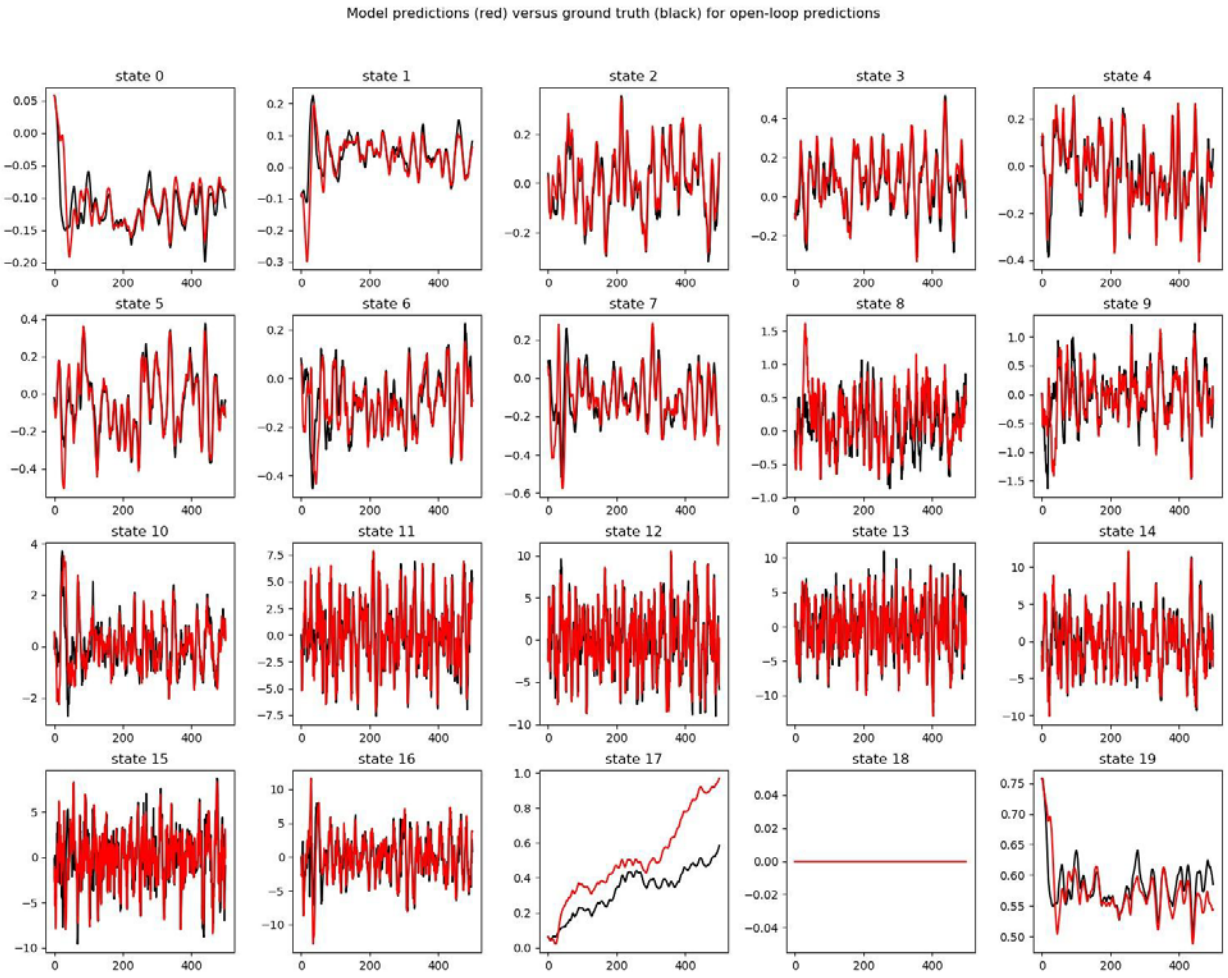
Problem 1

a) Provide a plot of the dynamics model predictions when the predictions are mostly accurate ¶



b) For (a), for which state dimension are the predictions the most inaccurate? Give a possible reason why the predictions are inaccurate

- Prediction for state dim-17 is the most inaccurate. Since it is a open loop prediction and state 17 has a upward trend, the small prediction errors will compound and the accumulation of them will make prediction deviate a lot from the actual states.

In [ ]:

Problem 2

Provide the ReturnAvg and ReturnStd for the random policy and for your model-based controller trained on the randomly gathered data.

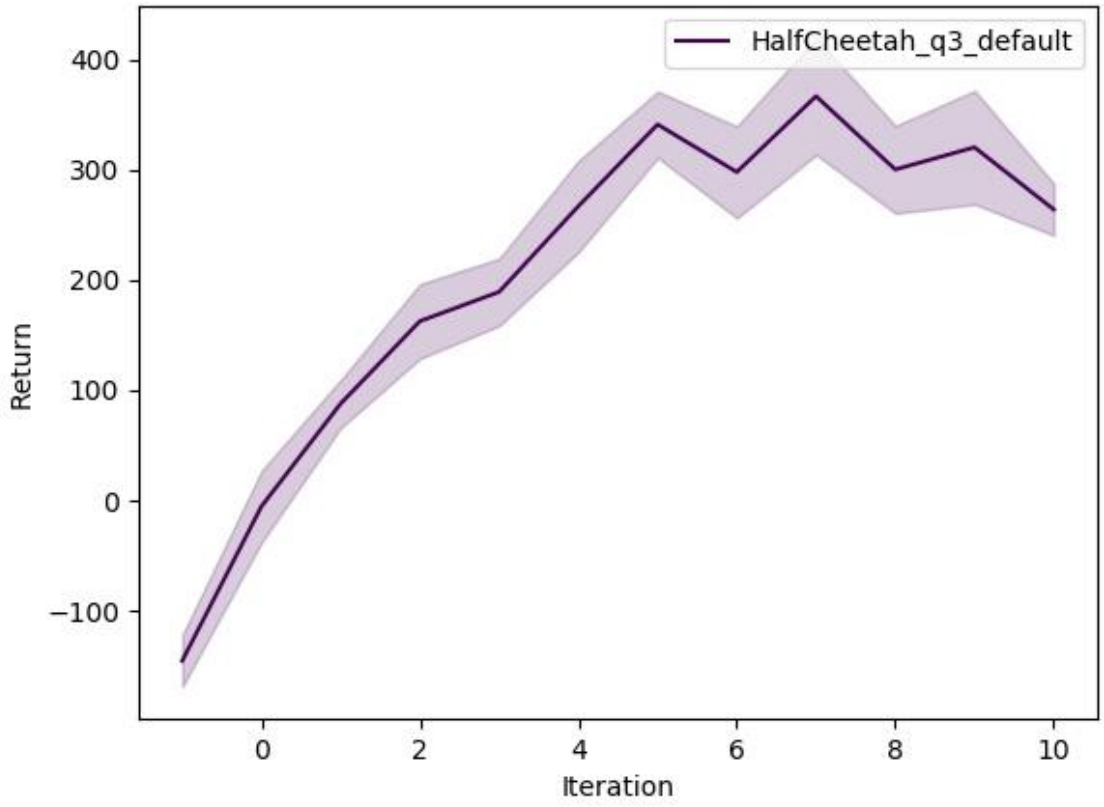
- 0 is random policy
- 1 is model based controller

	ReturnAv	ReturnStd	ReturnMi	ReturnMa	TrainingLc	TrainingLossFinal	
0	-135.907	37.13628	-216.31	-78.6554			
1	7.304251	22.37661	-29.6272	50.33616	1.064099	0.026369	

In [ ]:

Problem 3a

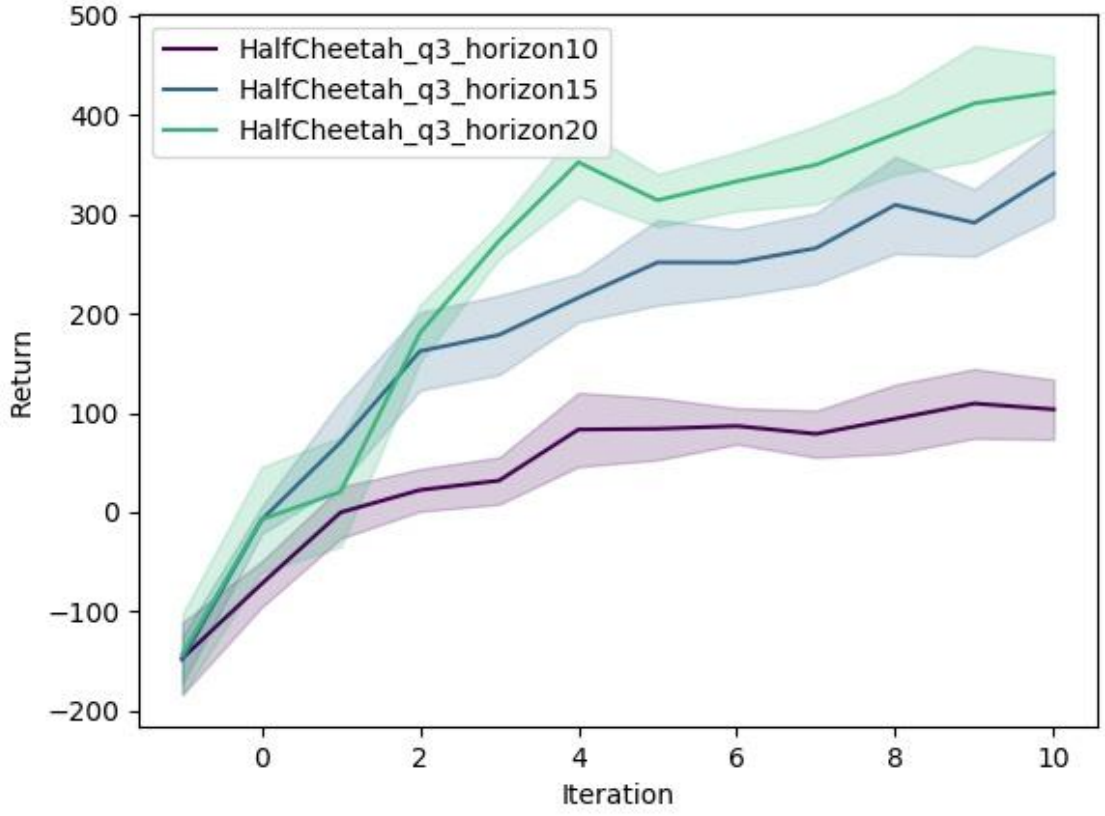
Plot of the returns versus iteration when running model-based reinforcement learning.



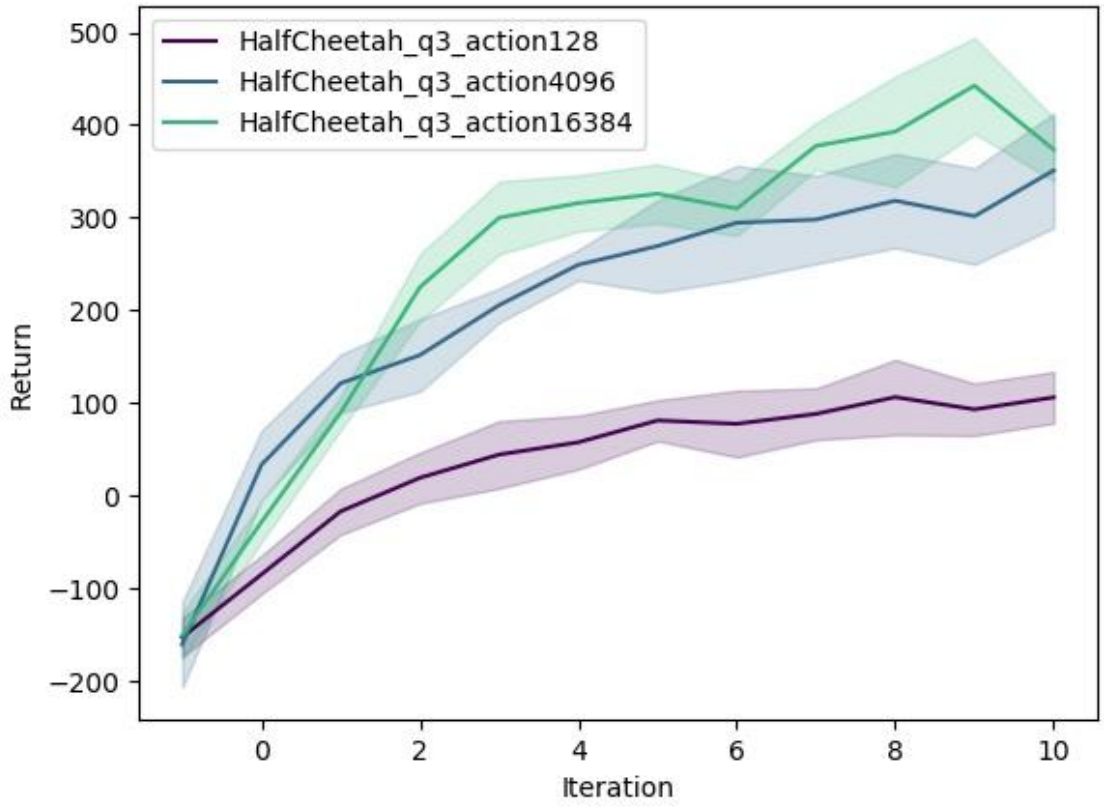
In [ ]:

Problem 3b

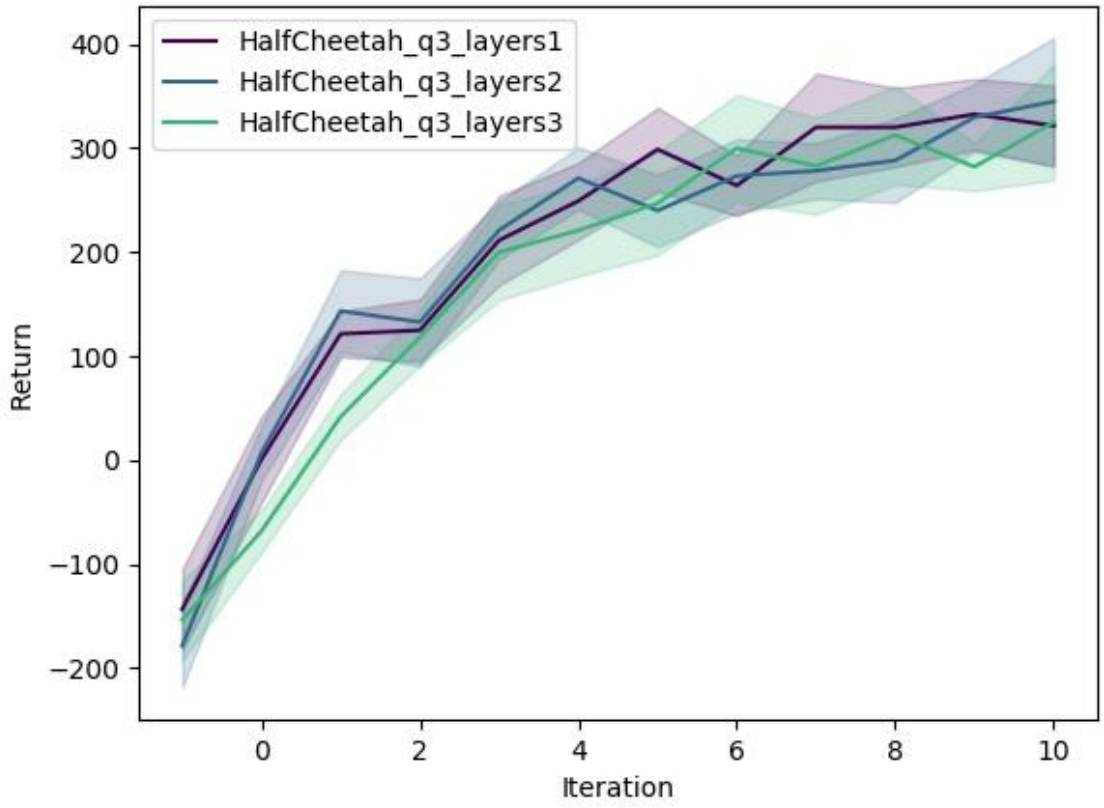
a) Plot comparing performance when varying the MPC horizon.



b) Plot comparing performance when varying the number of randomly sampled action sequences used for planning.



c) Plot comparing performance when varying the number of neural network layers for the learned dynamics model.


HalfCheetah_q3_layers1
HalfCheetah_q3_layers2
HalfCheetah_q3_layers3
Return
Iteration
400
300
200
100
0
−100
−200
0
2
4
6
8
10

In [ ]:

Extra Bonus

Plot comparing performance of either CEM to random for action selection.

- code is implemented in `model_based_policy.py line 204-line 233`

	TrainingLc	TrainingLc	ReturnAv	ReturnStd	ReturnMi	ReturnMax
0	1.091793	0.032194	-37.4549	5.17852	-45.853	-29.5091

In [ ]: