**Project B**

2017202109 刘洋凡

1. 实验目的

通过机器学习算法进一步强化cartpole，认识机器学习的具体实现

1. 环境配置

Anaconda、pip install gym

gym的核心接口是Env，作为统一的环境接口。Env包含下面几个核心方法：

1、reset(self):重置环境的状态，返回观察。

2、step(self,action):推进一个时间步长，返回observation，reward，done，info

3、render(self,mode=’human’,close=False):重绘环境的一帧。默认模式一般比较友好，如弹出一个窗口。

1. 算法思路

与上一个实验相比，除了考虑当前决策收益之外，DQN还考虑了对未来收益的影响，并通过多次训练找到接近最优的决策

1. 实验结果

