



STA303 Artificial Intelligence Course Project



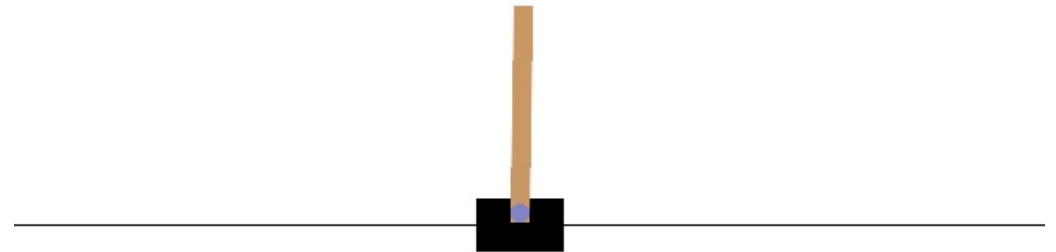
1. Introduction

- **CartPole Problem:**

The CartPole environment is a classic control problem where the goal is to balance a pole on a moving cart by applying forces to the left or right.

- **Project Goal:**

To design and train an intelligent agent that can successfully balance the pole in the CartPole environment using RL techniques.



2. Environment

- **Environment:** OpenAI Gymnasium – CartPole-v1
- **State Space:** 4-dimensional
(cart position, cart velocity, pole angle, pole angular velocity)
- **Action Space:** Discrete — {Move Left, Move Right}
- **Reward:** +1 for every time step the pole remains balanced
- **Termination:** Episode ends when the pole falls or the cart moves out of bounds



3. Tasks

- **Baseline:** A DQN implementation is provided as a starting point.
- **Basic requirement:**
 - Implement and train at least two different types of RL algorithms.
 - Value-based: DQN (provided as a baseline), double DQN, ...
 - Policy-based: REINFORCE, PPO, TRPO, etc.
 - Actor–Critic: A2C, DDPG, etc.
- **Performance:** If the average score exceeds a specified threshold, you will receive full marks for this part.



3. Tasks (cont'd)

- **Advanced**
- **Go beyond traditional online RL algorithms:**
 - **Offline RL:** train from a fixed dataset without new environment interaction (e.g. using BCQ, TD3+BC, or IQL).
 - **Imitation Learning:** mimic an expert's policy using Behavioral Cloning (BC) or GAIL.
 - **Hybrid or Novel Approaches:** combine methods creatively and analyze the result.



4. Encouraged Directions

You are **strongly encouraged to think critically and explore creatively.**

Beyond implementation, we expect you to discuss questions such as:

- How do important hyperparameters influence learning stability and convergence?
- Which module of the algorithm is most critical for performance and why?
- Can you propose a simple modification that improves the original version?



5. Team project

This is a **team project** — each team should consist of **three students**.

Team registration information will be announced **next week**.

