AIDL_B02
Advanced Topics in Deep Learning

Deep Reinforcement Learning Project

Presenter: Panagiotis Kasnesis

MSC Artificial Intelligence and Deep Learning
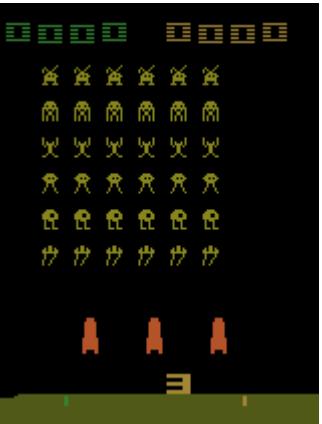
DQN

Game

!pip install "gymnasium[atari,accept-rom-license,other]" ale-py

"ALE/SpaceInvaders-v5"

List of available actions

Value	Meaning
0	NOOP
1	FIRE
2	RIGHT
3	LEFT
4	RIGHTFIRE
5	LEFTFIRE

https://github.com/ounospanas/AIDL_B02/blob/main/AIDL_B02_REINFORCE_Pong_PyTorch.ipynb

MSC Artificial Intelligence and Deep Learning

Instructions

- **4 stacked input frames**
- Better resize them and crop them (don't loose info!)
- Actions **6** in total (DON'T discard)
- **Add replay buffer!**
- **Add target network!**
- The DQN network architecture is your choice...you could try the original DQN

Objective

- Deliver 3 value-based implementations (play around with the hyperparameters, like image size, layers, neurons, lr, epsilon, even try other approaches like double DQN, dueling DQN, prioritized experience replay, etc)

- Receive **at least 500** in average (100 games) **at least one implementation,** with the other two reaching **at least 400**

- Keep track of the results to be presented

Hints
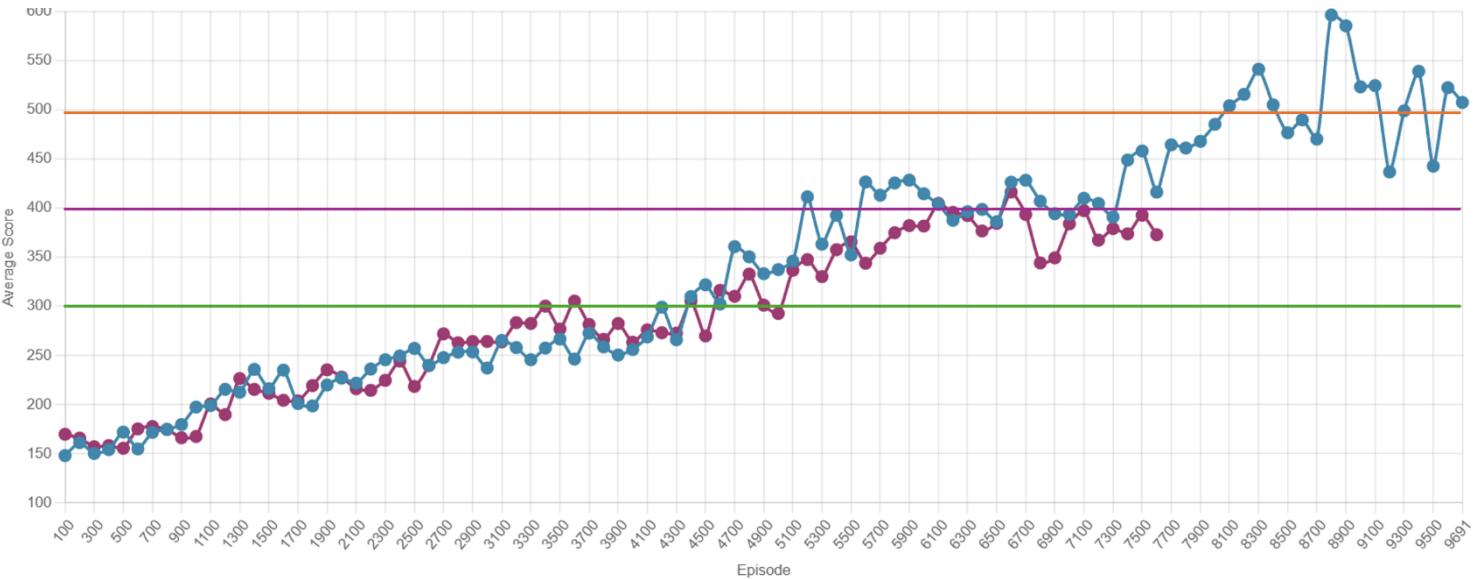
- You can use the exact same code for DQN provided in [AIDL_B02/AIDL_B02_Deep_Q_Network_Solution_LunarLander.ipynb at main · ounospanas/AIDL_B02](#) with the necessary modifications on the preprocessing step, network and replay buffer

- Lower the LR, that worked for me ☺

- Use a seed for the environment to simplify the game and be deterministic. Will converge faster

- It will probably took me > 5,000 epochs to reach over 400 (around 5 colab hours)

- Reduce a little the batch size or the replay buffer or the input images size
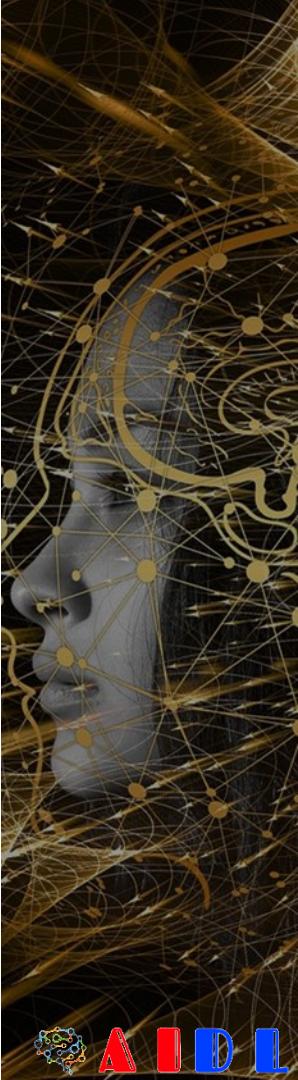
Project Presentation

Example

Present your NN architectures, training hyperparameters and the obtained results using average episode rewards

Remember to deliver 3 not 2

Scores

5 points in total

- Only 1 DQN (+replay buffer, target net): **3** points
  - Remember you should reach a score over 500 at 100 consecutive games

- 2 DQN (+replay buffer, target net): **2** points (1 each)
  - Remember you should reach a score over 400 at 100 consecutive games

Thank You!