CS 294 - Homework 4

Hatim Ezbakhe

Problem 1

a)

Model predictions (red) versus ground truth (black) for open-loop predictions
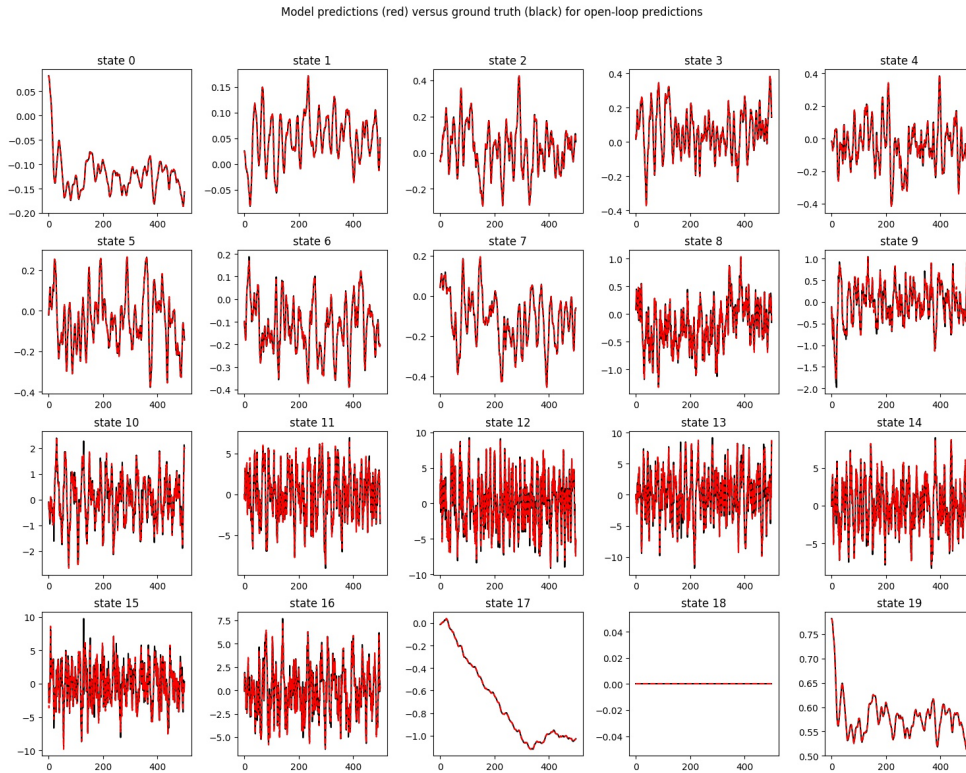
b)

It seems like the predictions are the most inaccurate for the state dimension 15 (although a couple other dimensions are equivalently inaccurate).

My first thought was that it was because the generated random policy did not explore too much that state dimension (all actions randomly taken happen to not give too much variance along that 15th dimension).

So I tried running the code for different seeds, but all gave the same results: the most inacurrate state dimensions are always the same.

A possible explanation for this could be that some state dimension are harder to "get in", making a random policy not exploring too many states along that dimension. For example, we could imagine a car driving on a long road surounded by walls, with only a small hole that gives acces to a perpendicular road: the car actually exploring this side road won't be very likely, and if that road is way smaller than the other one, even when the environment resets there's not too much chance of starting on it. Consequently, the random dataset generated does not have too many different values for this dimension which make the model results poorer on it.
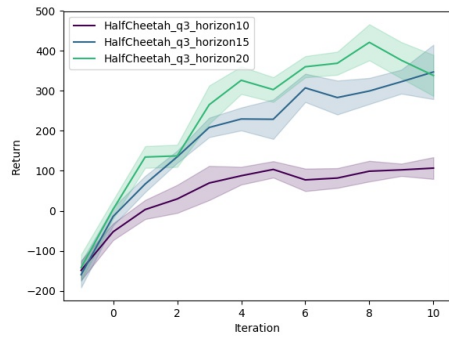
Problem 2

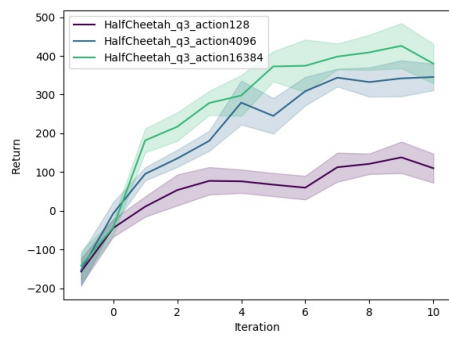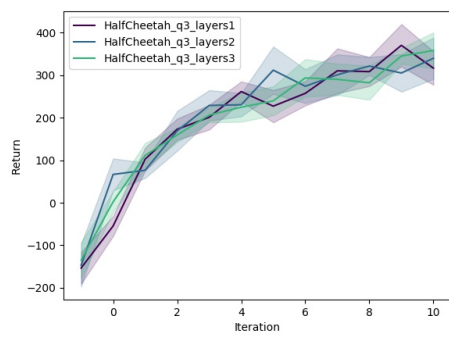| Type	ReturnAvg	ReturnMax	ReturnMin	ReturnStd	TrainingLossFinal	TrainingLossStart
Random Policy	-149.5	-84.6	-195.1	36.1		
Model Based Controller	-11.0	27.0	-70.5	28.5	0.02897	1.08080
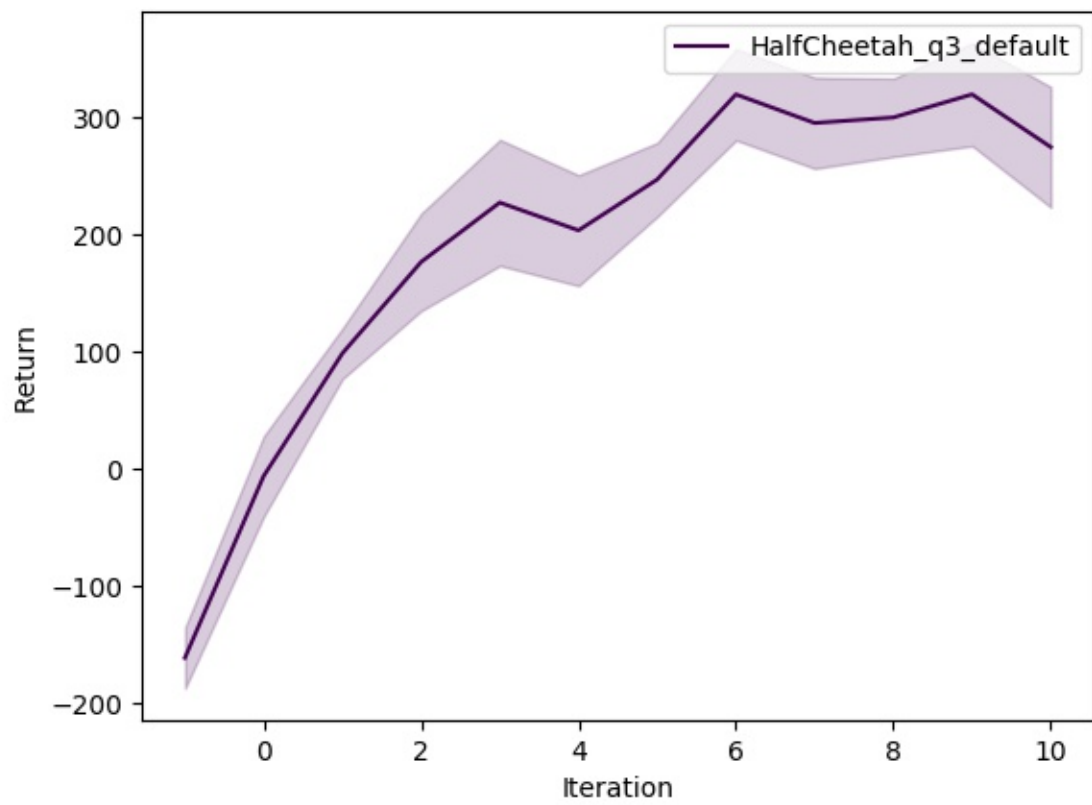
Problem 3a

Problem 3b

a)

b)

c)

Extra Credit a)

The code for the cross entropy method has been implemented in the model_based_policy.py file (lines 165-210). It might be wrong as I I do no have the time to plot the results before the submission deadline.