Sequential Decision Making and Reinforcement Learning – Final Project

Submission Instructions

1. **Submission Deadlines**: The project must be submitted no later then 1.3.2025
2. **Group Submissions**: Projects must be submitted in groups of 2
3. **Submission File Format**: Each team should submit a zip file named <ID1_ID2_ID3>, where <ID<i>> represents the ID number of each of the 2 team members. The zip file must include the following files:
   - <ID1_ID2>.pdf
   - <ID1_ID2>.MP4

   where the pdf file is the final project report and the MP4 file is a the project video presentation (both detailed next).
4. **Report Requirements**:
   - The report should include a link to a GitHub repository, as detailed in the instructions provided later.
   - The report (PDF file) must be written in English.
5. **Questions**:
   - Any general questions about the project should be posted in the Moodle forum.
   - Specific questions can be asked through the email to the staff. Please include an indicative title, mention who are the team members and write your questions clearly in a numbered list.
   - **Questions in the Moodle forum will be prioritized in the order of answering.**

1 Introduction

In this final project, you will apply the concepts covered in the course to a real world challange: optimizing the behavior of an agent (market players) in a dynamic environment (electricity market). The project will include the following sections:

1. **Implementation:** You will build the environment and agent according to the specifications in the project description, defining the state space, actions, and rewards. Tools such as Gymnasium, introduced in class, are encouraged.

2. **Experimentation:** Using techniques covered in the course, you will train the agent to find effective policies.

3. **Creative:** You'll explore your own approaches to improve the agent's performance. This could involve modifying existing methods or developing entirely new strategies.

2 Background: The Evolution of the Electricity Market

For many years, electricity systems operated in a straightforward manner: centralized government-run entities generated power, managed its distribution and set fixed prices for consumers. However, with the advent of renewable energy technologies and a push toward decentralization, this traditional structure is evolving. Private companies and individual operators are increasingly participating in the electricity markets, introducing flexibility but also adding complexity.

In Israel, the electricity market is undergoing a significant transformation. While the Grid System Operator (GSO), a government-affiliated company, still manages the grid's stability and operation, private energy producers are entering the market. These participants include private companies that use solar panels, wind turbines, and battery storage systems. Unlike traditional power producers, these entities operate independently, buying and selling electricity at prices determined by the GSO rather than following direct government orders.

2.1 Agent

The agent in this framework is a **market player** participating in the electricity market. The agent operates as a unit that includes components such as battery storage systems. The agent's goal is to maximize its profit by optimizing decisions regarding energy production, storage, and transactions with the grid.

2.1.1 Agent's Actions

At each timestep t, the agent can take a **single** continuous action.

1. **Charge or Discharge Battery Storage (a_t):**
   - Charge the battery using produced power (cannot charge more than the capacity of the battery).
   - Discharge the battery to meet internal load or sell power to the grid (cannot discharge more than the capacity).
   - Action value range: [-Battery Capacity, Battery Capacity]

2.2 Environment

The environment is the **electricity market**, operated by the system manager, which determines the prices (market price). The environment varies in complexity:

The agent observes the state of the environment, which includes:

- SoC: State of Charge - how much the battery is "full". Values range: [0, Capacity].
- $\mathcal{D}_t$: Demand of electricity, how much electricity the household consumes (and therefor cannot sell to the grid). Values range: $[0, \infty]$
- $\mathcal{P}_t$: Price, a stochastic random variable representing price at which electricity is sold to the grid. The price is drawn from a random bariable which changes according to the markek value of electricity. Values range: $[0, \infty]$

Policy Objective

Every time the agent discharges power, he first answers the demand of the household $\mathcal{D}$. Every unit of power discharged after answering the demand will be sold to the electric grid at the price observed by the environment.

$$\text{Maximize} \sum_{t=1}^{T} \mathcal{B}_t \cdot \mathcal{P}_t$$

where:

- $\mathcal{B}_t$: The amount of power discharged to the grid at time t
- $\mathcal{P}_t$: The current price of electricity at time t.

Notice that the agent will receive a reward for discharging power to the grid **only** after meeting the demand of the household.

3 Section One: Implement the Environment and Agent

In this section, you will:

1. **Define the Environment:** The environment models the electricity market. Implement the following components:
   - **State Variables:** Define the state space including SoC (State of Charge), $\mathcal{D}_t$ (electricity demand), and $\mathcal{P}_t$ (electricity price).
   - **Dynamics:** Model the stochastic evolution of the market price ($\mathcal{P}_t$) and demand ($\mathcal{D}_t$). The market price and demand should both be periodic functions with two "peaks". Both functions should be noisy, meaning that a random noise should be added to the function value at each timestep. An example of the demand function could be a combination of two normal distributions. For example: $f(x) = 100 \cdot e^{\frac{-(x-0.4)^2}{2\cdot(0.05)^2}} + 120 \cdot e^{\frac{-(x-0.7)^2}{2\cdot(0.1)^2}}$
   - **Reward Function:** Design a reward function to maximize profits while meeting household demand.
2. **Design the Agent:** The agent represents a market player. Implement the following:
   - **Actions:** Continuous action space for charging/discharging the battery, bounded by battery capacity.

4 Section Two: Train a Policy

1. **Training Framework:** Use reinforcement learning algorithms to train the agent. Suggested algorithms include Deep Q-Networks (DQN), Proximal Policy Optimization (PPO), or Soft Actor-Critic (SAC).
2. **Experiment Setup:** Configure the environment with realistic parameters for battery capacity, electricity demand, and price dynamics.
3. **Evaluate Policies:** Assess the agent's performance based on:
   - Total profit over a fixed time horizon.
   - The percentage of demand met from battery discharge.
   - Stability and robustness of the policy in varying market conditions.
4. **Visualization Analysis:** Analyze the performance of the agent. Plot training curves for rewards and analyze policy behavior over time and any other metric you think relevant.

5 Section Three: Research a New Training Paradigm

1. **Objective:** This section is intended to be a research-focused component of the project. After applying conventional techniques to train a policy, perform your own research and devise a novel training approach to improve policy performance.

2. **Experimentation:** Compare the proposed creative approach with baseline methods:
    - Train the agent using the newly researched paradigm and evaluate its performance against conventional techniques.
    - Conduct experiments to analyze the impact of key components in the new approach and justify their effectiveness.
3. **Insights:** Summarize findings from the research, discuss the advantages and limitations of the new approach, and suggest potential improvements or future directions for further exploration.

Submission

Your final submission should include a PDF report structured similarly to a short academic paper. The report should clearly convey the entire process you went through during your project and summarize all aspects of your work. The report must be a maximum of 6 pages, excluding the appendix. The appendix should list the main papers, videos, and blog posts (no need to add a formal bibliography). By and large, your report should be structured as follows:

- **Title and Author Information:** The name of your work, and the names and emails of the team members.
- **Abstract:** A brief summary of the project, including the problem addressed, the methods used, and the main findings.
- **Methodology:** Describe in detail the implementation of the agent and the environment and the methodology you employed. This is supposed to best explain your project.
- **Experiments:** Explain the experimental setup and evaluation metrics. Conduct experiments to demonstrate the key points conveyed in your work. Include tables, graphs, and other visual aids to support your findings. Discuss the performance of your approach compared to baselines.
- **Short Discussion:** Provide insights and potential improvements for future work. Summarize the key findings of your project.
- **Appendix:** List the main papers, videos, and blog posts you relied on. Include a link to your Github repository containing all code and related materials.

Code

Your GitHub repository is a major component of your project submission. It should demonstrate not only the functionality of your code but also the quality of your collaboration and documentation. The repository must adhere to the following requirements:

- **Code Quality and Structure:** Ensure that your code is well-structured and organized into meaningful files (and if needed, directories).
- **README file:** Include a `README` file that provides an overview of the project, setup instructions, and how to run your code.
- **Documentation:** Document all significant parts of your code using comments and docstrings.
- **Reproducibility:** Ensure that your code can be easily set up and run by others. This includes providing clear setup instructions and listing any dependencies.

Video Presentation

The video presentation should be attached to you submission. It should allow you to concisely communicate the key aspects of your project and demonstrate your work visually. The video length must be at most 5 minutes, in which you present the problem, motivation, methodology, experiments and insights of your work.

Grading

Your final project will be evaluated based on three main components: the written report, the code submission, and the video presentation. The grading distribution is as follows:

- **Report (60%):**
  - **Content and Clarity:** Your report should thoroughly explain the problem, methodology, experiments, results, and conclusions.
  - **Novelty:** The creative aspect cannot be a copy-paste of existing works. Your novelty should be well-stressed in the report.
  - **Effort:** The depth of your analysis, the comprehensiveness of your experiments, and the overall effort invested in the project will be assessed.
  - **Documentation:** Clear documentation of your process, including the rationale behind your choices and detailed explanations of your methods.
- **Code (30%):**
  - **Functionality:** Your code should run correctly and produce the results described in your report.
  - **Organization:** The code should be well-organized into meaningful directories and files.
  - **Documentation:** Code should be well-commented, with clear explanations of algorithms, functions, and any non-trivial steps.
- **Video Presentation (10%):**
  - **Clarity:** The presentation should be clear and easy to follow, highlighting the key aspects of your work.
  - **Demonstration:** If you implemented an entire system, briefly demonstrate it in the video.