

UDACITY

DEEP REINFORCEMENT LEARNING
NANODEGREE UDACITY

Continuous Control Project

1. Algorithm Details

Deep Deterministic Policy Gradient (DDPG) is a reinforcement learning algorithm designed for environments with continuous action spaces. It combines ideas from Q-learning and policy gradient methods and uses neural networks to approximate both the policy and the value function. Here are the key details and components of the DDPG algorithm

2. Network Architecture

A fully connected neural network with:

Input layer: 33 units

Hidden layers: Two hidden layers with 128 units each, ReLU activation

Output layer: 4 units

3. Hyperparameters

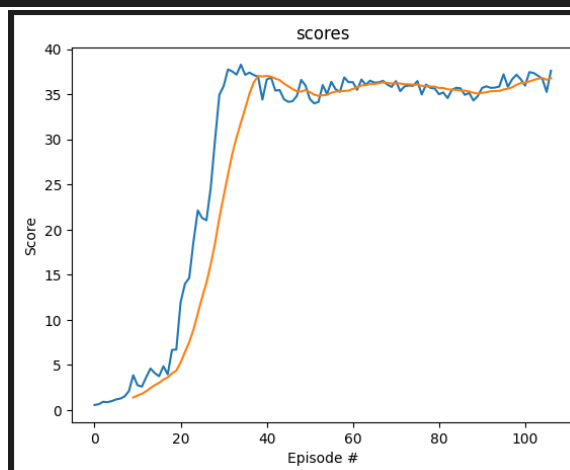
The following hyperparameters were used:

<code>BUFFER_SIZE = int(1e5)</code>	# replay buffer size
<code>BATCH_SIZE = 128</code>	# minibatch size
<code>LR_ACTOR = 1e-3</code>	# learning rate of the actor
<code>LR_CRITIC = 1e-3</code>	# learning rate of the critic
<code>GAMMA = 0.99</code>	# discount factor
<code>TAU = 1e-3</code>	# for soft update of target parameters
<code>WEIGHT_DECAY = 0</code>	# L2 weight decay

4. Plot of Rewards

I needed 7 episodes to solve the environment:

```
Episode 100   Average Score: 27.83
Episode 107   Average Score: 30.34
Environment solved in 7 episodes!   Average Score: 30.34
```



5. Ideas for Future Work

To improve convergence speed, the developments covered in the D4PG course can be used to help reduce overestimation of action values