length = 1000

rollout is done

Eval\_AverageReturn : 5345.263671875

Eval\_StdReturn : 0.0

Eval\_MaxReturn : 5345.263671875

Eval\_MinReturn : 5345.263671875

Eval\_AverageEpLen : 1000.0

Train\_AverageReturn : 5252.75634765625

Train\_StdReturn : 0.0

Train\_MaxReturn : 5252.75634765625

Train\_MinReturn : 5252.75634765625

Train\_AverageEpLen : 1000.0

Train\_EnvstepsSoFar : 12128

TimeSinceStart : 7.13176703453064

Training Loss : 0.010578212328255177

Initial\_DataCollection\_AverageReturn : 5566.845703125

Done logging...

\*\*\*\*\*\*\*\*\*\* Iteration 16 \*\*\*\*\*\*\*\*\*\*\*\*

Collecting data to be used for training...

batch\_size = 1000

sampling 0-th trajectory

steps = 402 ; max\_path\_length = 1000

rollout is done

envsteps\_this\_batch = 402

Relabelling collected observations with labels from an expert policy...

Training agent using sampled data from replay buffer...

Beginning logging procedure...

Collecting data for eval...

steps = 500

steps = 1000

steps = 1000 ; max\_path\_length = 1000

rollout is done

Eval\_AverageReturn : 5356.47509765625

Eval\_StdReturn : 0.0

Eval\_MaxReturn : 5356.47509765625

Eval\_MinReturn : 5356.47509765625

Eval\_AverageEpLen : 1000.0

Train\_AverageReturn : 1751.2069091796875

Train\_StdReturn : 0.0

Train\_MaxReturn : 1751.2069091796875

Train\_MinReturn : 1751.2069091796875

Train\_AverageEpLen : 402.0

Train\_EnvstepsSoFar : 12530

TimeSinceStart : 7.600750207901001

Training Loss : 0.008783700875937939

Initial\_DataCollection\_AverageReturn : 5566.845703125

Done logging...

\*\*\*\*\*\*\*\*\*\* Iteration 17 \*\*\*\*\*\*\*\*\*\*\*\*

Collecting data to be used for training...

batch\_size = 1000

sampling 0-th trajectory

steps = 305 ; max\_path\_length = 1000

rollout is done

envsteps\_this\_batch = 305

Relabelling collected observations with labels from an expert policy...

Training agent using sampled data from replay buffer...

Beginning logging procedure...

Collecting data for eval...

steps = 500

steps = 1000

steps = 1000 ; max\_path\_length = 1000

rollout is done

Eval\_AverageReturn : 5360.5908203125

Eval\_StdReturn : 0.0

Eval\_MaxReturn : 5360.5908203125

Eval\_MinReturn : 5360.5908203125

Eval\_AverageEpLen : 1000.0

Train\_AverageReturn : 1245.969970703125

Train\_StdReturn : 0.0

Train\_MaxReturn : 1245.969970703125

Train\_MinReturn : 1245.969970703125

Train\_AverageEpLen : 305.0

Train\_EnvstepsSoFar : 12835

TimeSinceStart : 8.066989183425903

Training Loss : 0.012322533875703812

Initial\_DataCollection\_AverageReturn : 5566.845703125

Done logging...

\*\*\*\*\*\*\*\*\*\* Iteration 18 \*\*\*\*\*\*\*\*\*\*\*\*

Collecting data to be used for training...

batch\_size = 1000

sampling 0-th trajectory

steps = 500

steps = 1000

steps = 1000 ; max\_path\_length = 1000

rollout is done

envsteps\_this\_batch = 1000

Relabelling collected observations with labels from an expert policy...

Training agent using sampled data from replay buffer...

Beginning logging procedure...

Collecting data for eval...

steps = 400 ; max\_path\_length = 1000

rollout is done

steps = 500

steps = 1000

steps = 1000 ; max\_path\_length = 1000

rollout is done

Eval\_AverageReturn : 3566.977783203125

Eval\_StdReturn : 1811.818115234375

Eval\_MaxReturn : 5378.7958984375

Eval\_MinReturn : 1755.15966796875

Eval\_AverageEpLen : 700.0

Train\_AverageReturn : 5383.228515625

Train\_StdReturn : 0.0

Train\_MaxReturn : 5383.228515625

Train\_MinReturn : 5383.228515625

Train\_AverageEpLen : 1000.0

Train\_EnvstepsSoFar : 13835

TimeSinceStart : 8.637742280960083

Training Loss : 0.006868940778076649

Initial\_DataCollection\_AverageReturn : 5566.845703125

Done logging...

\*\*\*\*\*\*\*\*\*\* Iteration 19 \*\*\*\*\*\*\*\*\*\*\*\*

Collecting data to be used for training...

batch\_size = 1000

sampling 0-th trajectory

steps = 500

steps = 1000

steps = 1000 ; max\_path\_length = 1000

rollout is done

envsteps\_this\_batch = 1000

Relabelling collected observations with labels from an expert policy...

Training agent using sampled data from replay buffer...

Beginning logging procedure...

Collecting data for eval...

steps = 500

steps = 1000

steps = 1000 ; max\_path\_length = 1000

rollout is done

Eval\_AverageReturn : 5389.3037109375

Eval\_StdReturn : 0.0

Eval\_MaxReturn : 5389.3037109375

Eval\_MinReturn : 5389.3037109375

Eval\_AverageEpLen : 1000.0

Train\_AverageReturn : 5372.25390625

Train\_StdReturn : 0.0

Train\_MaxReturn : 5372.25390625

Train\_MinReturn : 5372.25390625

Train\_AverageEpLen : 1000.0

Train\_EnvstepsSoFar : 14835

TimeSinceStart : 9.177519083023071

Training Loss : 0.008911849930882454

Initial\_DataCollection\_AverageReturn : 5566.845703125

Done logging...