

Gabriel Ewing

I was intrigued by the paper we read in class on October 31 (Incentivizing Exploration in Reinforcement Learning with Deep Predictive Models, Stadie et al.). My first goal is to reproduce their results on at least some of the games where they successfully learned, hopefully including some of those that the DeepMind group was unable to beat in 2013. I have the Arcade Learning Environment set up on my laptop and I think TensorFlow should be sufficient for the techniques they used.

As we discussed in class, Stadie et al. trained a model of the environment dynamics on the autoencoded state, but then used it for exploration outside of the Deep Q-Network setup. I want to use the same model to improve the learning directly. I plan to start by just using the model output as a parallel input to the Q-network and iterate from there. I think this should be able to yield significant improvements over the standard DQN formulation.

Using the emulator means I don't need to worry about finding a data set or cleaning up the data. If my laptop proves to not be powerful enough, I may request HPC or AWS access.