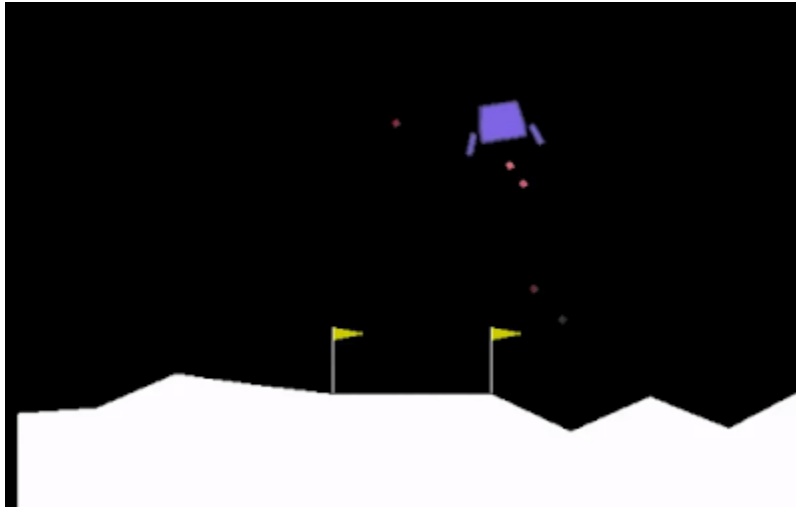In this project you will implement a reinforcement learning (RL) algorithm for playing a variant of a game called "lunar lander".

In the lunar lander environment, a *lander* must be controlled to land between two flags on the ground. When controlling the lander, the goal is to successfully land upright on its two legs, between the flags and in a limited timeframe, whilst expending minimal fuel. Your task is to implement an algorithm that, by practicing on a simulator, learns a control policy for the lander.

The `.pdf` of the handout contains the following additional sections. Environment and Scoring Details will give a more thorough description of the Lunar Lander environment, along with the requirements for passing the baseline. Solution Details will guide you through a set of minimal modifications to `solution.py` for passing the task baseline. We also provide an appendix on policy gradients which will briefly review the lecture material and give you the additional information needed for completing the TODOs in the code for this task.