Programming Assignment 5

Modified implementation of SARSA and Q-learning control algorithms

1. As the first task you have to modify the SARSA algorithm implementation that was given in class. Modify the code so that the policy improvement step is done in every time slot rather than at the end of every episode.
2. Confirm whether your implementation is correct by comparing the obtained $q(s, a)$ estimates to what was obtained using the prior version of the code.
3. Similarly, modify the Q-learning algorithm implementation that was given in class. Modify the code so that the policy improvement step is done in every time slot rather than at the end of every episode.
4. Confirm whether your implementation is correct by comparing the obtained $q(s, a)$ estimates to what was obtained using the prior version of the code.

Using SARSA and Q-learning for FrozenLake problem

In this series of tasks, you will also get acquainted with the OpenAI gym environments.

1. Install the Gymnasium environment (see https://gymnasium.farama.org/).
2. Read the documentation for the FrozenLake environment https://gymnasium.farama.org/environments/toy_text/frozen_lake/
3. Use your implementation of Q-learning to try and obtain a policy for this problem.
4. Plot the total reward obtained as a function of the number of episodes
5. How does your agent compare with the results obtained previously (see https://github.com/openai/gym/wiki/Leaderboard)

Submission of this assignment will be through Moodle. For this assignment you have to submit a Jupyter notebook. The submission file should have the format "[SCCODE]_[NAME]_PA5.ipynb" (where you have to insert your SCCODE and NAME).