

# P2- Continues Control: Reacher

For solving this environment, I used DDPG algorithm for training neural network agent and predicting action against each different state.

## Learning from this project

Input of the DDPG is game state vector from Unity Agent and output of the DDPG is action space for different state

### a. Key Point about DDPG

- **Actor**
  1. Fully-connected layer - input: 33 (state size) output: 128
    - Linear Neural Network with followed by ReLU activation function
  2. Batch Normalization
    - To stabilize neural network, we used this. Please find more detail about it from [here](#)
  3. Hidden Fully-connected layer - input: 128 output: 128
    - Linear Neural Network with followed by ReLU activation function
  4. Fully-connected layer - input: 128 output: 4 (action size)
    - Linear Neural Network with followed by Tanh activation function
  5. Maximum steps per episode: 1000
  6. Update neural network after each step and batch size of 128
  7. Discount factor: 0.99
- **Critic**
  1. Fully-connected layer - input: 33 (state size) output: 128
    - Linear Neural Network with followed by ReLU activation function
  2. Batch Normalization
    - To stabilize neural network, we used this. Please find more detail about it from [here](#)
  3. Hidden Fully-connected layer - input: 128 output: 128
    - Linear Neural Network with followed by ReLU activation function
  4. Fully-connected layer - input: 128 output: 4 (action size)
    - Linear Neural Network with followed by Tanh activation function
  5. Maximum steps per episode: 1000
  6. Update neural network after each step and batch size of 128
  7. Discount factor: 0.99

- **Added Noise** to modify action against state to add some twist.
- **Replay Buffer Size:** 1e5
- **Why Batch Normalization is used?**
  - To reduce impact of very large weight or outlier data on next level neural network and Stabilize neural network we used this. Find more details [here](#)
- **Why ReLU is used at hidden layer?**
  - The ReLU is the most used activation function in the world right now. Since, it is used in almost all the convolutional neural networks or deep learning.
  - Relu has range between 0 to infinity
- **Why Tanh is used at last layer?**
  - tanh is also like logistic sigmoid but better. The range of the tanh function is from (-1 to 1). tanh is also sigmoidal (s - shaped).
  - tanh and logistic sigmoid activation functions are used in feed-forward nets.
  - The tanh function is mainly used classification between two classes.

## Plot of Rewards

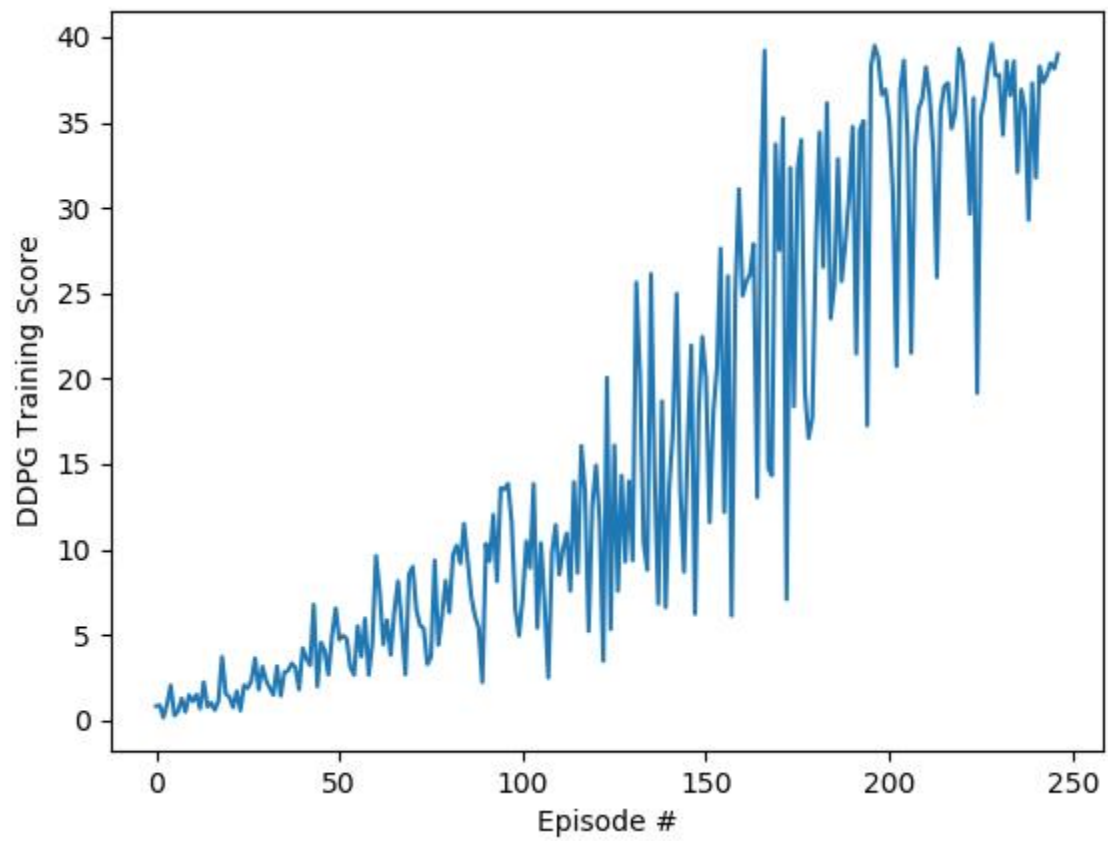
Plotting Average Score against Episode of the game.

### DDPG

#### 1. Training Phase

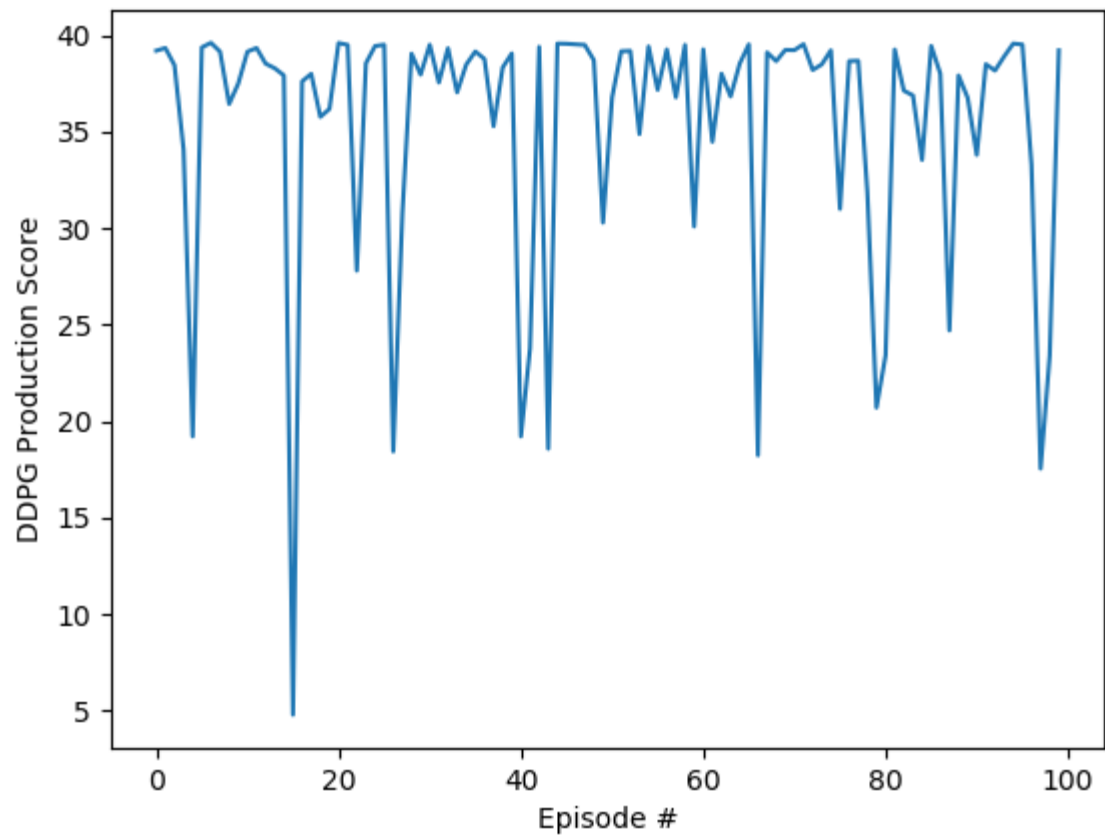
---

```
Episode 50      Average Score: 2.18 Score: 2.18
Episode 100     Average Score: 4.56 Score: 4.56
Episode 150     Average Score: 9.67e Score: 9.67
Episode 200     Average Score: 19.33 Score: 19.33
Episode 247     Score: 39.00      Average Score: 30.02
Environment solved in 147 episodes!      Average Score: 30.02
```



## 2. Production Phase

Episode 50      Average Score: 35.38    Score: 35.38  
Episode 100     Average Score: 35.48    Score: 35.48



### Future Implementations

1. DDPG training with multiple agents
2. Changing neural network size
3. Try different kind of activation function
4. Tune Hyper parameters for better learning