Behavioral Cloning

1.2: For this question, I used the following setting of parameters in both environments:

Expert Data	
ep_len	1000
num_agent_train_steps_er_iter	2000
n_iter	1
batch_size	2000
Batches & Buffers	
eval_batch_size	5000
train_batch_size	200
max_replay_buffer_size	1000000
Network	
n_layers	3
size	128
learning_rate	5e-3

a. In the Ant environment

Eval. Return Avg.	Eval. Return Std.	Expert Reward	% of performance
3498	1343	4714	74%

b. In the Walker environment:

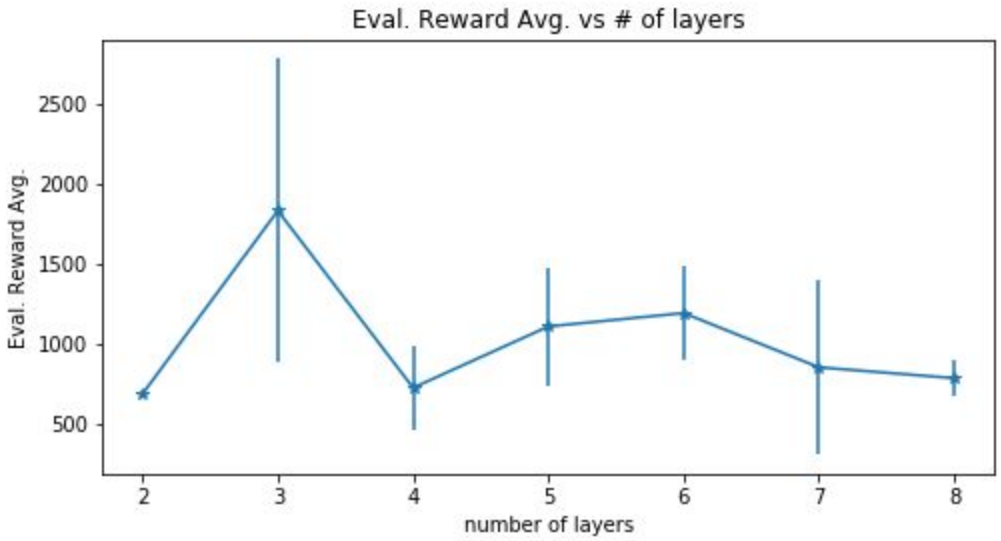
Eval. Return Avg.	Eval. Return Std.	Expert Reward	% of performance
1565	1371	5567	28%

1.3: In this experiment, I explore how the number of layers of MLP in the policy affect the performance of the agent in the Ant environment. I varied the number of layers from 2 to 8 with step size 1. The rest of parameters are the same as the previous session. The purpose of choosing to vary the number of layers in MLP is to see whether the increase in the number of parameters in neural networks can help the agent perform better.

Here are the parameters I used:

Expert Data	
ep_len	1000
num_agent_train_steps_er_iter	1000
n_iter	1
batch_size	1000
Batches & Buffers	
eval_batch_size	5000
train_batch_size	100
max_replay_buffer_size	1000000
Network	
n_layers	3
size	128
learning_rate	5e-3

Here is the result, where the error bars indicate the standard deviations:

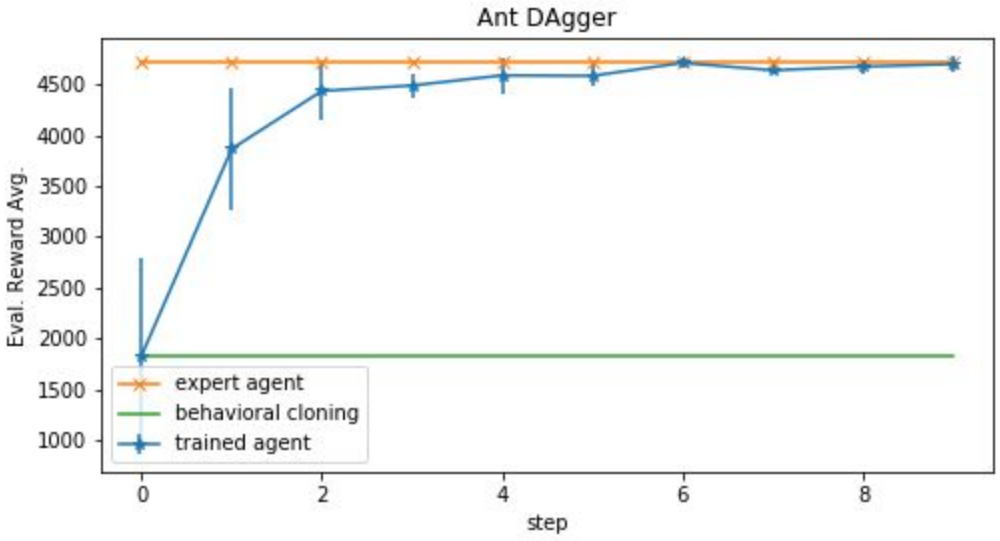


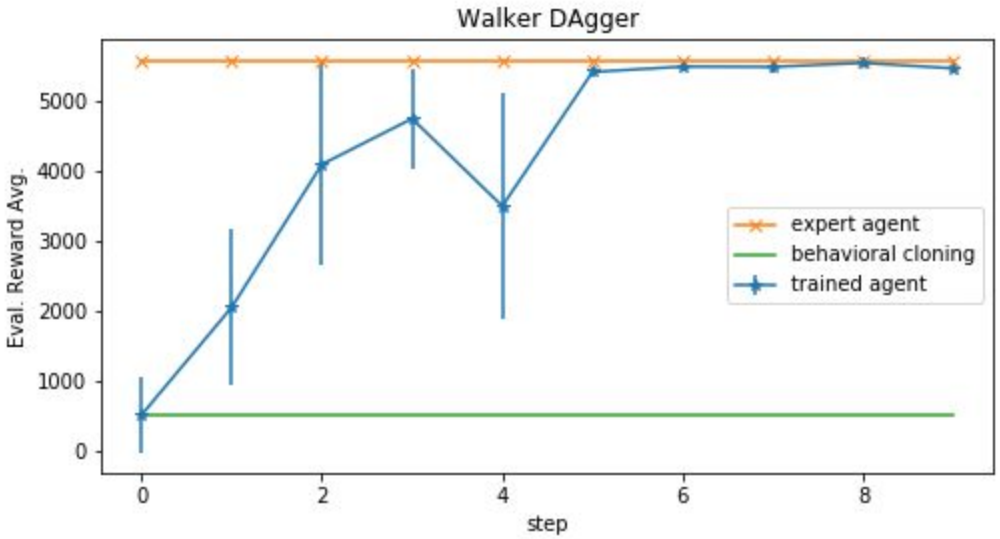
The result of this experiment is telling us that simply increasing the number of layers won't help to learn a good agent. But it appeared that more layers can reduce variance of prediction.

DAgger

For both environments, we use the following parameters

Expert Data	
ep_len	1000
num_agent_train_steps_er_iter	1000
n_iter	10
batch_size	1000
Batches & Buffers	
eval_batch_size	5000
train_batch_size	100
max_replay_buffer_size	1000000
Network	
n_layers	3
size	128
learning_rate	5e-3


Ant DAgger
Eval. Reward Avg.
4500
4000
3500
3000
2500
2000
1500
1000
0
2
4
6
8
step
expert agent
behavioral cloning
trained agent


Walker DAgger
Eval. Reward Avg.
5000
4000
3000
2000
1000
0
0
2
4
6
8
step
expert agent
behavioral cloning
trained agent