Overview

In this assignment, you will gain experience working with OpenAI Gym, which is a set of problems that can be explored with different reinforcement learning algorithms. This assignment is designed to help you apply the concepts you have been learning about Q-learning algorithms to the "cartpole" problem, a common reinforcement learning problem.

Note: The original code referenced in this assignment was written in Python 2.x. You have been given a zipped folder containing an updated Python 3 version of the code that will work in the Apporto environment. To make this code work, some lines have been commented out. Please leave these as comments.

Reference: Surma, G. (2018). *Cartpole*. Github repository. Retrieved from https://github.com/gsurma/cartpole.

Prompt

Access the Virtual Lab (Apporto) by using the link in the **Virtual Lab Access** module. It is recommended that you use the Chrome browser to access the Virtual Lab. If prompted to allow the Virtual Lab access to your clipboard, click "Yes", as this will allow you to copy text from your desktop into applications in the Virtual Lab environment.

1. Review the following reading: Cartpole: Introduction to Reinforcement Learning. In order to run the code, upload the Cartpole.zip folder into the Virtual Lab (Apporto). Unzip the folder, then upload the unzipped folder into your Documents folder in Apporto. Refer to the Jupyter Notebook in Apporto (Virtual Lab) Tutorial to help with these tasks.

   Note: The Cartpole folder contains the Cartpole.ipynb file (Jupyter Notebook) and a scores folder containing score_logger.py (Python file). It is *very* important to keep the score_logger.py file in the scores folder (directory).

2. Open Jupyter Notebook and open up the Cartpole.ipynb and score_logger.py files. Be sure to review the code in both of these files. Rename the Cartpole.ipynb file using the following naming convention:

   <YourLastName>_<YourFirstName>_Assignment5.ipynb

   Thus, if your name is Jane Doe, please name the submission file "Doe_Jane_Assignment5.ipynb".

3. Next, run the code in Cartpole.ipynb. The code will take several minutes to run and you should see a stream of output while the file runs. When you see the following output, the program is complete:

   ```
   Solved in _ runs, _ total runs.
   ```

   Note: If you receive the error "NameError: name 'exit' is not defined" after the above line, you can ignore it.

4. **Modify the values for the exploration factor, discount factor, and learning rates in the code** to understand how those values affect the performance of the algorithm. Be sure to place each experiment in a different code block so that your instructor can view all of your changes.

   Note: Discount factor = GAMMA, learning rate = LEARNING_RATE, exploration factor = combination of EXPLORATION_MAX, EXPLORATION_MIN, and EXPLORATION_DECAY.

5. Create a Markdown cell in your Jupyter Notebook after the code and its outputs. In this cell, you will be asked to analyze the code and relate it to the concepts from your readings. You are expected to include resources to support your answers, and must include **citations** for those resources.

   Specifically, you must address the following rubric criteria:
   - **Explain how reinforcement learning concepts apply to the cartpole problem.**
     - What is the goal of the agent in this case?
     - What are the various state values?

        - What are the possible actions that can be performed?
        - What reinforcement algorithm is used for this problem?
    - **Analyze how experience replay is applied to the cartpole problem.**
        - How does experience replay work in this algorithm?
        - What is the effect of introducing a discount factor for calculating the future rewards?
    - **Analyze how neural networks are used in deep Q-learning.**
        - Explain the neural network architecture that is used in the cartpole problem.
        - How does the neural network make the Q-learning algorithm more efficient?
        - What difference do you see in the algorithm performance when you increase or decrease the learning rate?

Guidelines for Submission

Please submit your completed IPYNB file. Make sure that your file is named as specified above, and that you have addressed all rubric criteria in your response. Sources should be cited in APA style.

Module Five Assignment Rubric

Criteria	Exemplary (100%)	Proficient (85%)	Needs Improvement (55%)	Not Evident (0%)	Value
Modifies Values for Exploration Factor, Discount Factor, and Learning Rate	Exceeds proficiency in an exceptionally clear, insightful, sophisticated, or creative manner	Modifies the values for the exploration factor, discount factor, and learning rates in the code with the code working correctly	Shows progress toward proficiency, but with errors or omissions	Does not attempt criterion	10
Reinforcement Learning Concepts for the Cartpole Problem	Exceeds proficiency in an exceptionally clear, insightful, sophisticated, or creative manner	Explains how reinforcement learning concepts are applied to the cartpole problem	Shows progress toward proficiency, but with errors or omissions	Does not attempt criterion	25
Experience Replay in the Cartpole Problem	Exceeds proficiency in an exceptionally clear, insightful, sophisticated, or creative manner	Analyzes how experience replay is applied to the cartpole problem	Shows progress toward proficiency, but with errors or omissions	Does not attempt criterion	25
Neural Networks in the Cartpole Problem	Exceeds proficiency in an exceptionally clear, insightful, sophisticated, or creative manner	Analyzes how neural networks are used in deep Q-learning	Shows progress toward proficiency, but with errors or omissions	Does not attempt criterion	30
Articulation of Response	Exceeds proficiency in an exceptionally clear, insightful, sophisticated, or creative manner	Clearly conveys meaning with correct grammar, sentence structure, and spelling, demonstrating an understanding of	Shows progress toward proficiency, but with errors in grammar, sentence structure, and spelling, negatively impacting readability	Submission has critical errors in grammar, sentence structure, and spelling, preventing understanding of ideas	5

		audience and purpose			
Citations and Attributions	Uses citations for ideas requiring attribution, with few or no minor errors	Uses citations for ideas requiring attribution, with consistent minor errors	Uses citations for ideas requiring attribution, with major errors	Does not use citations for ideas requiring attribution	5
				Total:	100%