LLM Policies for Text-based Reinforcement Learning: An Interactive Tutorial

Mauricio Tec
mauriciogtec@hsph.harvard.edu
Harvard University

Abstract

This tutorial considers key challenges and techniques in utilizing large-language model (LLM) policies for reinforcement learning. We present an interactive notebook demonstrating how to train an agent in the Textworld environment. The tutorial covers key topics such as (1) parameterizing a policy using an LLM for text generation, (2) supervised fine-tuning with expert demonstrations, and (3) reinforcement learning with proximal policy optimization. The tutorial highlights strategies for efficient computation and memory management, including quantization and low-rank adaptation, which are crucial for scenarios with limited computational resources. This tutorial is designed for researchers familiar with reinforcement learning but with possibly limited hands-on experience in training and fine-tuning LLMs. The tutorial is available as an interactive Colab notebook at https://colab.research.google.com/drive/17oQqcbIJeM3EIruP4T2ju_-LNQmuqqYg?usp=sharing and can run on freely available GPUs.