Project 2: Continuoues Control

Featuring the reacher environment

# Learning Algorithm

**Deep Deterministic Policy Gradient (DDPG)**

Deep-Deterministic-Policy-Gradient is an Actor-Critic algorithm which is model-free as well as off-policy. The method consists of two main components. The Actor and the Critic. Both are represented by Neural Networks. It uses concepts from both the Deep-Q-Network (DQN) as well as policy methods. By combining elements of both, DDPQ (or actor-critic models in general) are able to apply the concepts of DQNs on continuous action spaces through application of policy evaluation and noise, while also improving the stability and robustness of policy methods.

**The Actor** is represented by a neural network (NN). It is used to evaluate the best action to take considering the current state of the environment, as calculated by the current policy.

**The Critic** is also represented by a NN and is used to improve the learning of the actor. To accomplish this, the critic is also trained on the same pool of data as the actor but is then used to calculate the loss of the current actor model. This improves the trajectory of the learning.

## Parameters:

## Components:

## Functions:

# Training

## Performance:

## Challenges:

# Plot of rewards:

# Ideas for Future Work