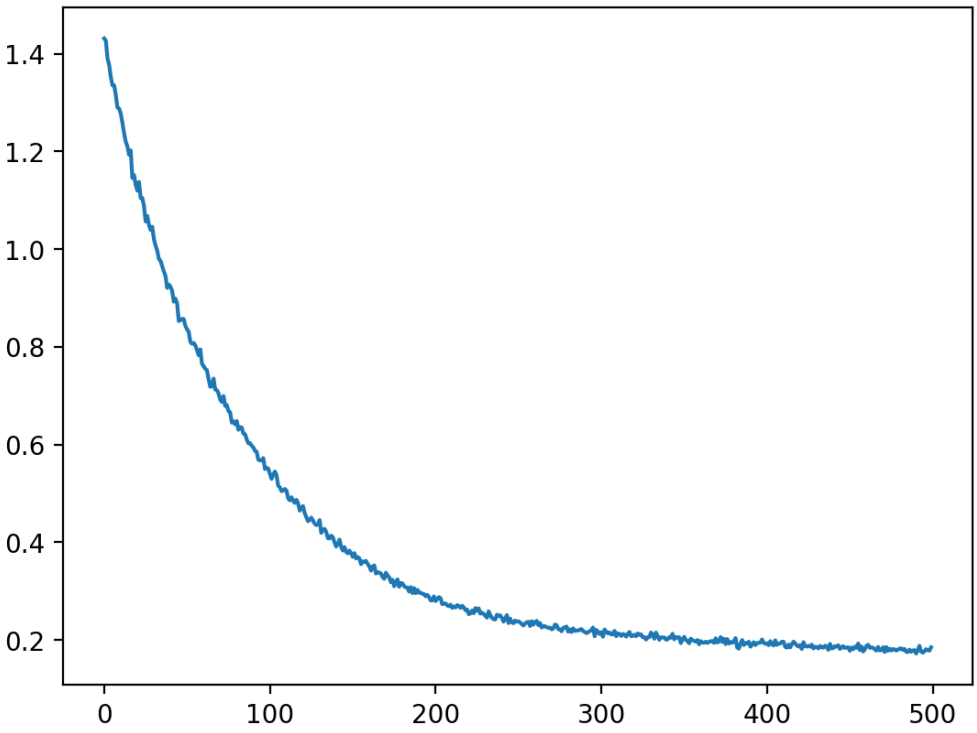
HW4 Report

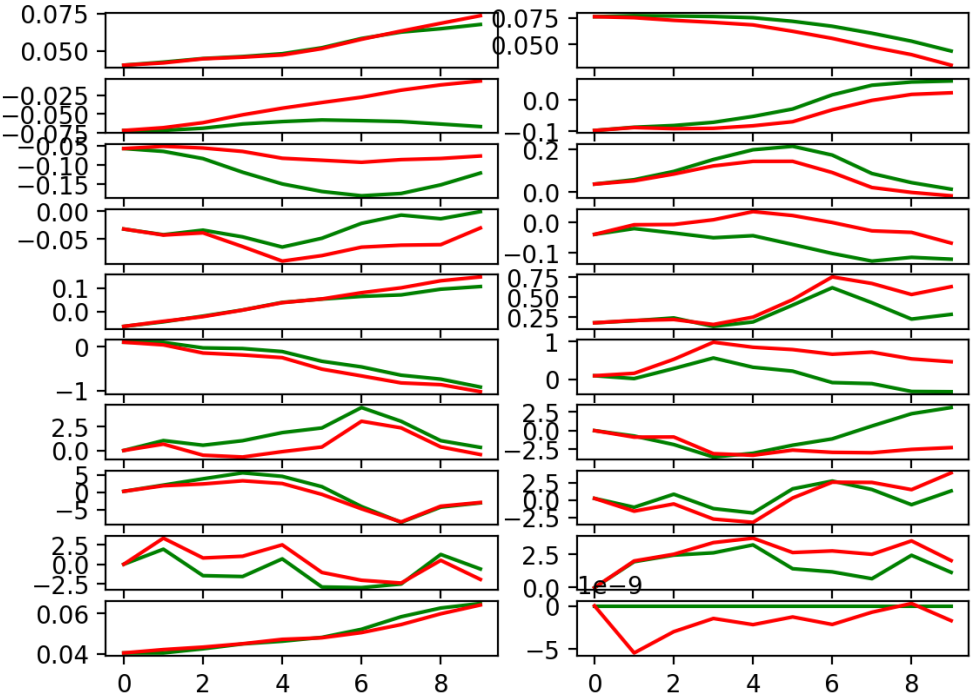
P1: Neural Network Dynamics Model for Cheetah

1 layer 32 neurons/layer 500 training steps per iteration

```
python cs285/scripts/run_hw4_mb.py --exp_name cheetah_n500_arch1x32 --env_name
cheetah-cs285-v0 --add_sl_noise --n_iter 1 --batch_size_initial 20000 --
num_agent_train_steps_per_iter 500 --n_layers 1 --size 32 --scalar_log_freq -1
--video_log_freq -1
```

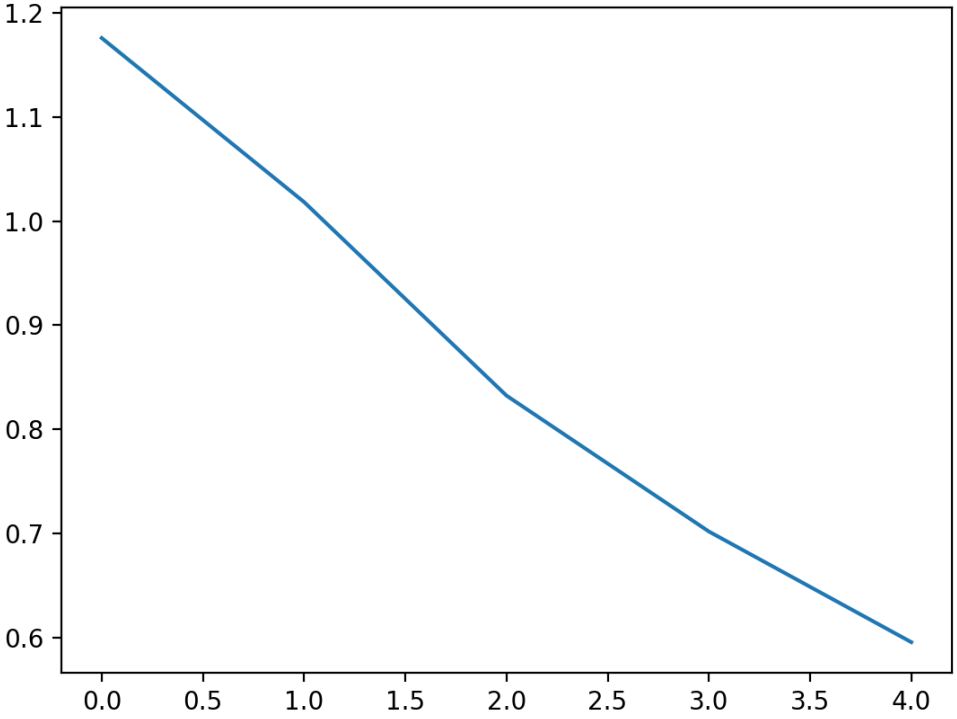


MPE: 0.7646447

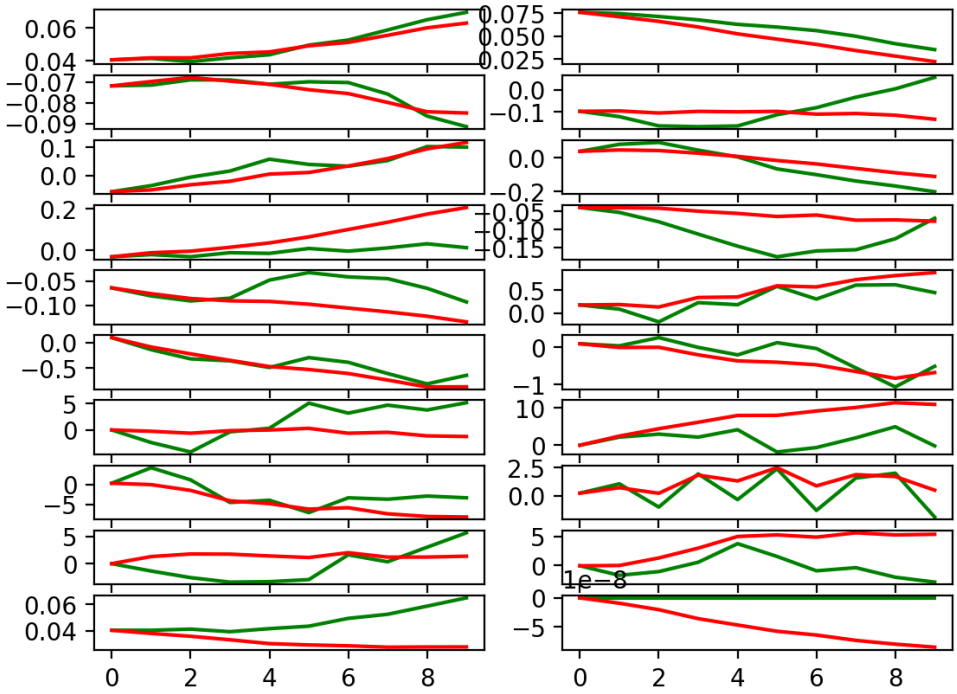


2 layer 250 neurons 5 training steps per iteration

```
python cs285/neurons/layer/run_hw4_mb.py --exp_name cheetah_n5_arch2x250 --
env_name cheetah-cs285-v0 --add_sl_noise --n_iter 1 --batch_size_initial 20000
--num_agent_train_steps_per_iter 5 --n_layers 2 --size 250 --scalar_log_freq -1
--video_log_freq -1
```

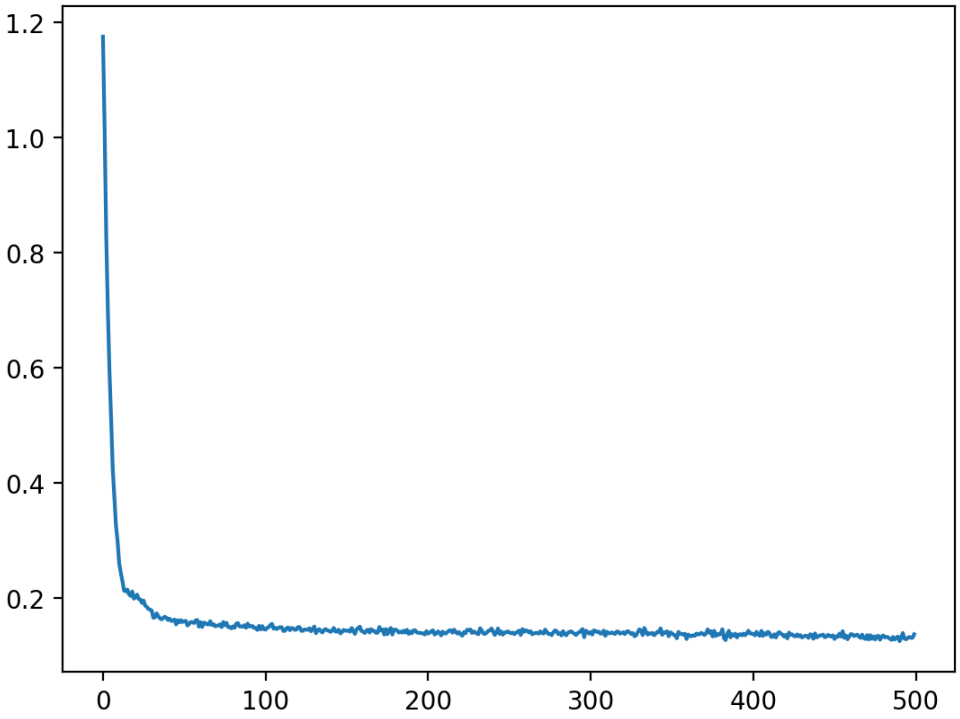


MPE: 5.010764

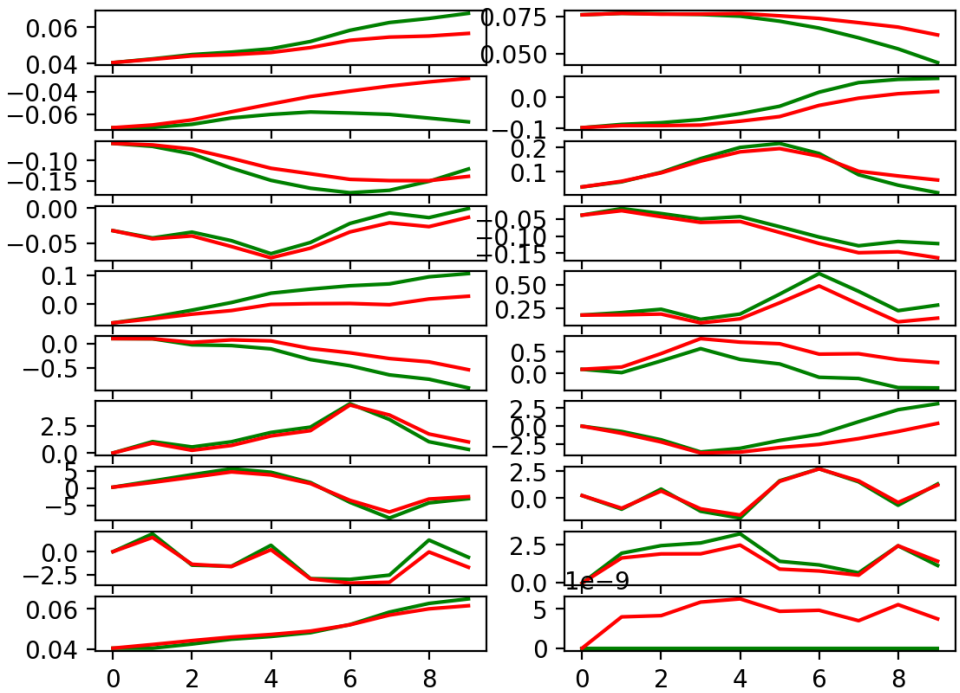


2 layer 250 neurons/layer 500 training steps per iteration

```
python cs285/scripts/run_hw4_mb.py --exp_name cheetah_n500_arch2x250 --env_name
cheetah-cs285-v0 --add_sl_noise --n_iter 1 --batch_size_initial 20000 --
num_agent_train_steps_per_iter 500 --n_layers 2 --size 250 --scalar_log_freq -1
--video_log_freq -1
```



MPE: 0.2010182



P2: Model based action selection

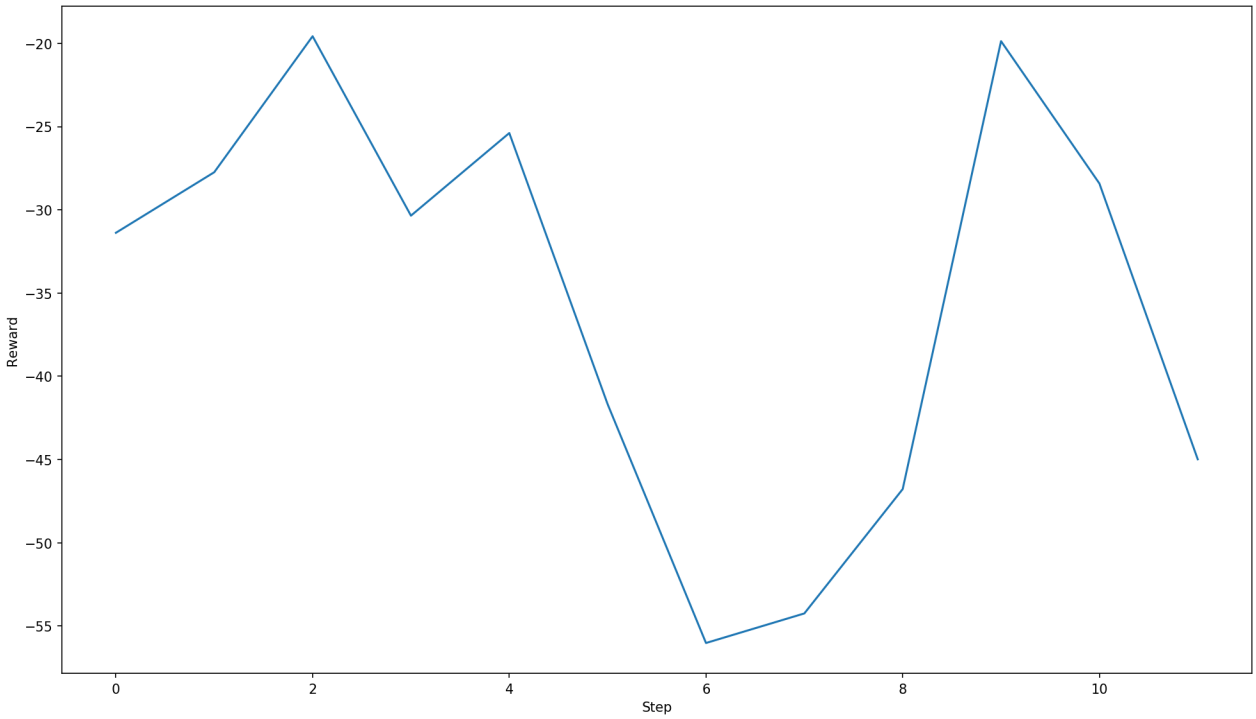
```
python cs285/scripts/run_hw4_mb.py --exp_name obstacles_singleiteration
--env_name obstacles-cs285-v0 --add_sl_noise --
num_agent_train_steps_per_iter 20 --n_iter 1 --batch_size_initial
5000 --batch_size 1000 --mpc_horizon 10
```

Train_AverageReturn(Random policy): -163.49
Eval_AverageReturn(Model based policy): -47.02

P3: MBRL with on-policy data collection and iterative model training

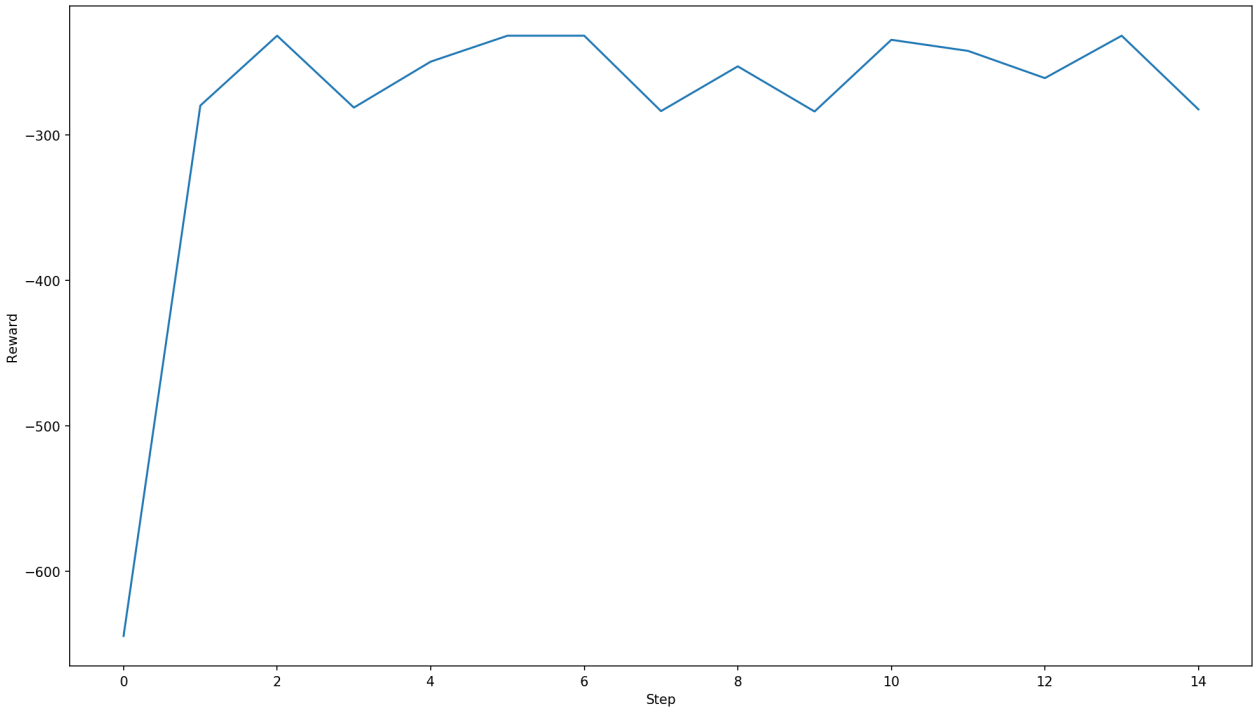
Obstacles

```
python cs285/scripts/run_hw4_mb.py --exp_name obstacles --env_name obstacles-
cs285-v0 --add_sl_noise --num_agent_train_steps_per_iter 20 --batch_size_initial
5000 --batch_size 1000 --mpc_horizon 10 --n_iter 12 --video_log_freq -1
```



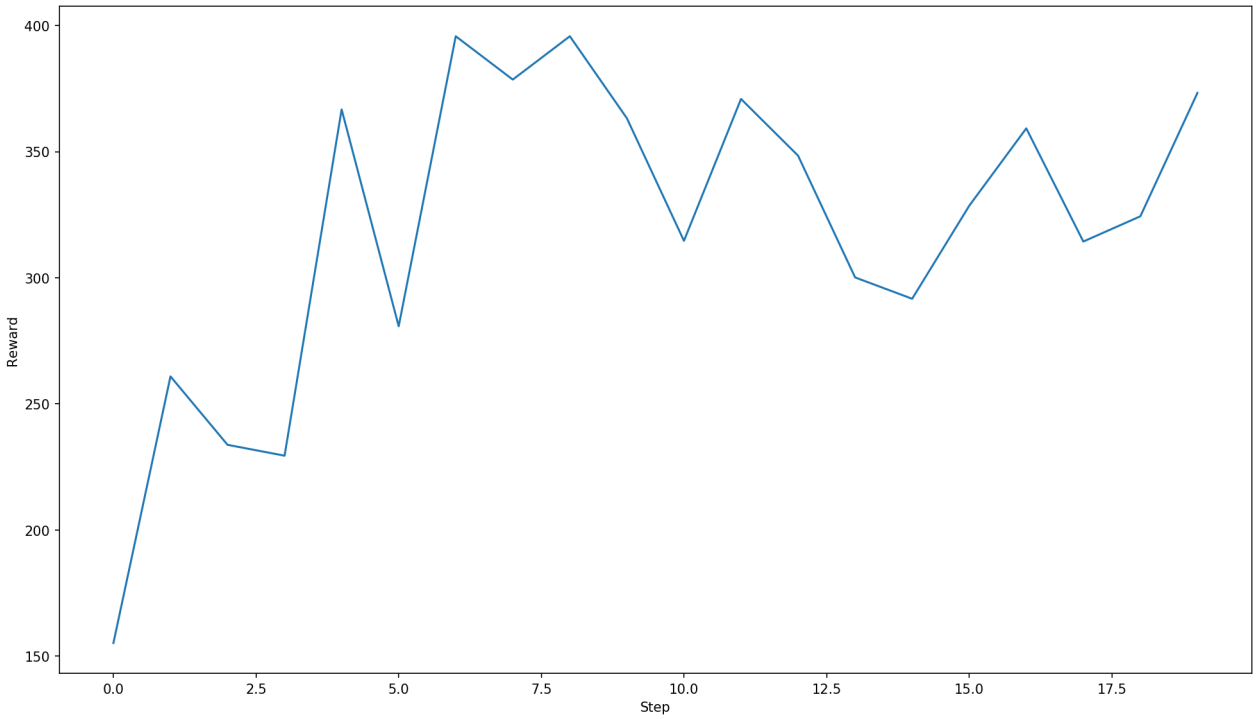
Reacher

```
python cs285/scripts/run_hw4_mb.py --exp_name reacher --env_name reacher-cs285-v0 --add_sl_noise --mpc_horizon 10 --num_agent_train_steps_per_iter 1000 --batch_size_initial 5000 --batch_size 5000 --n_iter 15 --video_log_freq -1
```



Cheetah

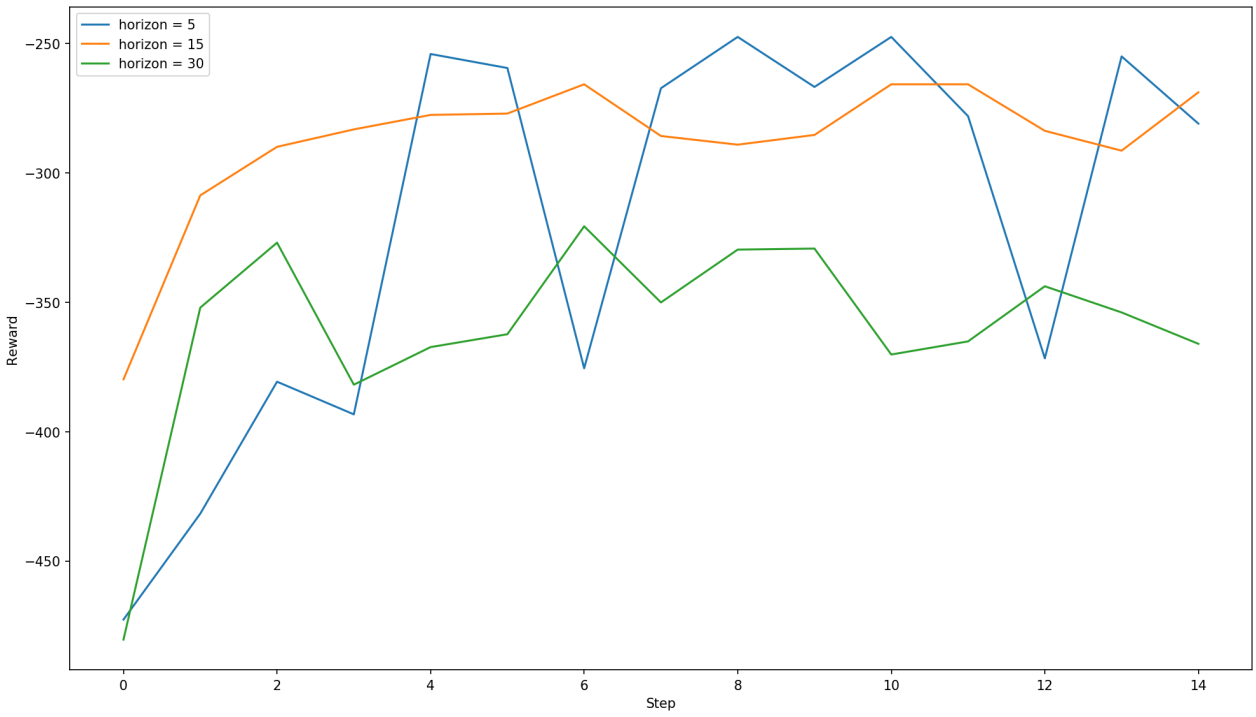
```
python cs285/scripts/run_hw4_mb.py --exp_name cheetah --env_name cheetah-cs285-v0 --mpc_horizon 15 --add_sl_noise --num_agent_train_steps_per_iter 1500 --batch_size_initial 5000 --batch_size 5000 --n_iter 20 --video_log_freq -1
```



P4: Hyperparameters

Horizon

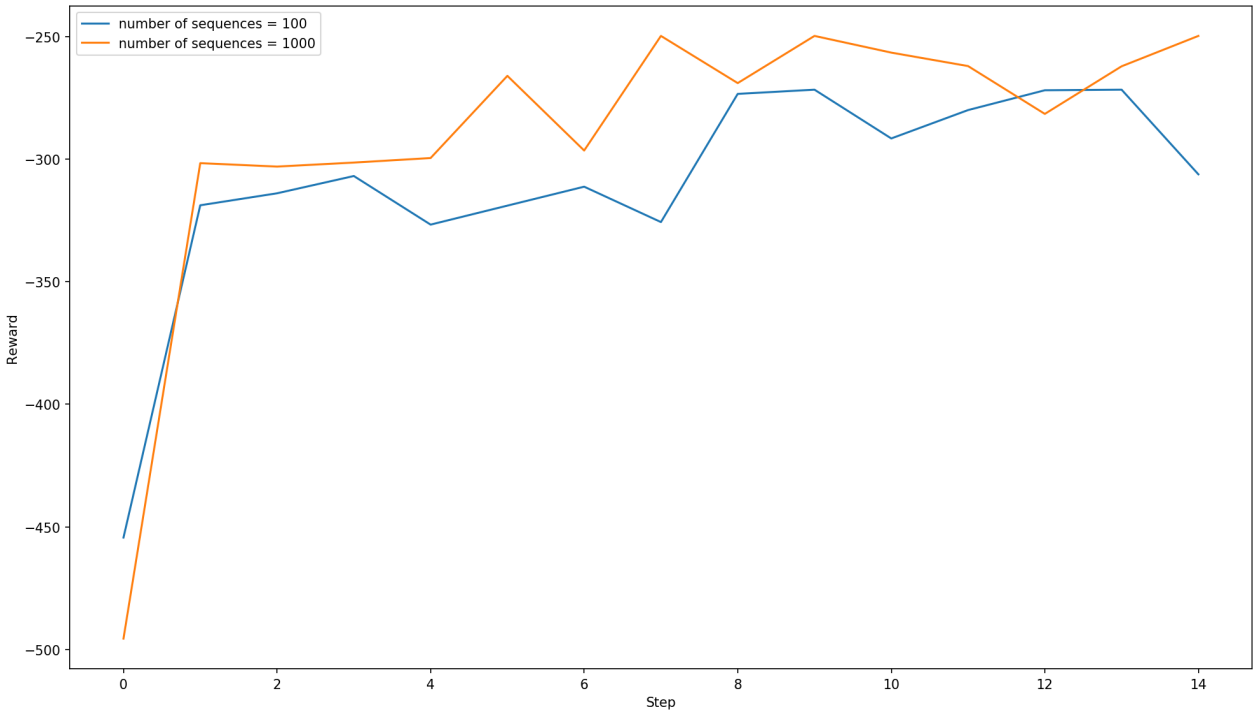
```
python cs285/scripts/run_hw4_mb.py --exp_name q5_reacher_horizon5 --
env_name reacher-cs285-v0 --add_sl_noise --mpc_horizon 5 --
num_agent_train_steps_per_iter 1000 --batch_size 800 --n_iter 15
python cs285/scripts/run_hw4_mb.py --exp_name q5_reacher_horizon15 --
env_name reacher-cs285-v0 --add_sl_noise --mpc_horizon 15 --
num_agent_train_steps_per_iter 1000 --batch_size 800 --n_iter 15
python cs285/scripts/run_hw4_mb.py --exp_name q5_reacher_horizon30 --
env_name reacher-cs285-v0 --add_sl_noise --mpc_horizon 30 --
num_agent_train_steps_per_iter 1000 --batch_size 800 --n_iter 15
```



Increasing the horizon decreases variance in performance but does not necessarily increase rewards as one might expect.

Number of action sequences to choose from

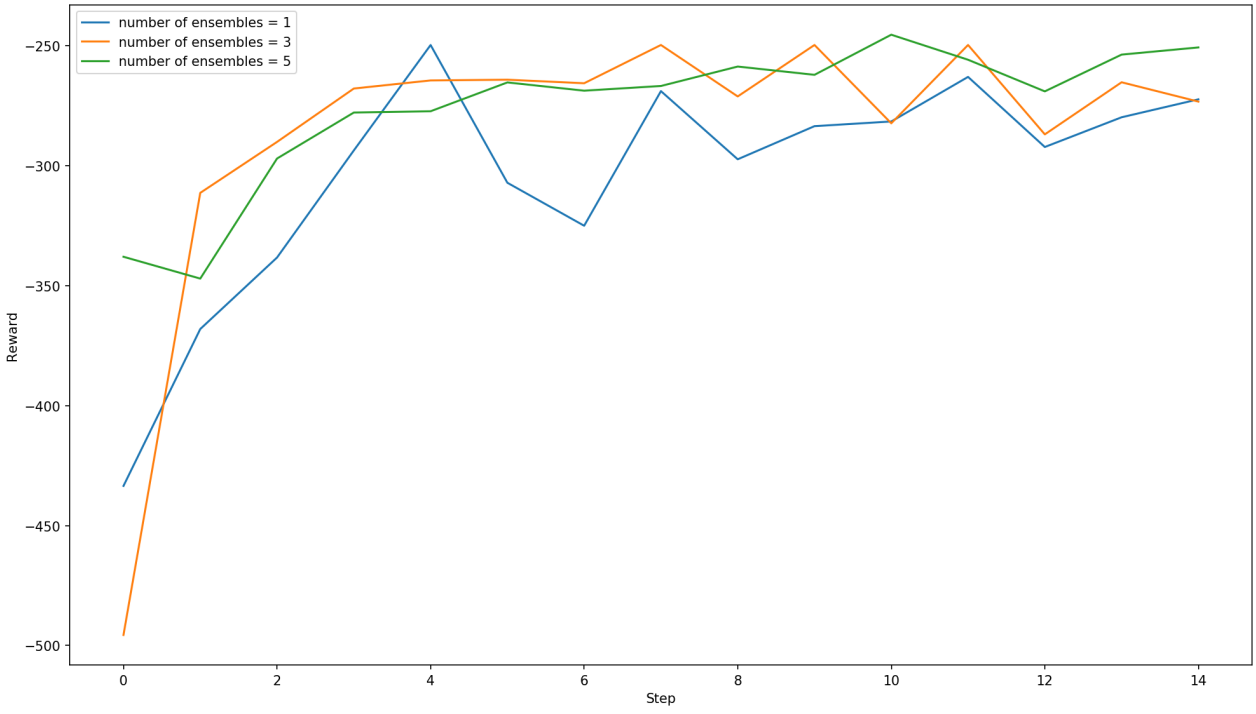
```
python cs285/scripts/run_hw4_mb.py --exp_name q5_reacher_numseq100 --
env_name reacher-cs285-v0 --add_sl_noise --mpc_horizon 10 --
num_agent_train_steps_per_iter 1000 --batch_size 800 --n_iter 15 --
mpc_num_action_sequences 100
python cs285/scripts/run_hw4_mb.py --exp_name q5_reacher_numseq1000 --
env_name reacher-cs285-v0 --add_sl_noise --mpc_horizon 10 --
num_agent_train_steps_per_iter 1000 --batch_size 800 --n_iter 15 --
mpc_num_action_sequences 1000
```



Reward increases with the number of action sequences to choose from as expected.

Ensemble size

```
python cs285/scripts/run_hw4_mb.py --exp_name q5_reacher_ensemble1 --
env_name reacher-cs285-v0 --ensemble_size 1 --add_sl_noise --
mpc_horizon 10 --num_agent_train_steps_per_iter 1000 --batch_size
800 --n_iter 15
python cs285/scripts/run_hw4_mb.py --exp_name q5_reacher_ensemble3 --
env_name reacher-cs285-v0 --ensemble_size 3 --add_sl_noise --
mpc_horizon 10 --num_agent_train_steps_per_iter 1000 --batch_size
800 --n_iter 15
python cs285/scripts/run_hw4_mb.py --exp_name q5_reacher_ensemble5 --
env_name reacher-cs285-v0 --ensemble_size 5 --add_sl_noise --
mpc_horizon 10 --num_agent_train_steps_per_iter 1000 --batch_size
800 --n_iter 15
```



Increasing ensemble size increases reward and decreases variance as expected.