HW1 REPORT

Q1 Behavioral Cloning

	Ant	Hooper
`Eval_AverageReturn`	4775.42041015625	1113.5980224609375
`Eval_StdReturn`	64.87841033935547	33.29758834838867
`Train_AverageReturn`	4713.6533203125	3772.67041015625
`Train_StdReturn`	12.196533203125	1.9483642578124
`Training Loss`	0.0013712793588638306	0.007351272273808718

Table 1. Comparison between results of **Ant** and **HalfCheetah** environment, with parameters `n_iter=1` , `eval_batch_size=10000` , `n_layers=2` , `size=64` , `learning_rate=5e-3` .

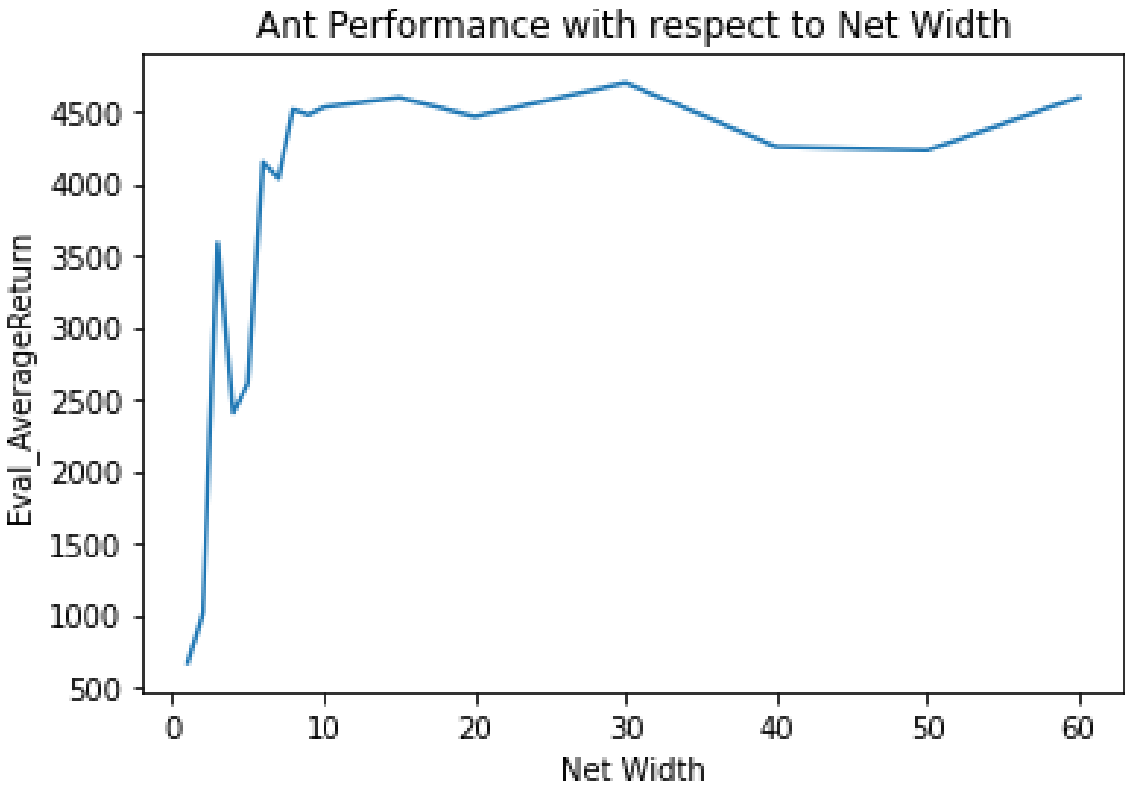

Figure 1. **Ant** environment performance with respect to Net Width(`--size`). The other parameters are the same with those in Table 1. The reason why I chose net width is that I want to know how wide is our network needs to be to learn this environment. The result shows that a size of 9 is approximately enough for best performance. And below that, the network is too small to learn all the features to imitate properly.

Q2 DAgger

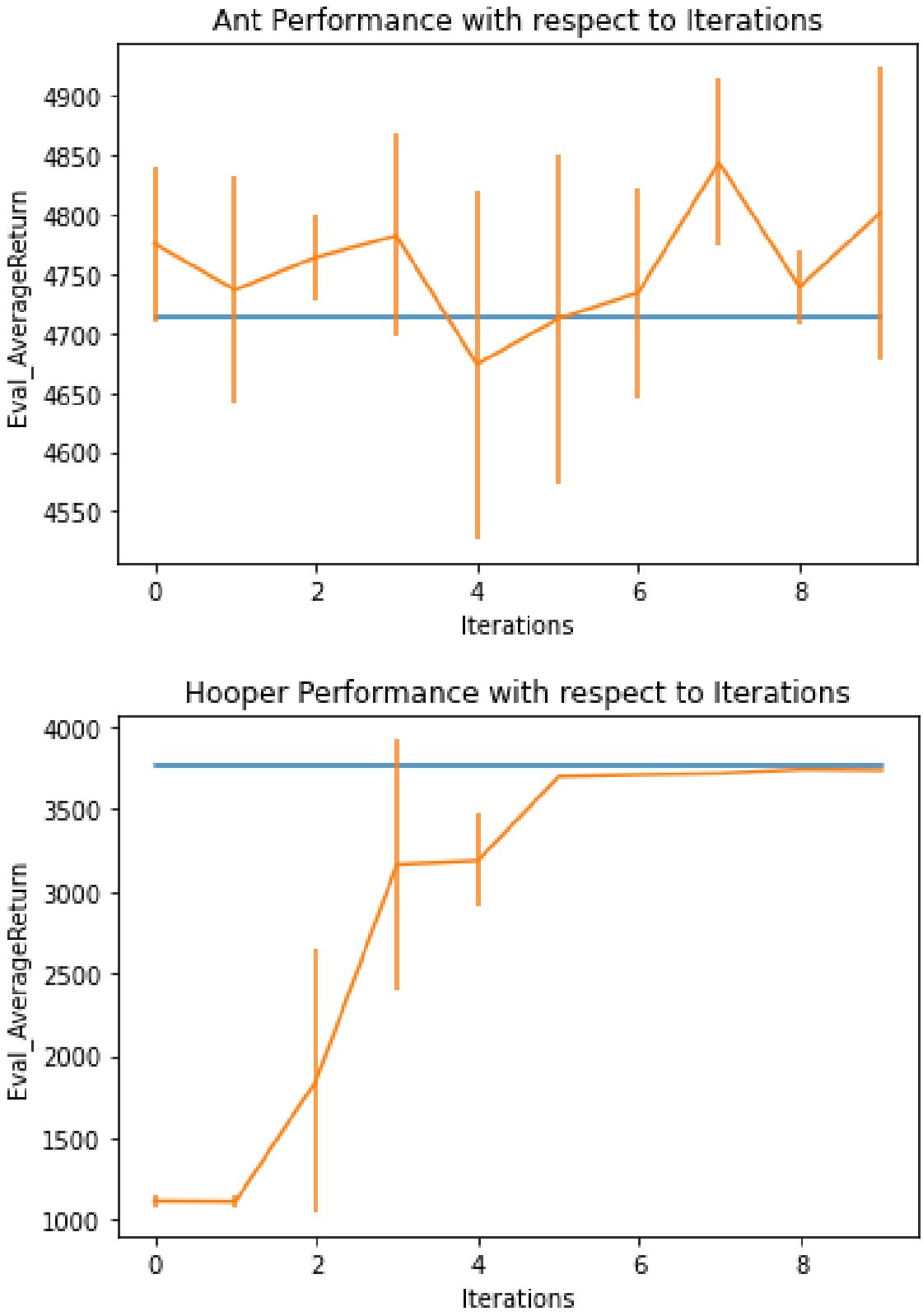

Figure2. Two Environments' performance with respect to Iterations with expert performance.