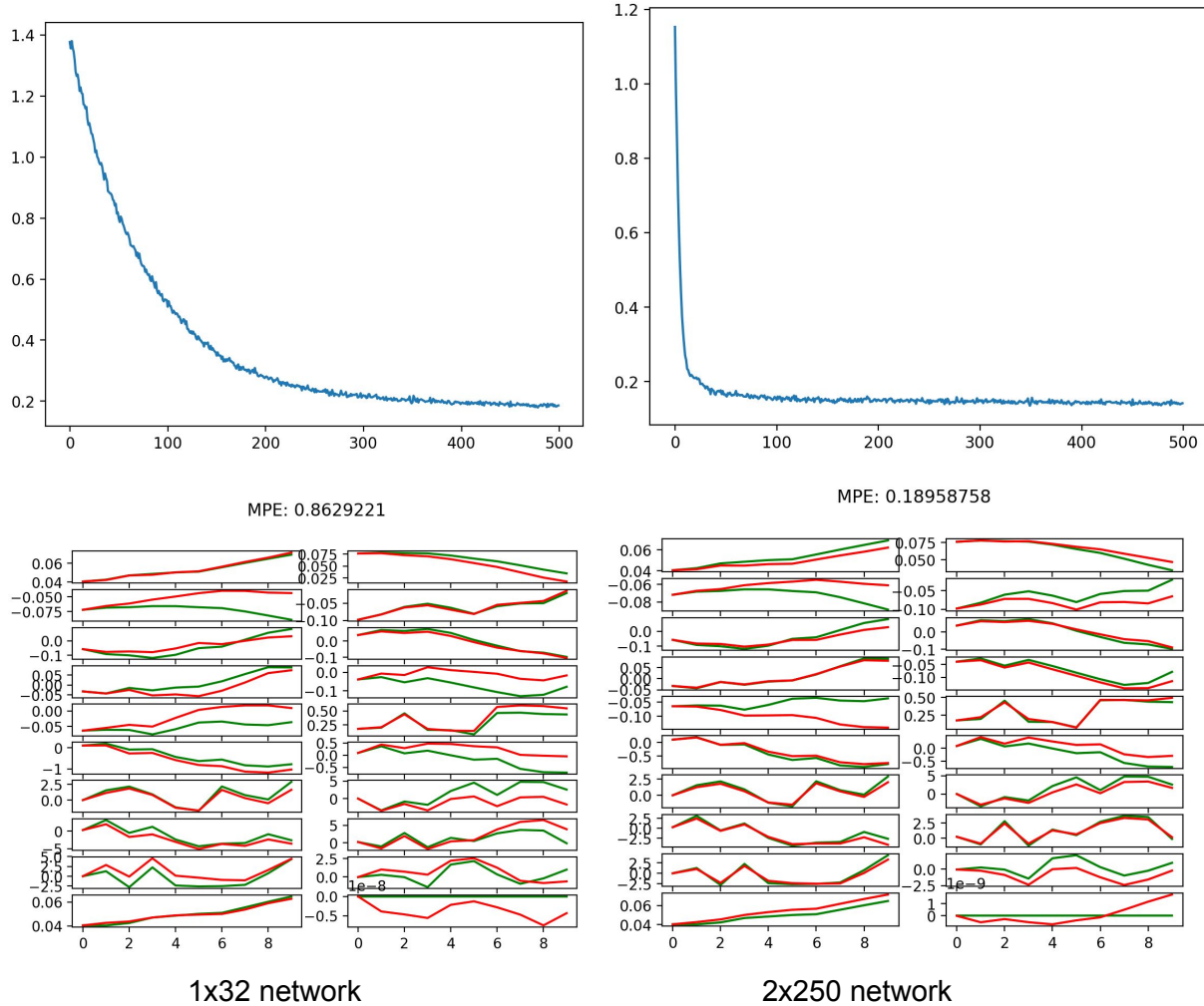


Deep RL hw4 report

(Scripts executed exactly as they appear in the problem formulations if not stated otherwise)
(execute plotnik.ipynb files at logs folders to build the plots)

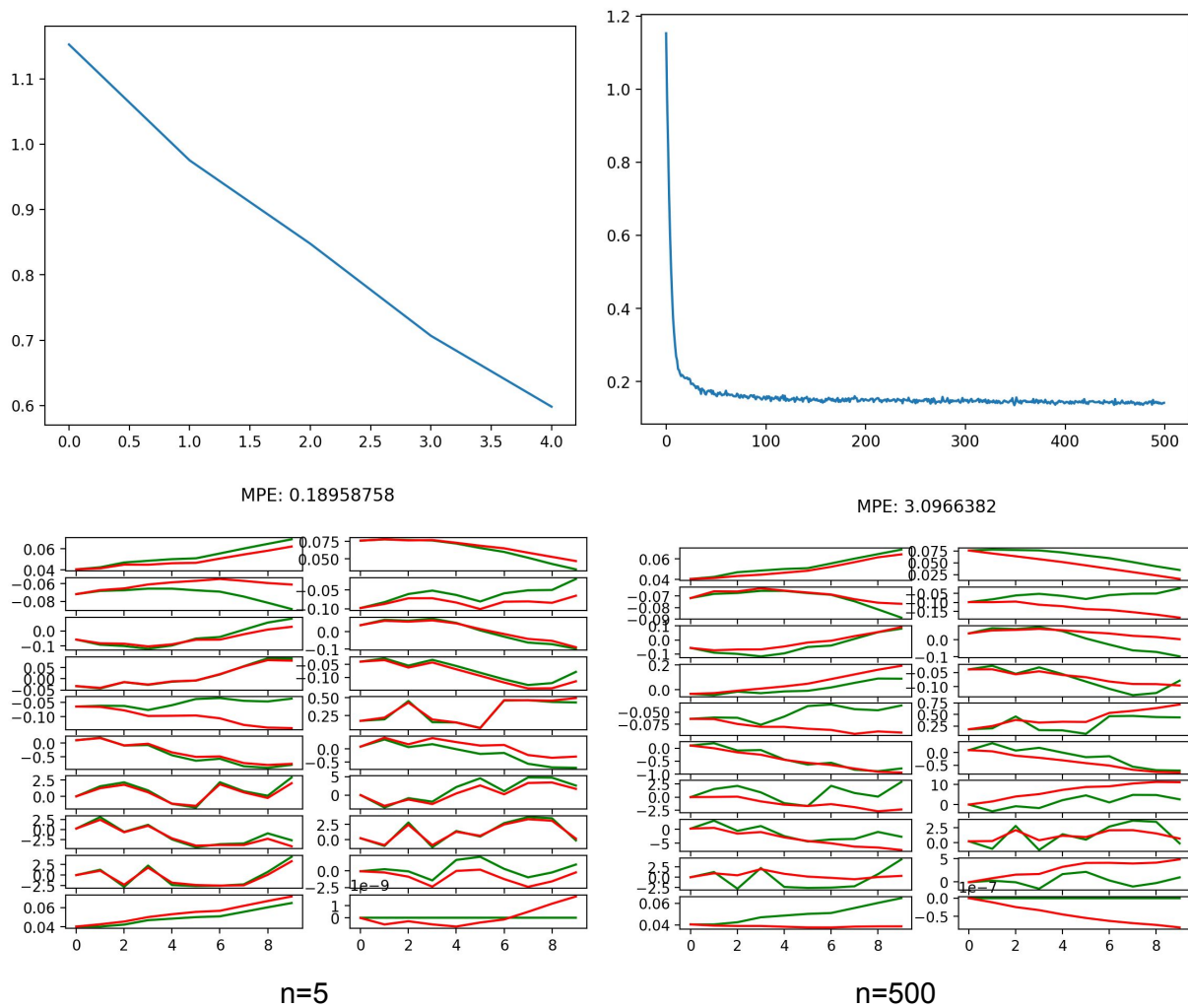
Question 1

- Loss and qualitative model prediction plots for the 2 network sizes:



Qualitative model prediction plots look extremely similar, although the convergence and MPE is much better for the bigger network

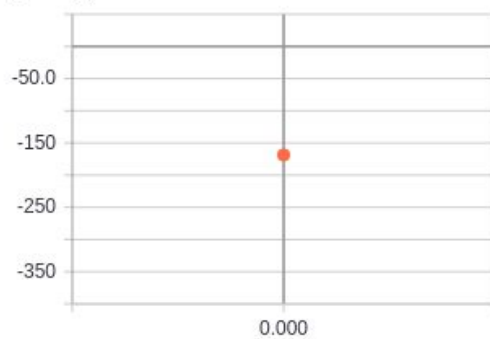
- Loss and qualitative model prediction plots for the 2 of training steps:



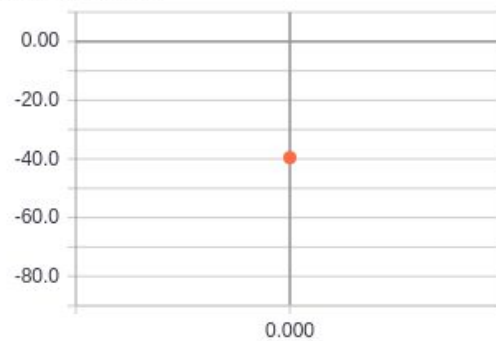
Qualitative model predictions look better for $n=5$, despite the large loss. Overfitting might be the reason for that.

Question 2

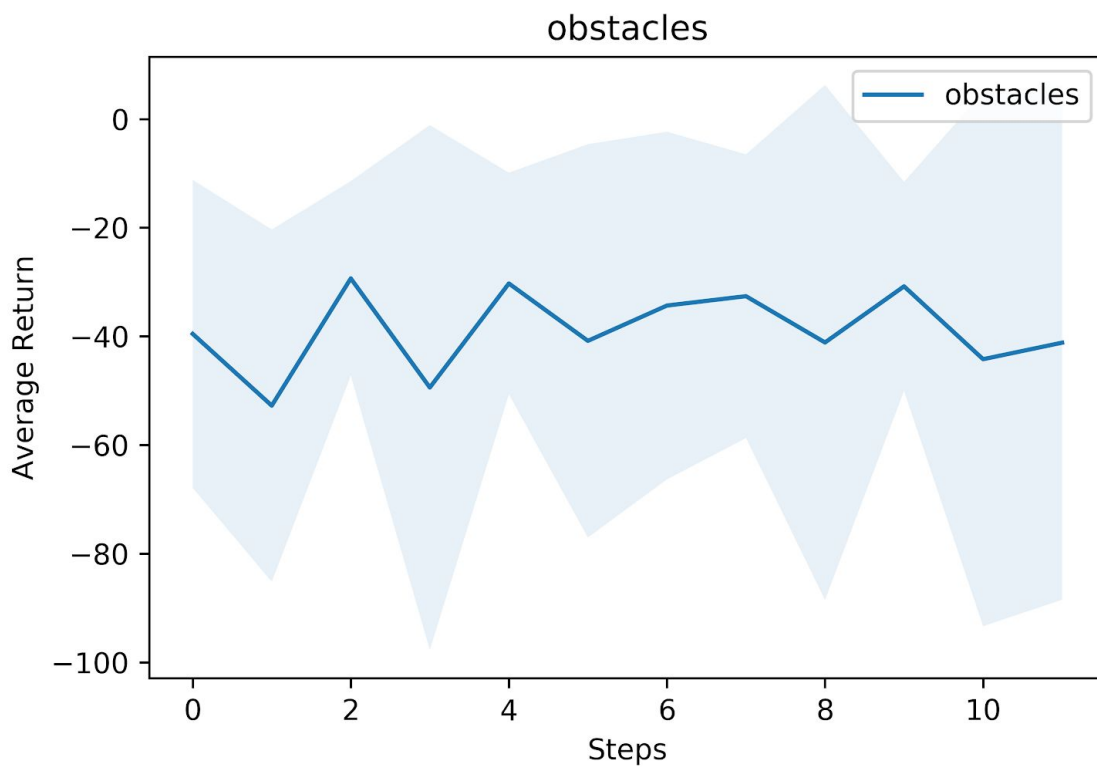
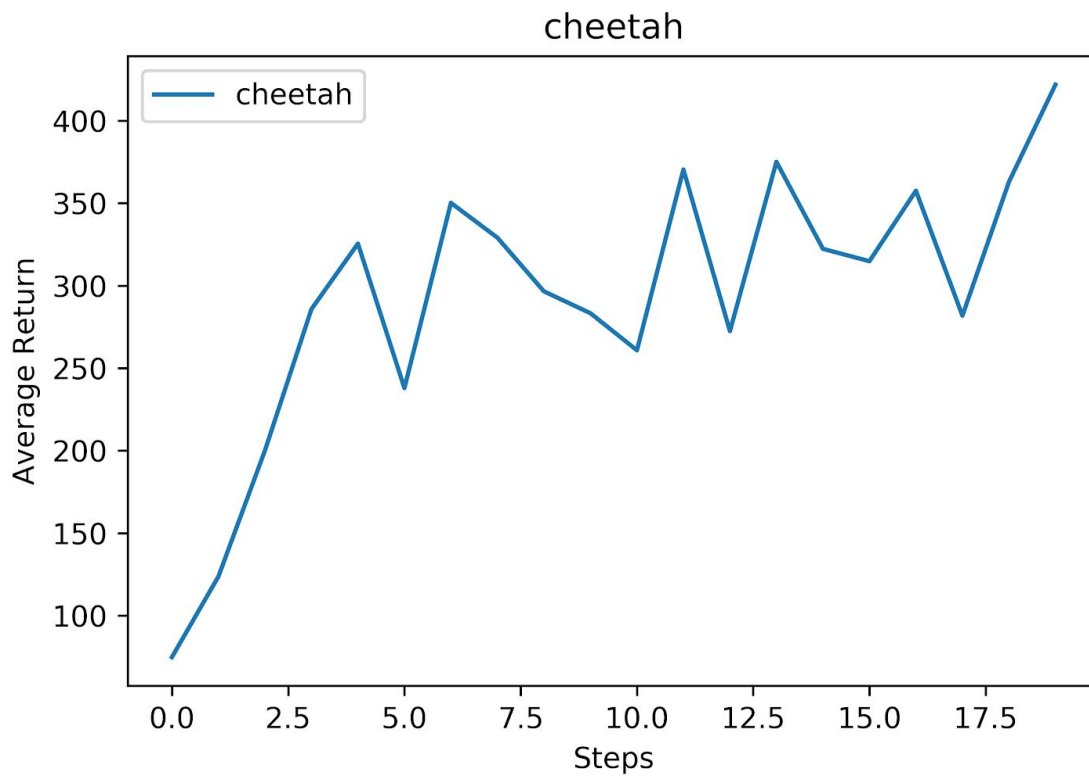
Train_AverageReturn

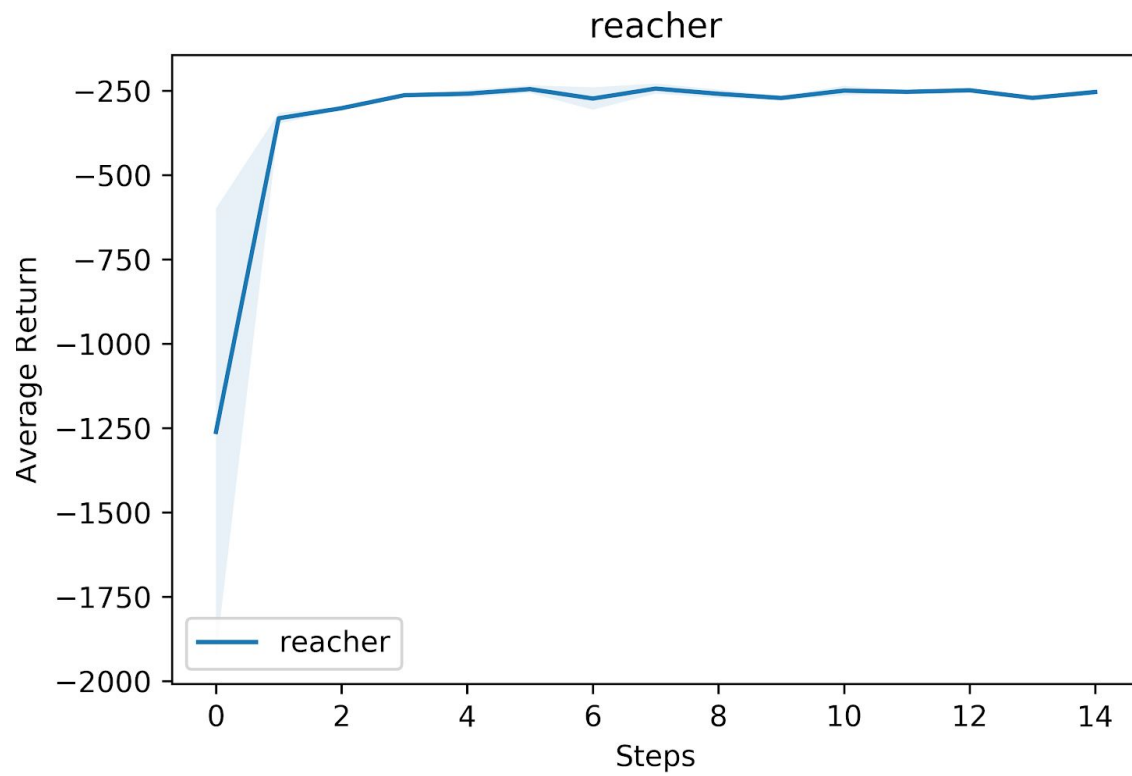


Eval_AverageReturn

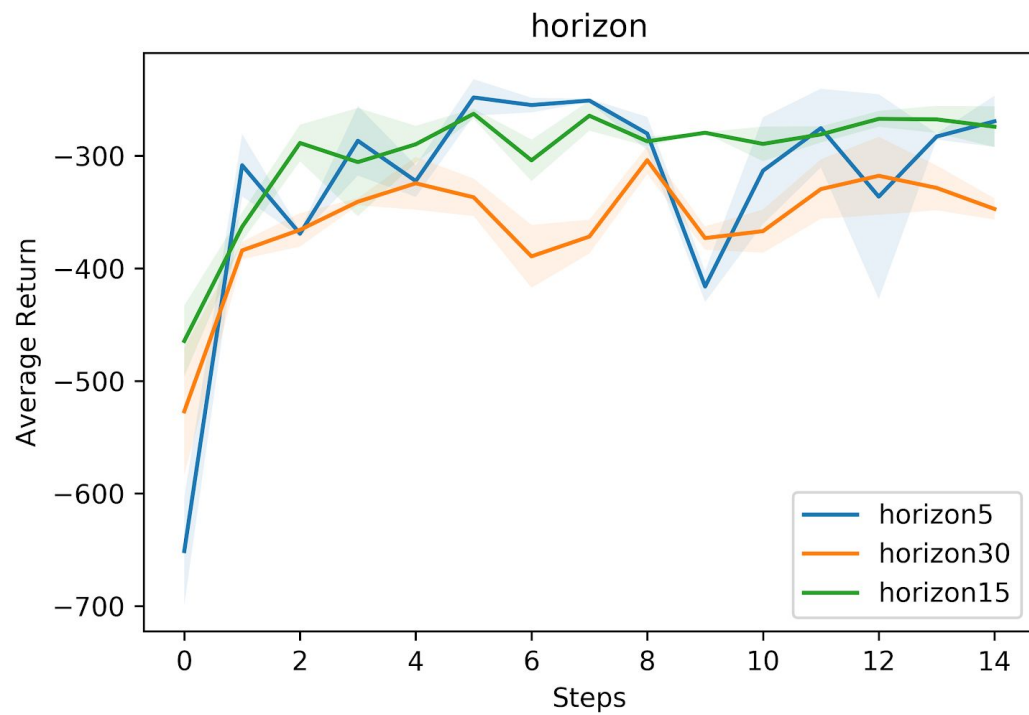


Question 3

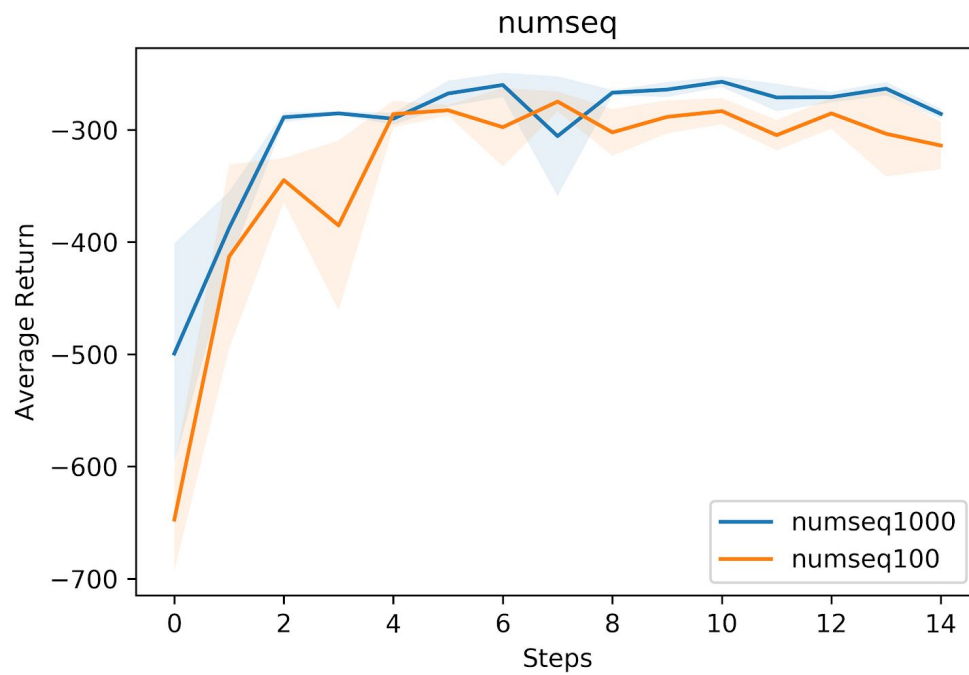




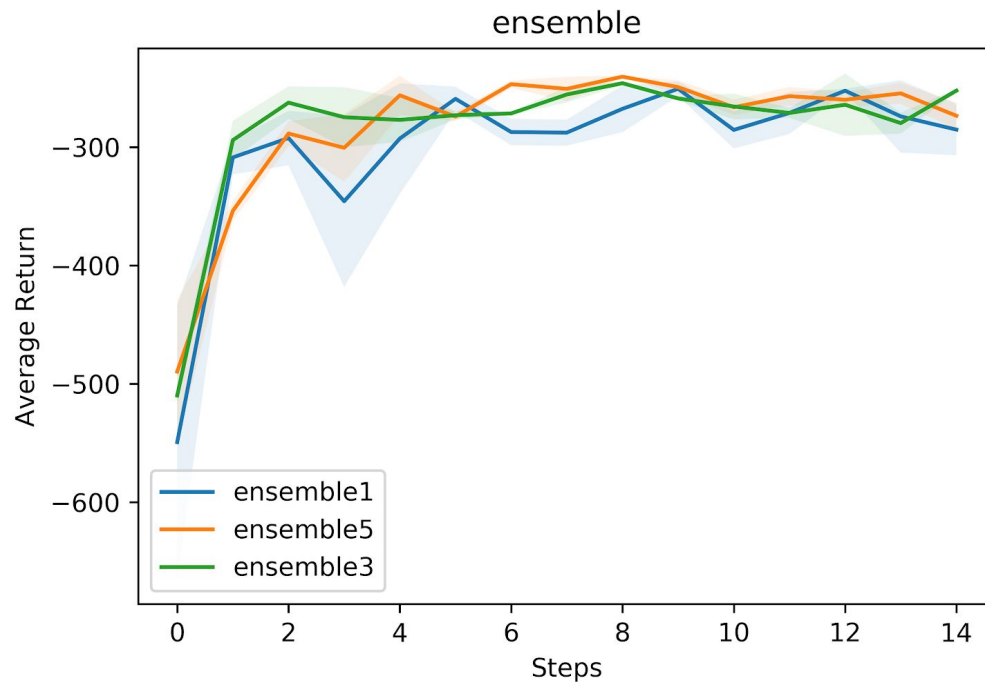
Question 4



Horizon of 15 appears to be the optimal value for the parameter



As could be expected, the number of sequences to be optimized over increases the computational time and boosts the performance



The performance appears to be very stable with respect to the number of ensembles