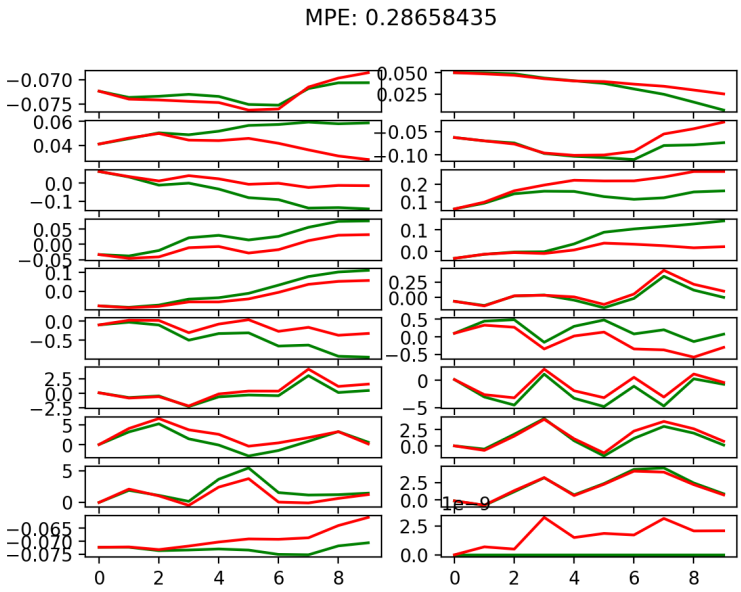
# Assignment 4: Model-Based RL and Exploration

**Andrew ID:** `kkabeer`
**Collaborators:** `asenathi`
**NOTE:** Please do **NOT** change the sizes of the answer blocks or plots.

## 1 Problem 1: Dynamics Model Training

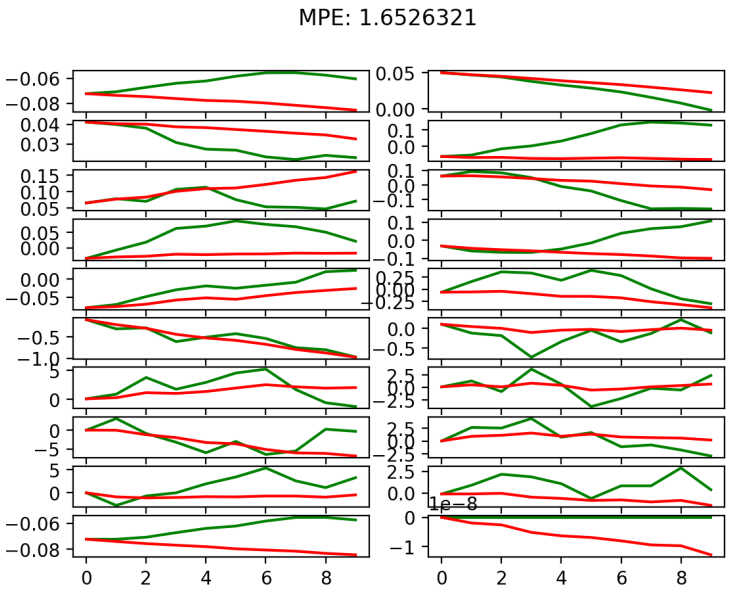
The model having 2 layers with each layer having a size of 250, when trained for 500 iterations, performs best. This is to be expected as the model has more weights and is trained for a longer time, and hence is able to learn a highly non-linear dynamics model.

500 iterations and 1x32 architecture



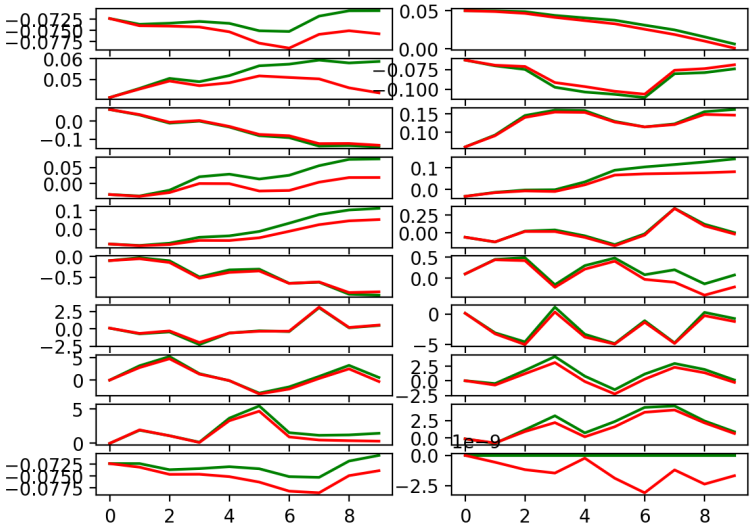
# 1 Problem 1: Dynamics Model Training

5 iterations and 2x250 architecture



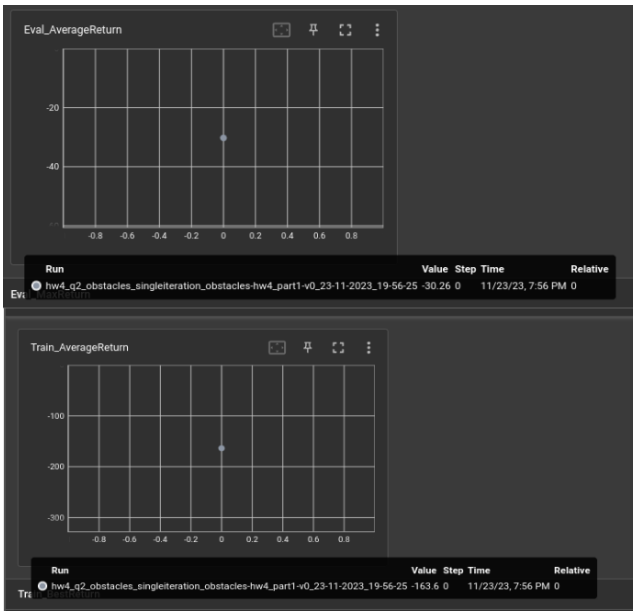

# 1 Problem 1: Dynamics Model Training

500 iterations and 2x250 architecture



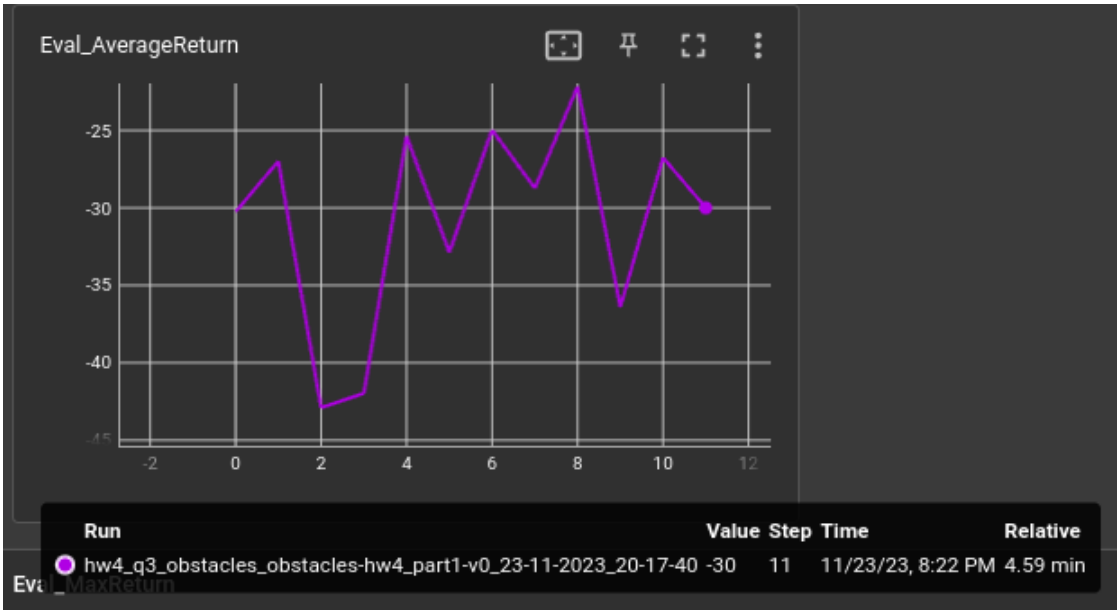
## 2 Problem 2: Action Selection

Eval and Train returns



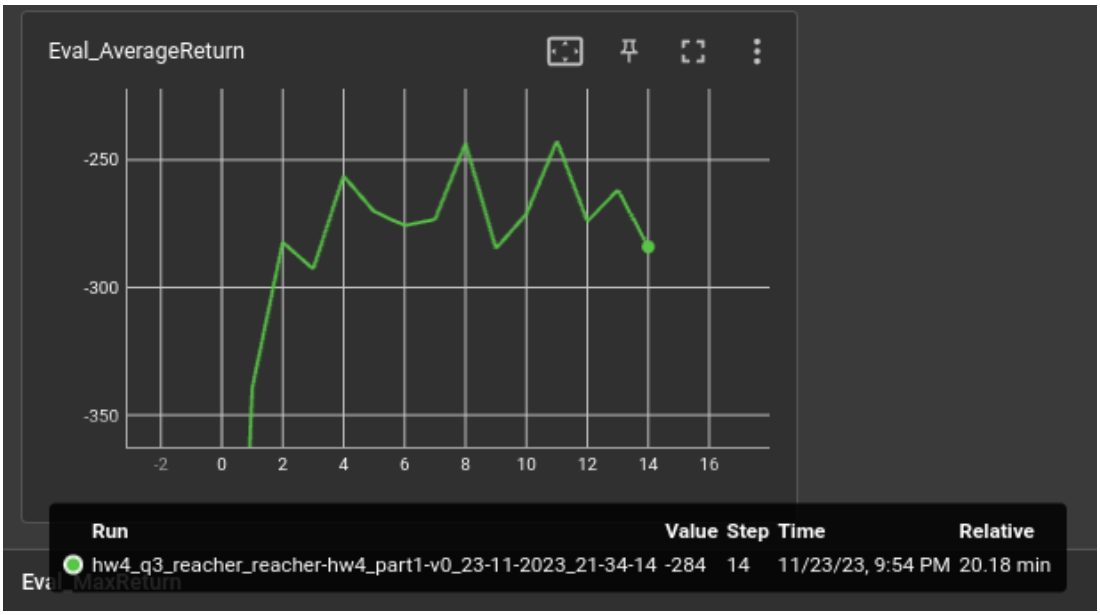

## 3 Problem 3: Iterative Model Training

Eval returns for obstacle



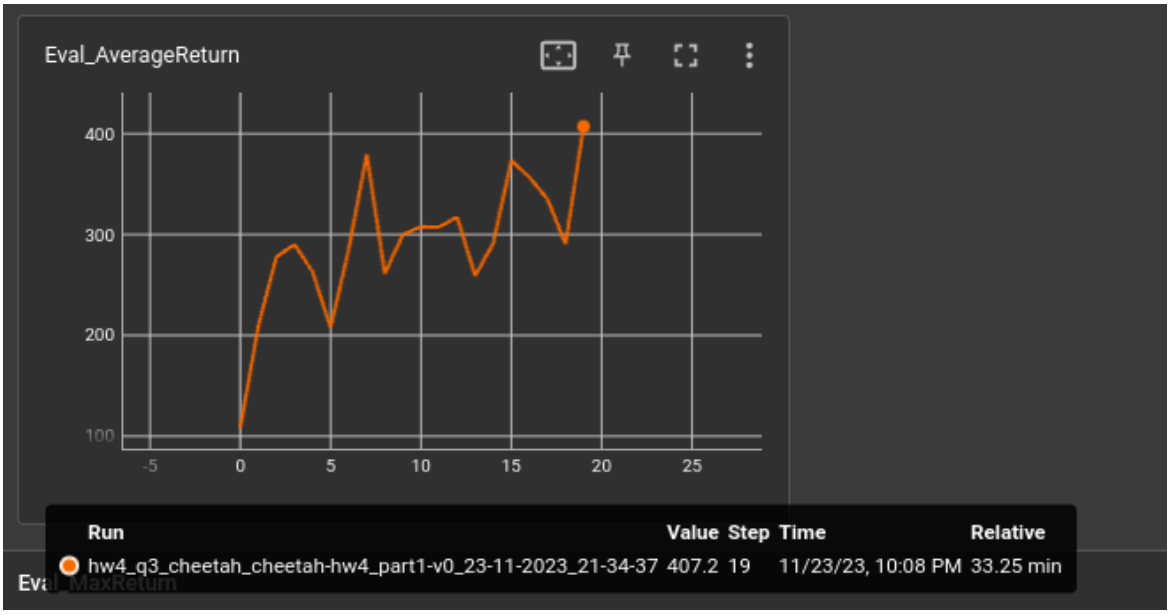
# 4 Problem 3: Iterative Model Training

Eval returns for reacher




# 5 Problem 3: Iterative Model Training

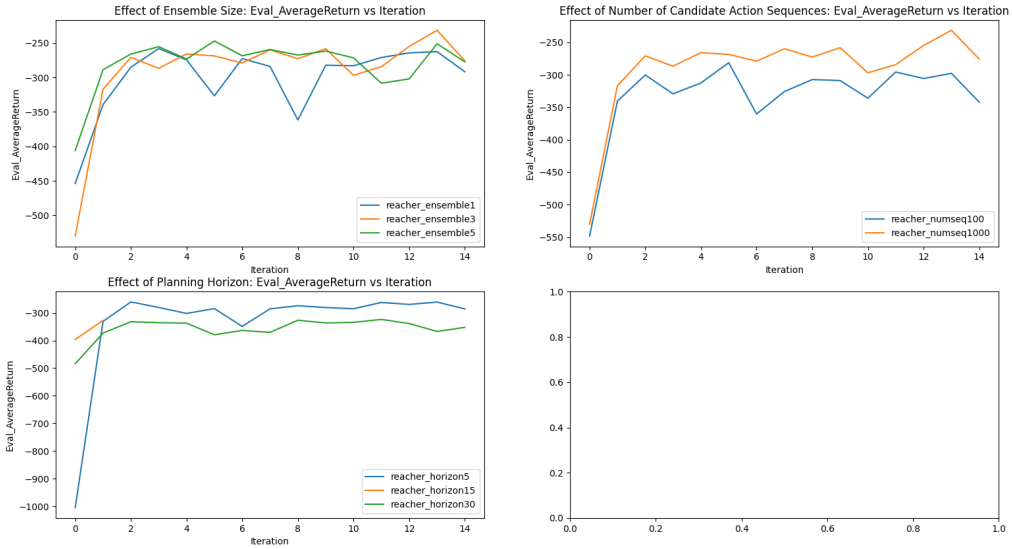
Eval returns for cheetah



# 6 Problem 4: Hyper-parameter Comparison

Ensemble size has a slight effect on the performance - performance increases as ensemble size increases Increasing number of candidate action sequences leads to increased performance Increasing the length of the planning horizon also leads to increase performance

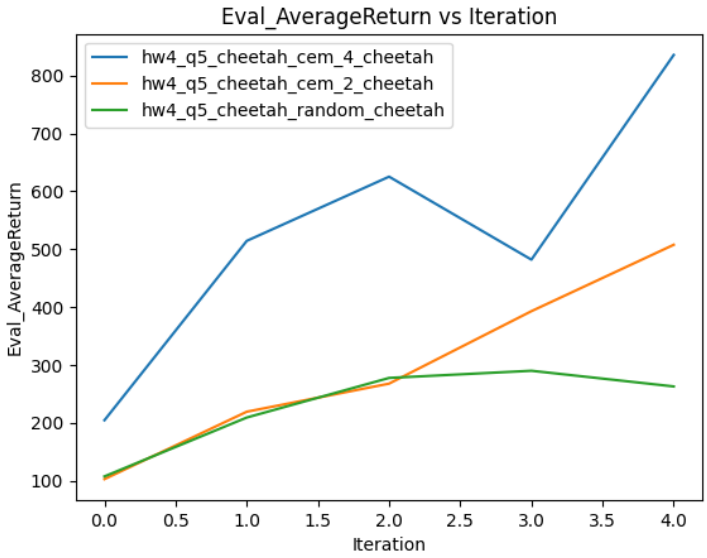
Effects of various hyperparameters on performance




# 7 Problem 5: Hyper-parameter Comparison

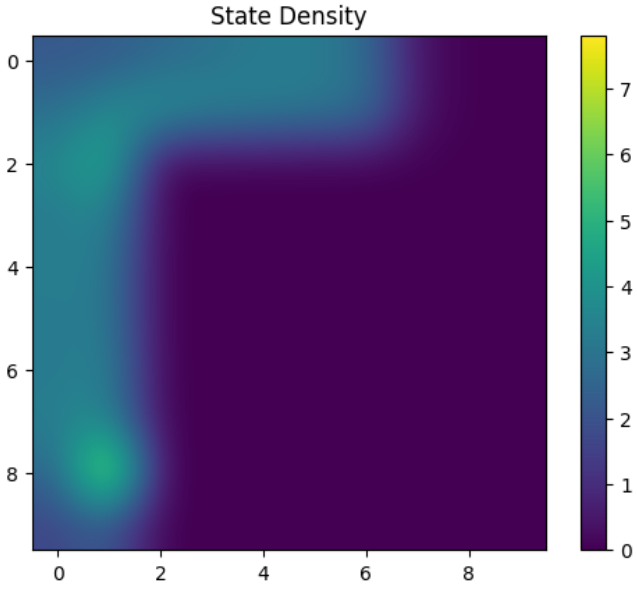
CEM with higher number of sampling iterations performs better. CEM with both 2 and 4 sampling iterations performs better than random sampling.

CEM vs Random Sampling



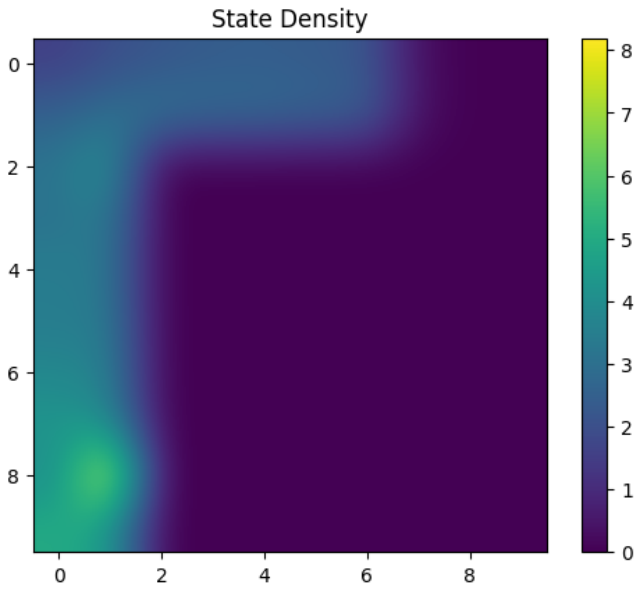

# 8 Problem 6: Exploration

Point Mass Easy with Random Network Distillation



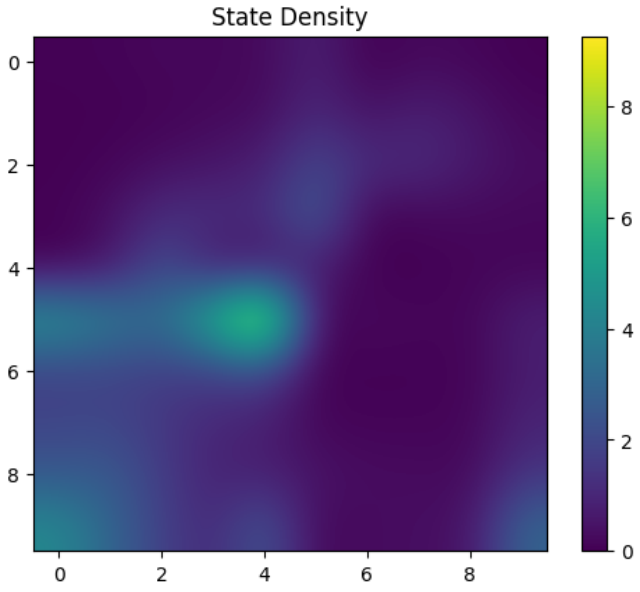
# 9 Problem 6: Exploration

Point Mass Easy with Random Exploration



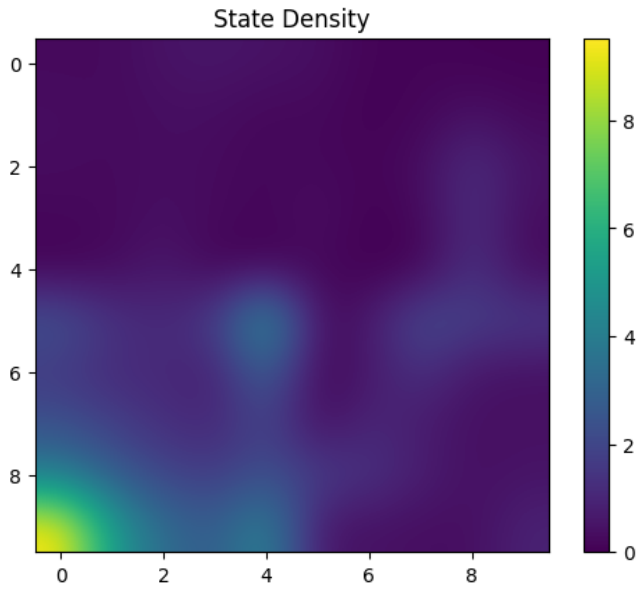

# 10 Problem 6: Exploration

Point Mass Hard with Random Network Distillation



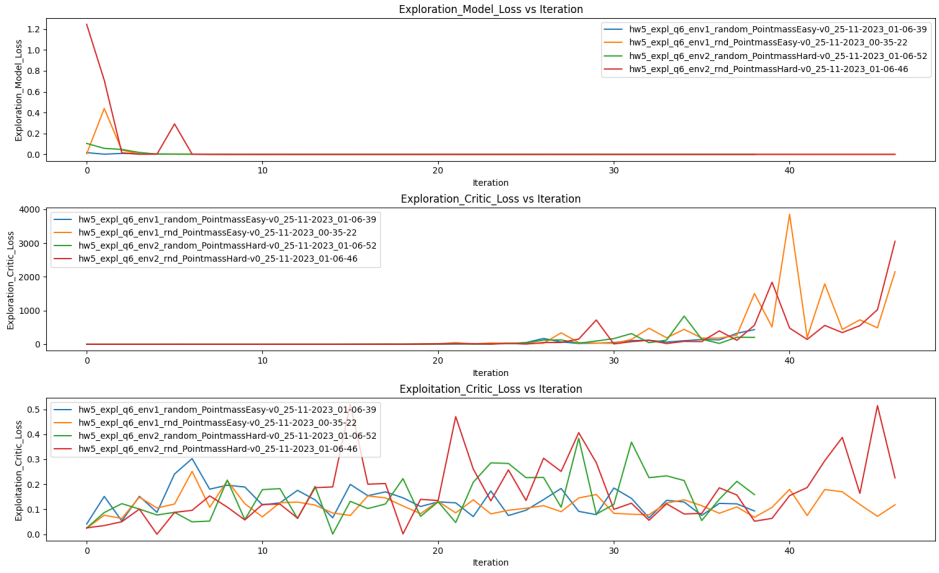
# 11 Problem 6: Exploration

Point Mass Hard with Random Exploration




# 12 Problem 6: Exploration

Model and Critic Performance



## 13 Problem 6: Bonus

Plot 4