Assignment: Compare PG, TRPO, and PPO

- Environment: 1. Mujoco https://gymnasium.farama.org/environments/mujoco/
2. Box2d   https://gymnasium.farama.org/environments/box2d/
3. Classic control https://gymnasium.farama.org/environments/classic_control/
Pick any 5 environments of your choice(try at least one from Mujoco)
- Performance comparison (reference): Run at least 10^4 steps
https://spinningup.openai.com/en/latest/spinningup/bench.html
- Algorithms to implement and compare: Policy gradient (PG), (2)TRPO, (3) PPO-clip
  - Implement Generalized Advantage Estimator for gradient computation
  - Use **Pytorch**. Standard functions (backprop, loss function, optimizer) can be used
- Deliverables: A Zip file with
  - One Jupyter Notebook for each algorithm (total 3 notebooks). Notebooks should be self-explanatory with proper documentation. At the end, show performance graphs for each of the 5 datasets.
  - A 1-page doc file with graphs comparing all 3 algorithms on 5 datasets
- Deadline 1 (Zip file): 10th October
- Deadline 2 (A 10-minute video explaining your code): 12th October