For this project, I implemented the DDPG Actor-Critic algorithm for continuous action spaces. It was the right algorithm to implement due to it’s combination of policy gradient and value function approximation. The combination of the two lead to better stability and results than either algorithm alone. Below are the steps for implementation.

1. Initialize replay memory buffer
2. Initialize local actor and critic network as pytorch neural nets
3. Implement a soft-weights copy function
4. Train the actor and the critic

Ideas for future work:

1. Prioritized experience replay
2. Asynchronous Actor Critic