# UDACITY DEEP REINFORCMENT LEARNING – PROJECT 3 REPORT "COLLABORATION AND COMPETITION; TENNIS ENVI"

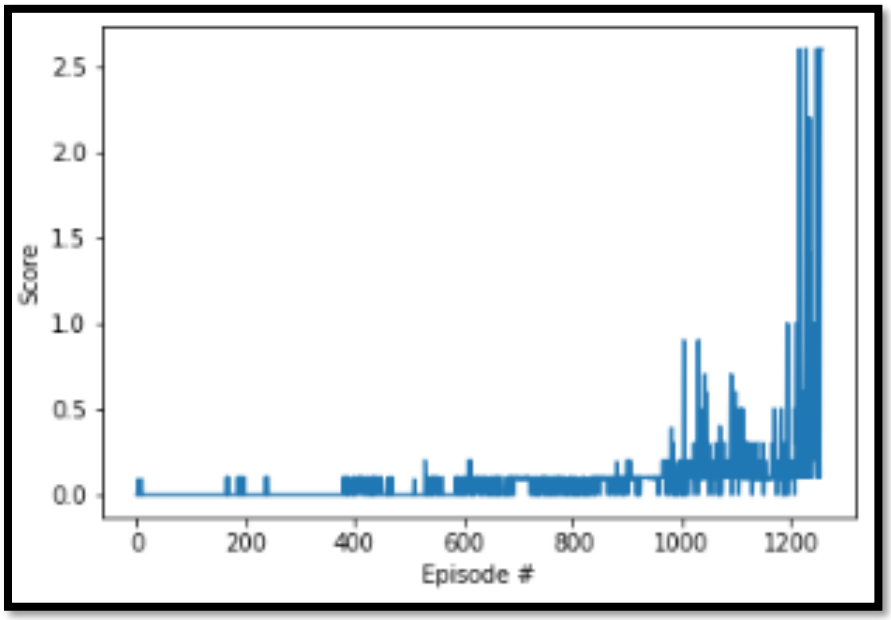
Mahmoud Adel Nour

m.adel.nour@gmail.com

# Report

1. Learning Algorithm
    - Building DDPG network using 3 hidden full connected layers for the actor and 4 hidden full connected layers for the critic, with ReLu activations ending with tanh and the optimizer is ADAM
    - Using Experience replay with batch size 1024 and buffer size 1,000,000 which is shared between actor and critic
    - Using different values of gamma
    - Creating target weights and current(local) weights knowing that the target weights changes with small steps from the current weights to keep soft update concept
    - I tried to use add on noise and it gives better values without noise
    - I tried all the combination of the hidden layer numbers of actor (3,4,5) and critic(3,4) and neurons(64,32,16) and finally reached the best combination for the actor with 3 layers and (64,32,2) neurons and critic with 4 layers and (64,64,32,1) neurons,
    - The current implementation achieved mean score over 100 episodes greater than 0.5 after 1257 episodes of training
    - Hyper parameters:
        - BUFFER_SIZE = 1,000,000 : replay buffer size
        - BATCH_SIZE = 1024 : batch size to start learning over the stored replays
        - GAMMA = 0.99 : discount factor
        - TAU = 0.001 : soft update factor which control the change of the weights
        - LR_ACTOR = 0.0001 : learning rate for the optimizer “ADAM” of the Actor
        - LR_CRITIC = 0.001 : learning rate for the optimizer “ADAM” of the Critic

2. Plot of Rewards



3. Ideas for Future Work
    - Increase the number of layers and agents to learn faster
    - Change the model of the actor and critic
    - To try PPO algorithm