

Deep Reinforcement Learning that Matters

1 Main Ideas

This paper studies the reproducibility of works done in the field of Reinforcement Learning. They show that the interest in RL is growing fast. However, if the working criteria in this field won't change to the direction of making a reproducible work, people will lose their interest in working on RL.

They have multiple suggestions in order to reach to this goal. In the following, we will see these factors and a brief summary of how they can matter.

1.1 Hyperparameters

Tuning hyperparameters are of a great importance in the performance of our algorithm. It is suggested that the proposed algorithm should narrow down the choice of hyperparameters by itself. So, we want our algorithm to be hyperparameter agnostic. In addition, it is preferred that the result is reported using a variety of hyperparameters, so that the effect of them becomes clear.

1.2 Network Architecture

Choosing a good network architecture and activation function is crucial for the success of the algorithm. Different architecture will lead to significantly different results. This is where having hyperparameter agnostic algorithm comes in handy. Because we don't want to deal with choosing the hyperparameters too, when we are trying to find a good network architecture, which, itself, is a hard task and deals with selecting multiple parameters.

1.3 Reward Scale

In gradient based algorithms, scaling the rewards up will cause some issues regarding to saturation and inefficiency in learning. So, what we need is to have a common definition for the reward, so that we can use it as a baseline in other researches in RL.

1.4 Random Seeds

The stochasticity in the environment and the learning process has a drastic influence on the result of our algorithm. It's common to report an average of

the results across multiple trial, or take the max over multiple trial. But, those could also be misleading. We need more research in this area to find a more reliable approach in reporting results with randomness in their nature.

1.5 Environment

Most of the works that are done in the branch of control in RL, suffer from the fact that the environment can play an important role, yet it's often neglected. It is common to only report the results for one set of environmental variables. However, changing this variables can drastically change the performance.

1.6 Codebases

Often times a proposed algorithm is tested against other algorithms that are implemented by the authors themselves. Subtle difference in implementation can lead to a huge difference in performance. So, we need to have a united platform in RL to test our proposed algorithms in a more fair way.