Computer Engineering Department
CMPS458: Reinforcement Learning Fall 2025
Ayman AboElHassan, PhD

Assignment 3

The goal of this assignment is to implement advanced Policy Gradient models using PyTorch, train RL agents on multiple classical environments, and use hyperparameter tuning to find the most suitable parameter set.

Summary

1. Update Gymnasium environment run script to apply hyperparameter tuning for
    - Discount Factor
    - Epsilon Decay Rate
    - NN Learning Rate
    - Replay Memory Size
    - Learning Batch Size
2. Implement A2C, SAC, PPO Models
3. Train the 3 models in the following classical environments
    - CartPole-v1
    - Acrobot-v1
    - MountainCar-v0
    - Pendulum-v1
4. Run trained agents 100 tests per environment and track the test episode duration
5. Use “RecordVideo” Wrapper to record the RL agent in-action
6. Answer the following questions

Questions

1. For each of the classical environments:
    1. What is the difference between RL models in terms of training time and performance?
    2. How stable are the trained agents? Show with test episode duration figures.
    3. Explain from your point of view how well-suited Policy Gradient is to solve this problem.
2. Compare Policy Gradient results to DDQN results from the previous Assignment.
3. Does the hyperparameter tuning results match the best hyperparameters used in the previous Assignment? Describe your interpretation.

Deliverables

Use this report template to deliver the following requirements
(https://www.overleaf.com/read/dsmwrcyzyyjs#ef6078)

1. GitHub repository with Python codes (Gym Environment, A2C/SAC/PPO models)
2. The recorded video of the trained agent in action
3. The Experiment charts generated by Weights and Biases
4. Report with the outcome summary and answers to the questions asked

Due date

4 December 2025

Helping Materials

- https://www.geeksforgeeks.org/machine-learning/actor-critic-algorithm-in-reinforcement-learning/
- https://gibberblot.github.io/rl-notes/single-agent/actor-critic.html
- https://spinningup.openai.com/en/latest/algorithms/ppo.html
- https://spinningup.openai.com/en/latest/algorithms/sac.html
- https://wandb.ai/site/articles/intro-to-mlops-hyperparameter-tuning/
- https://medium.com/@bhatadithya54764118/day-21-hyperparameter-tuning-basics-grid-search-random-search-d6f19b44294c