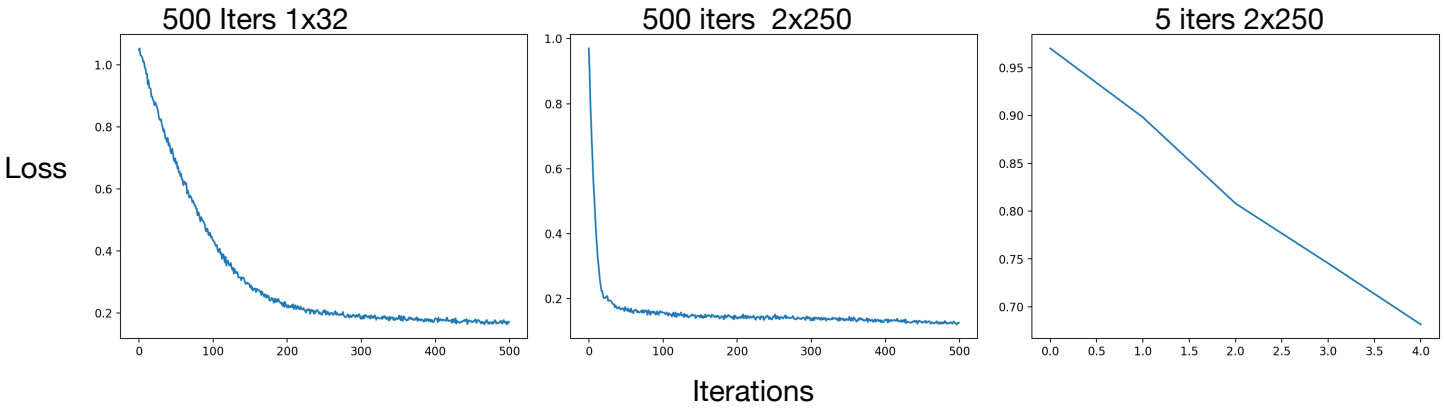
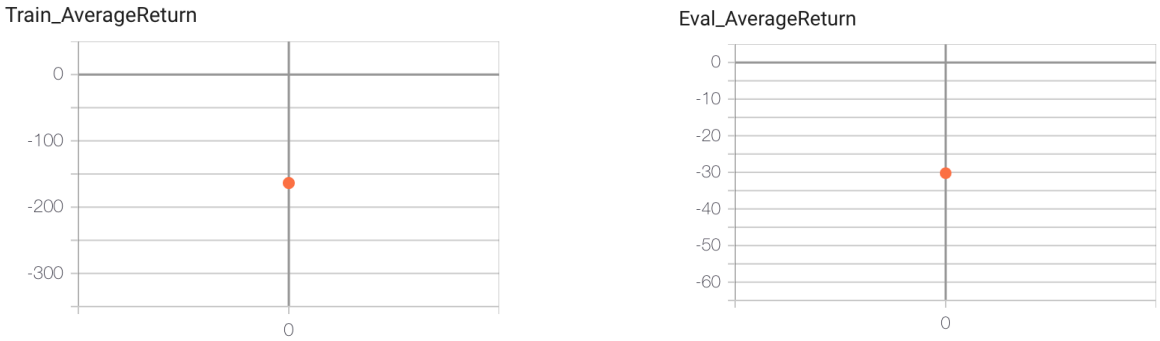
CS285 HW4 Report

Q1

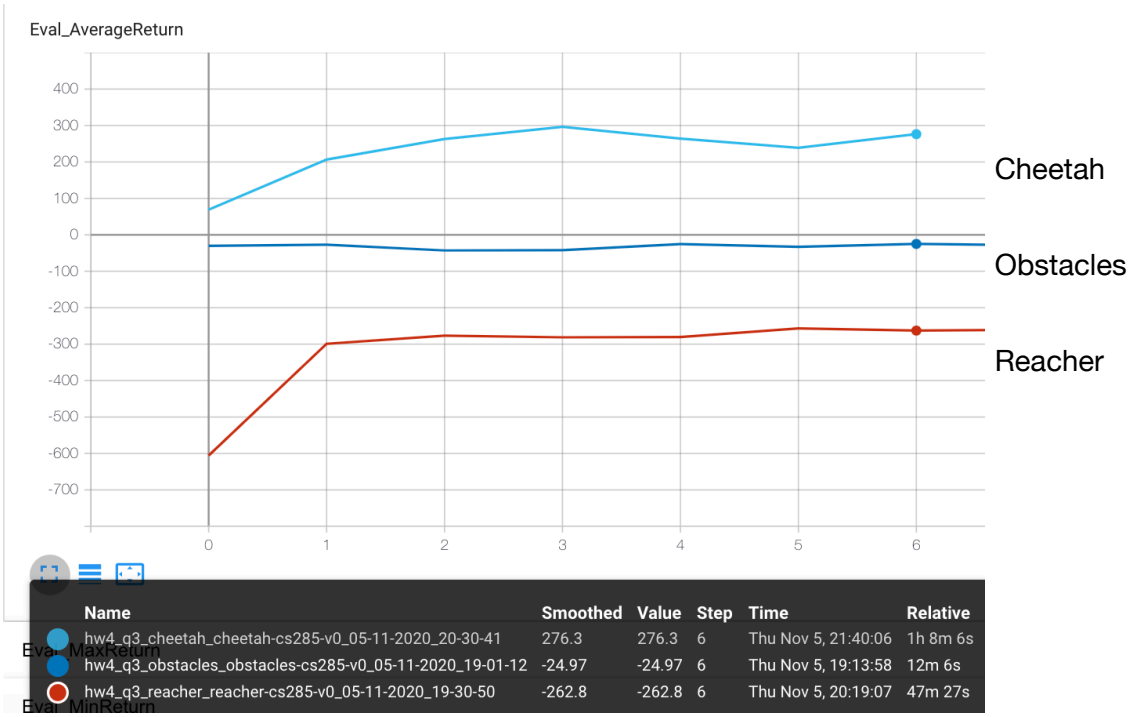


The best fit was found by the 500 iters and 2x250. This is because the model is complex enough to fit the underlying dynamics, yet we trained it for enough iterations that it did not overfit

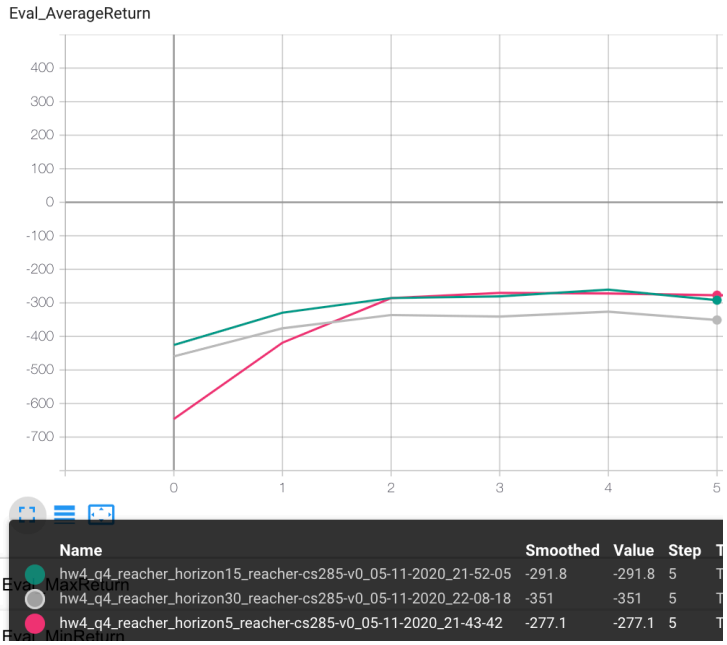
Q2



Q3



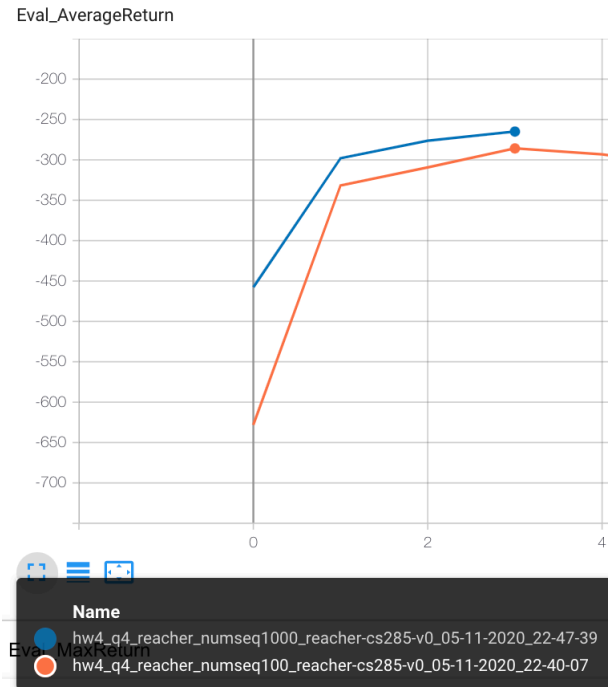
Q4a



I observed that a horizon of 10 was the best performing in this environment (reacher)

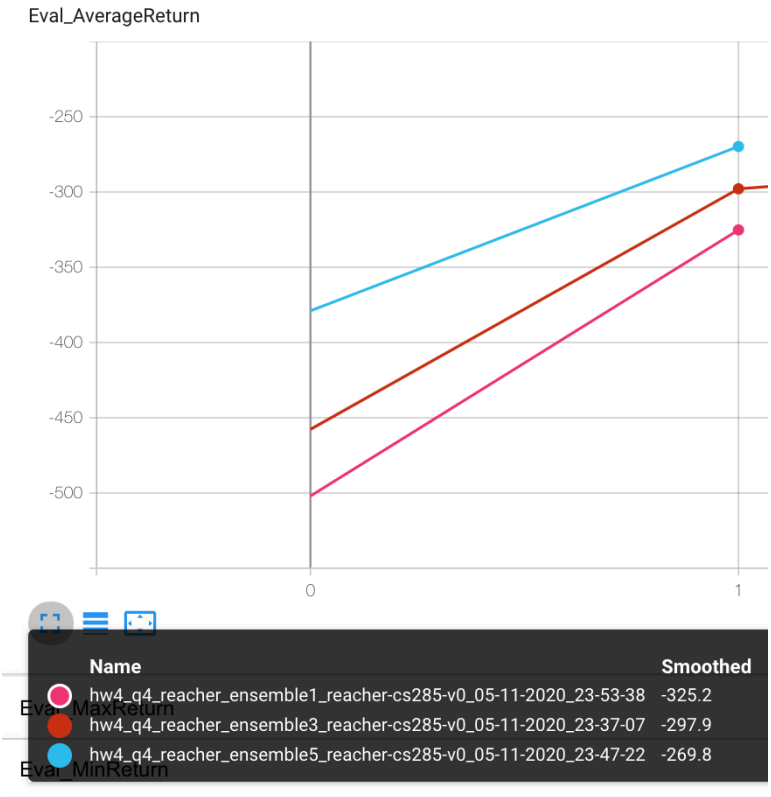
I believe this is because a horizon of 5 is too small to find rewards that are large horizons away from the current position, and a horizon of 15 is so large that the variance is much too high for consistent good results

Q4b



As expected, the nurseq1000 trial performed better in this experiment

Q4c



As can be seen from the graph, and as per our expectations, larger ensemble sizes improve performance