

# Project Report

Deep Deterministic Policy Gradients Algorithm has been used to move a double jointed arm to target positions. As the action space is continuous, I decided to use the Deep Deterministic Policy Gradients Algorithm.

I have used two neural networks, the actor network and the critic network. The actor network contains two hidden layers of 256 units and 128 units respectively. Relu activation has been applied on both the layers.

Critic network also has two hidden layers of sizes 128 units and 64 units with relu activation on both the layers. Batch normalization is also applied on the input to critic network.

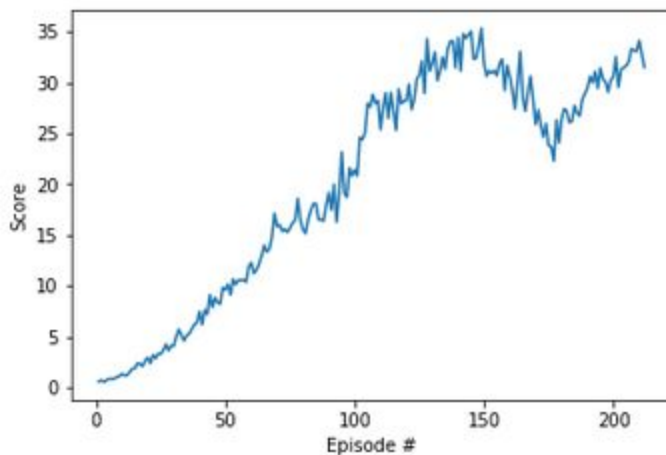
A lot of tuning and tweaking was done with the hyperparameters. Finally, I ended up using the following hyperparameters.

1. Buffer Size -  $1e-6$
2. Batch Size - 1024
3. Gamma - 0.99
4. Tau -  $1e-3$
5. Learning Rates - 0.001 (Actor and Critic)

## Plot of rewards

Episode: 212  
Min Score: 24.03 Max Score: 41.58  
Score: 31.50  
AvgScore: 30.00  
Time: 43.38 secs

```
fig = plt.figure()
ax = fig.add_subplot(111)
plt.plot(np.arange(1, len(scores)+1), scores)
plt.ylabel('Score')
plt.xlabel('Episode #')
plt.show()
```



## Work possible in future

- Changing the actor and critic network sizes in DDPG.
- Trying algorithms other than DDPG such as PPO, A3C, D4PG.