Unlocking Autonomous Telescopes through Reinforcement Learning: An Offline Framework and Insights from a Case Study

Franco Terranova[1,2], Maggie Voetberg[2], Brian Nord[2,3,4,5], Eric Neilsen[2]

1. University of Pisa, Department of Information Engineering, Pisa, IT 2. Fermi National Accelerator Laboratory, Scientific Computing Division, Batavia, IL 3. Department of Astronomy and Astrophysics, University of Chicago, Chicago, IL 4. Kavli Institute for Cosmological Physics, University of Chicago, Chicago, IL 5. Laboratory for Nuclear Science, MIT, Cambridge, MA

FERMILAB-POSTER-23-327-CSAID

Introduction

Reinforcement Learning (RL) is a paradigm of Machine Learning used to train an autonomous agent capable of behaving optimally in a given environment.

What if this paradigm could be used to optimize the potential of telescopes as sky watchers? An autonomous telescope, unburdened by human biases and complications, could be able to discover solutions we've been missing this whole time.

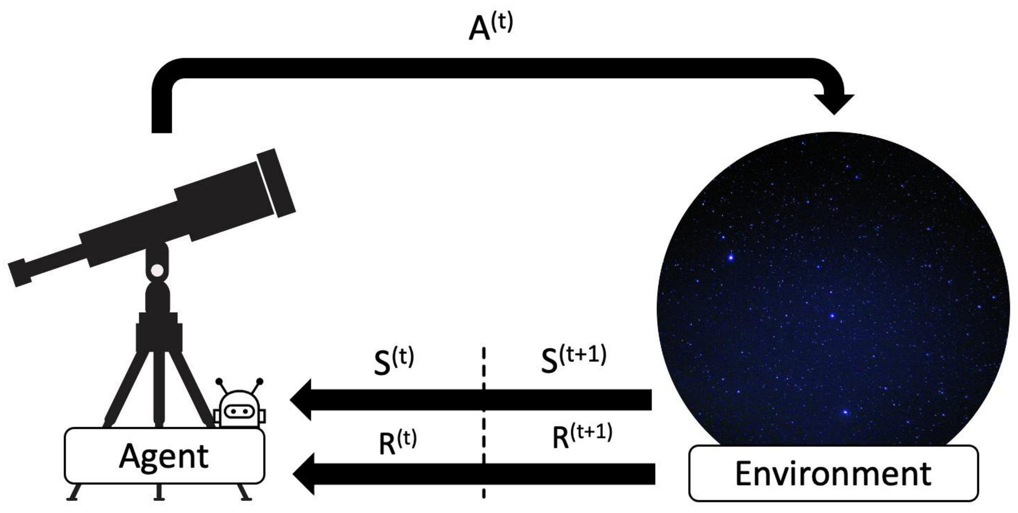


Figure 1. The RL paradigm explained: the presence of the agent, environment, states, actions, and rewards. The objective of the agent is learning an optimal or nearly-optimal policy able to map states to actions, with the goal of maximizing the expected discounted cumulative reward, also known as expected return.

Methodology

A framework has been developed for leveraging a dataset containing examples of interactions between a telescope and the sky as an environment. The dataset should contain (state, action, reward, next state) records that will be sampled to enable learning. Experiment configurations will personalize the training and testing phases.

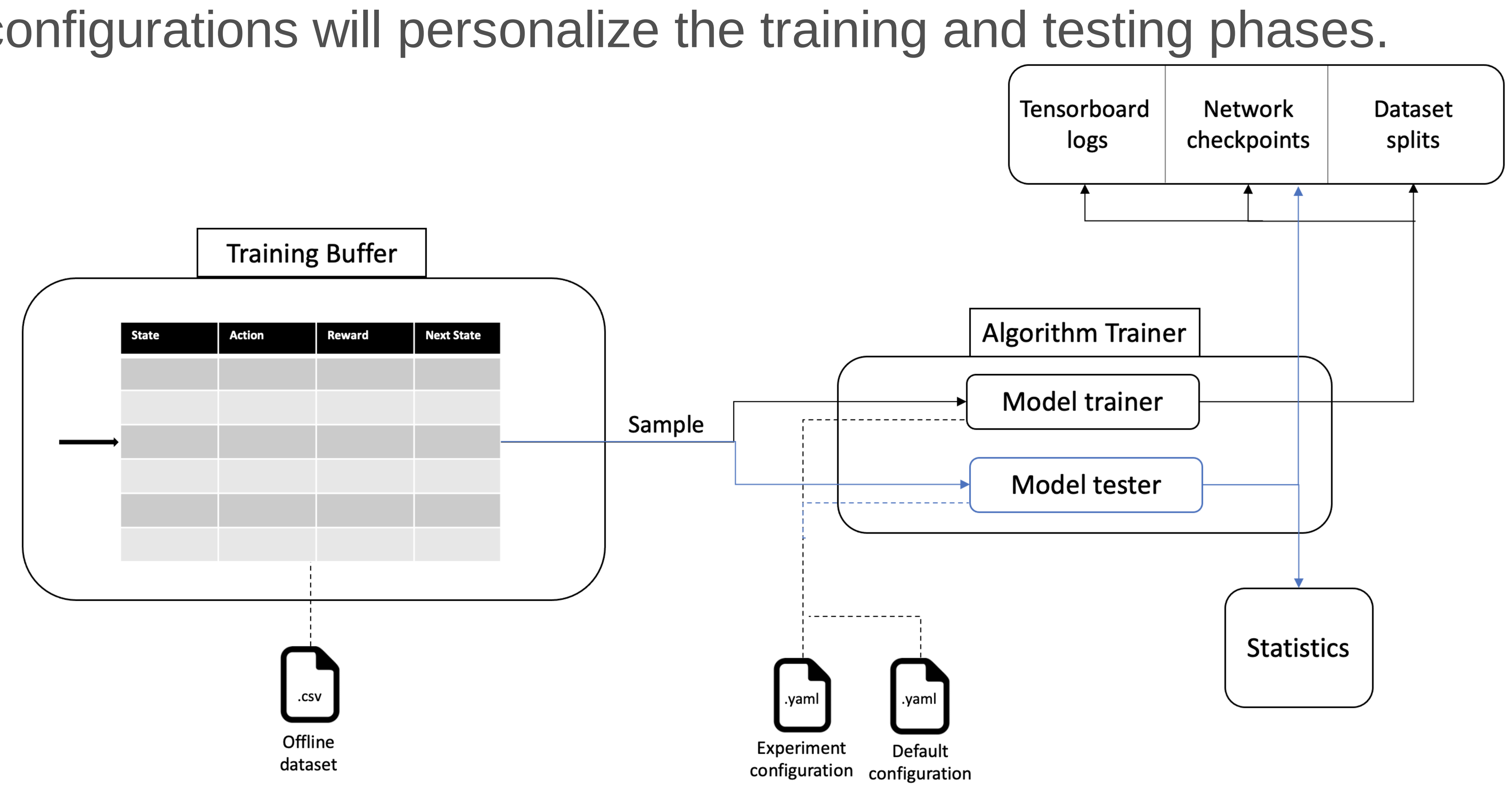


Figure 2. The RL framework enabling learning: a training buffer will wrap the user dataset as an environment, while experiment configurations will guide the process. A model trainer and tester will be used to train and evaluate the resulting self-driving telescope.

Our results have been obtained using a dataset simulated at the Stone Edge Observatory (SEO) with a state space containing variables related to the telescope, the moon, and the sun, an action space consisting of right ascension/declination pairs, and t-effective as a reward variable.

The action space has been properly discretized in an optimized number of bins and angular distance mapping is used to gather rewards based on predictions. This mechanism ensures consistency with possible lack of action samples in front of a portion of the states.

Several methods for discrete action spaces have been compared using a symmetric network structure and hyper-parameters.

Results

Several methods in the class of policy-based, value-based, and evolutionary approaches have been compared based on a holdout method. Using the average effectiveness distribution as a metric, obtained by playing many episodes and averaging the normalized reward gained by the trained network, value-based methods have shown remarkable success.

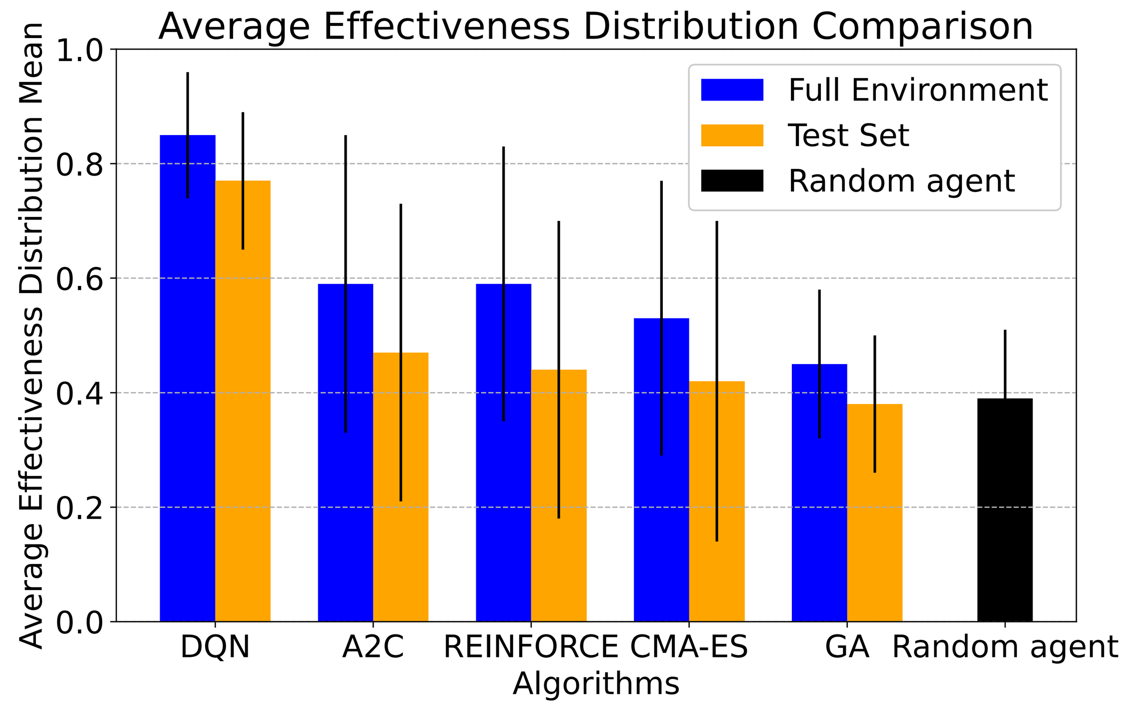


Figure 3. Average effectiveness distribution comparison among some value-based, policy-based, and evolutionary computation strategies.

Several extensions to the Deep Q-Network method (DQN) have been added to the model, ensuring higher performance and generalization capabilities. The extensions have been further combined into a Rainbow DQN network.

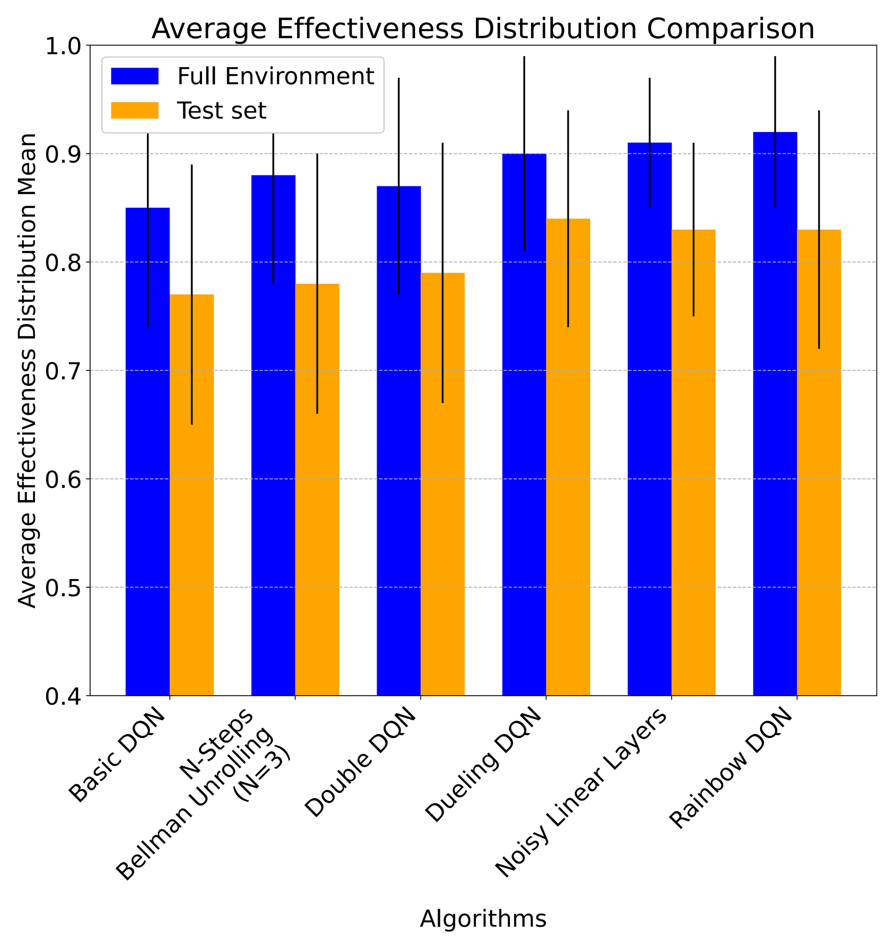
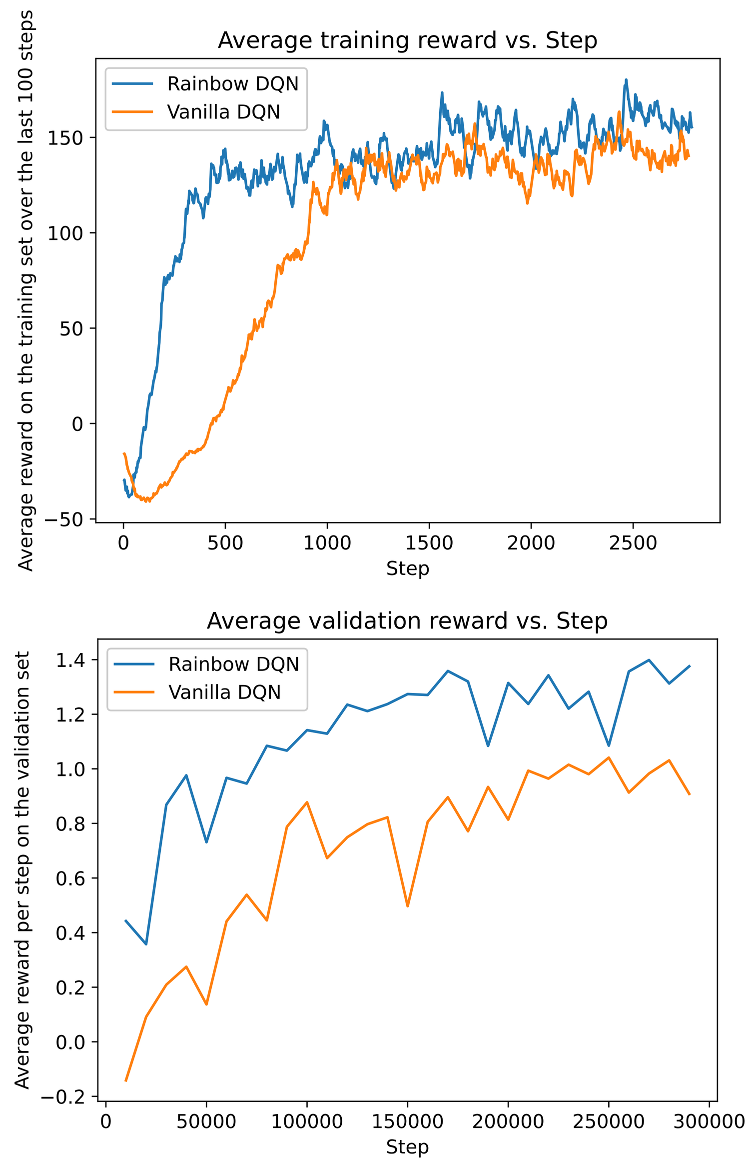


Figure 4. Average effectiveness distribution comparison among some extensions to the DQN method.



Figures 5 & 6. Average reward on the training and validation sets: comparison between the Rainbow DQN and the Basic DQN method.

Conclusions

A proper framework has been developed for the purpose of wrapping offline datasets with the goal of using them as environments and enabling RL training. Our results on the SEO dataset suggest that RL algorithms can be used to optimize the sequential schedule of a telescope survey.

In partnership with:

This manuscript has been authored by Fermi Research Alliance, LLC under Contract No. DE-AC02-07CH11359 with the U.S. Department of Energy, Office of Science, Office of High Energy Physics.