

Домашнее задание 3

- Дискретизировать (можно использовать `numpy.round()`) пространство состояний и обучить Агента решать CartPole-v1, Acrobot-v1, MountainCar-v0, или LunarLander-v2 (одну на выбор) методами MC, SARSA и Q-Learning. Изобразить скорость обучения этих методов и метода Deep Cross Entropy на графиках. Прислать код.