Project Report on "Continuous Control"

Learning Algorithm

I used Deep Deterministic Policy Gradients for this project on teaching robotic arms to move to target positions.

Hyperparameters

I moderately experimented with the hyperparameters as I was able to train the agent. The hyperparameters used are as follows:

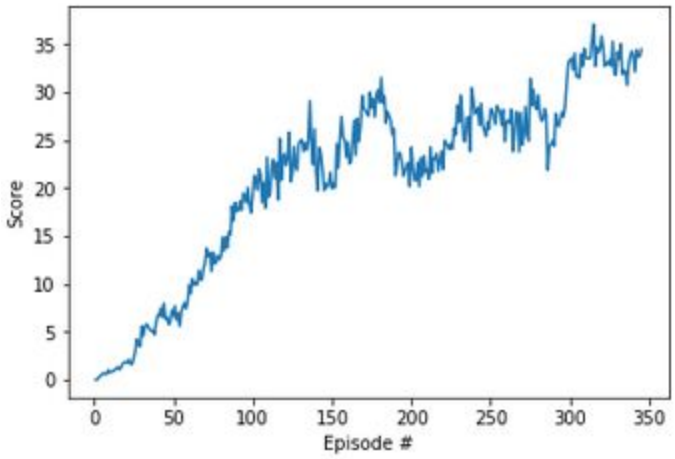
- BUFFER_SIZE = int(1e6) # replay buffer size
- BATCH_SIZE = 1024 # minibatch size
- GAMMA = 0.99 # discount factor
- TAU = 1e-3 # for soft update of target parameters
- LR_ACTOR = 1e-4 # learning rate of the actor
- LR_CRITIC = 1e-3 # learning rate of the critic
- WEIGHT_DECAY = 0.0000 # L2 weight decay

Neural Network Architectures

- Used 2 hidden layers in the actor network. Layer 1 has 256 nodes and layer 2 has 128 nodes. Both the layers have relu activation.
- Critic network has three hidden layers. Layer 1 has 256 nodes, layer 2 has 128 nodes and layer 3 has 64 nodes. All the layers use relu activation.

Plot of rewards

Average score of 30+ was achieved at around 350th episode.



Ideas for future works

Several other algorithms like PPO, D4PG have been specified in the classroom to solve the project. It would be good to experiment with these algorithms in future and see how they perform.