1. Learning Algorithm:

The source code and weights of a trained agent are placed in the main directory. The algorithm is implemented following the next guidelines:

- The reinforcement learning algorithm is the Double DQN (DDQN)

- Hidden layers are composed of State -> 64 -> LeakyReLU -> 64 -> LinearReLU -> Action. The optimizer is ADAM.

- Experience Replay is also implemented. In this technique, DDQN model is trained by mini-batch from a replay buffer.

- Agent selects next action based on Epsilon Greedy. The value of epsilon is set initially to 1, and decreases at a rate of epsilon_decay = 0.95 with time until 0.000001.
  The parameters just mentioned can be of course changed. These have been empirically found to work well for a rapid convergence.

2. Agent Results:

The following plot of rewards shows the training results. At Episode 278, agent performance met the criteria and stopped training. (mean scores of last 100 episodes is above +13).