Reinforcement Learning Project

Overview

Welcome to your Reinforcement Learning (RL) course project! You will work in groups of up to **5 students** to develop two RL agents capable of solving different environments from the **Gymnasium** library (https://gymnasium.farama.org/). **The project deadline has been set to the 20th of June.**

The main objective of this project is to explore the design and implementation of RL algorithms in environments with diverse state and action spaces. You are required to choose **two different environments** from Gymnasium that were **not covered in class**. At least **one environment must have a non-discrete/non-tabular observation space** (e.g., continuous or mixed). However, students are **allowed to discretize it** if they want to apply tabular methods (such as Q-learning or Monte Carlo).

Project Requirements

1. **Environments**:
    - Choose two different environments from the Gymnasium library (**not used in class**).
    - At least **one environment must have a non-discrete observation space**.
    - Document your choice in the report with a brief explanation of each environment.
2. **Algorithms**:
    - Implement **two different RL algorithms** (e.g., DP, Monte Carlo, Q-Learning, SARSA, DDPG, PPO, etc.) **in each environment chosen**.
    - You may choose algorithms seen in class or explore new ones.
3. **Evaluation and Analysis**:
    - Train your agents and record performance metrics, e.g.:
        - Total reward per episode/Return.
        - Convergence speed.
        - Exploration vs. exploitation balance.
    - Provide **visualizations** of learning progress (e.g., reward curves, success rates).
    - Perform a **comparative analysis** of both algorithms and environments.

Project Objectives

1. **Solve Two Distinct Environments Using Different Algorithms:**

- Your task is to demonstrate a solid understanding of RL by solving two distinct environments, applying distinct algorithms to each. For instance, one agent could use a model-free RL method, while the other uses a policy gradient method.

2. **Optimize Decision-Making Through Effective Exploration and Exploitation Strategies:**

- A key aspect of RL is balancing exploration (trying new actions to discover better strategies) and exploitation (relying on known strategies to maximize reward). Both strategies are crucial for solving the environments you choose.

3. **Analyse and Visualize Learning Progress, Convergence, and Policy Effectiveness:**

- It is essential to track your agent's learning progress. Visualize metrics such as rewards, episode lengths, and success rates. These visualizations will not only help you understand your agent's performance but will also form an important part of the report.

4. **Interpret and Explain Agent Behaviour in Each Environment:**

- It is important to understand how your agent learns to solve each task. In your report, provide an analysis of the agent's behaviour—what actions was it choosing, what patterns emerge in its learning process, and what can be improved in terms of the agent's policy.
- Analysing agent behaviour is crucial for interpreting whether the agent is exploring effectively and exploiting its knowledge efficiently.

Deliverables

1. Code Implementation

- Two separate implementations for each agent/environment combination.
- Include all scripts and utilities required for reproducibility.
- Provide clear instructions for running the experiments.

2. Project Report (Max 10 pages or 3000 words, excluding references & images)

- **Introduction:** Overview of the environments and algorithms chosen.
- **Methodology:** Detailed explanation of:
    - Environment descriptions.
    - Discretization techniques (if applied).
    - Algorithms used and why they were chosen.
- **Evaluation:**
    - Performance metrics (reward curves, convergence analysis, etc.).
    - Comparative analysis of algorithms.
    - Limitations and challenges faced.
- **Visualizations:** Meaningful plots to represent learning and performance.

- **Conclusion:** Summary of findings and potential improvements.

Project Evaluation Criteria

- The usage and argumentation of different learning methods will be the focus of the project's evaluation process. Choose appropriate learning algorithms for each observation and/or action type and argue their relevance and usefulness in solving the problem.
- While the overall performance of your RL agent is not to be part of the evaluation, it is rather essential to understand how different learning methods improve the performance of your agent.
- It is important that you develop a comprehensive evaluation to the extent to which the agent's reasoning can be understood. Assess the clarity and interpretability with visualizations.

Submission Guidelines

- Submit a compressed folder in Moodle containing:
    - Group members in a clear manner.
    - All code files and dependencies.
    - A well-structured project report in PDF format.
    - Any additional notes or documentation.