Report

Purpose of the project:
To teach a reinforcement learning algorithm how to solve a banana collection game.

Journey:
I started by trying to optimize the solution for the problem without clear direction. I ended up struggling with implementation of prioritized experience replay, dueling DQN, but still not able to optimize the solution.

Therferoe, I clicked the scary button of reset data under the Udacity menu and started from scratch. I reused code from the DQN lesson which implements a Double DQN algorithm with decaying epsilon and random experience replay. I was able to solve with clear and coherent steps the problem in 248 episodes. The model weights are stored under checkpoint.pth.

Model Architecture:
The model consists of three layers of fully connected neurons. The first layer takes a state size and outputs 64 units. The second layer takes 64 units and outputs 64 units. The last layer outputs action through 4 neurons.

Hyperparameters:
Below are hyperparameters:

BUFFER_SIZE = 10000
BATCH_SIZE = 64
GAMMA = 0.99
TAU = .001
LR = .0005
UPDATE_EVERY = 4
Eps decay = .995
Eps start = 1

Future Plans:
Future plan is to use prioritized experience replay with dueling DQN. It is also to implement policy state functions such as A2C or GAE.

Conclusion:
KISS, or keep it simple stupid. I had to remove everything and start with a simplistic model to solve a problem.