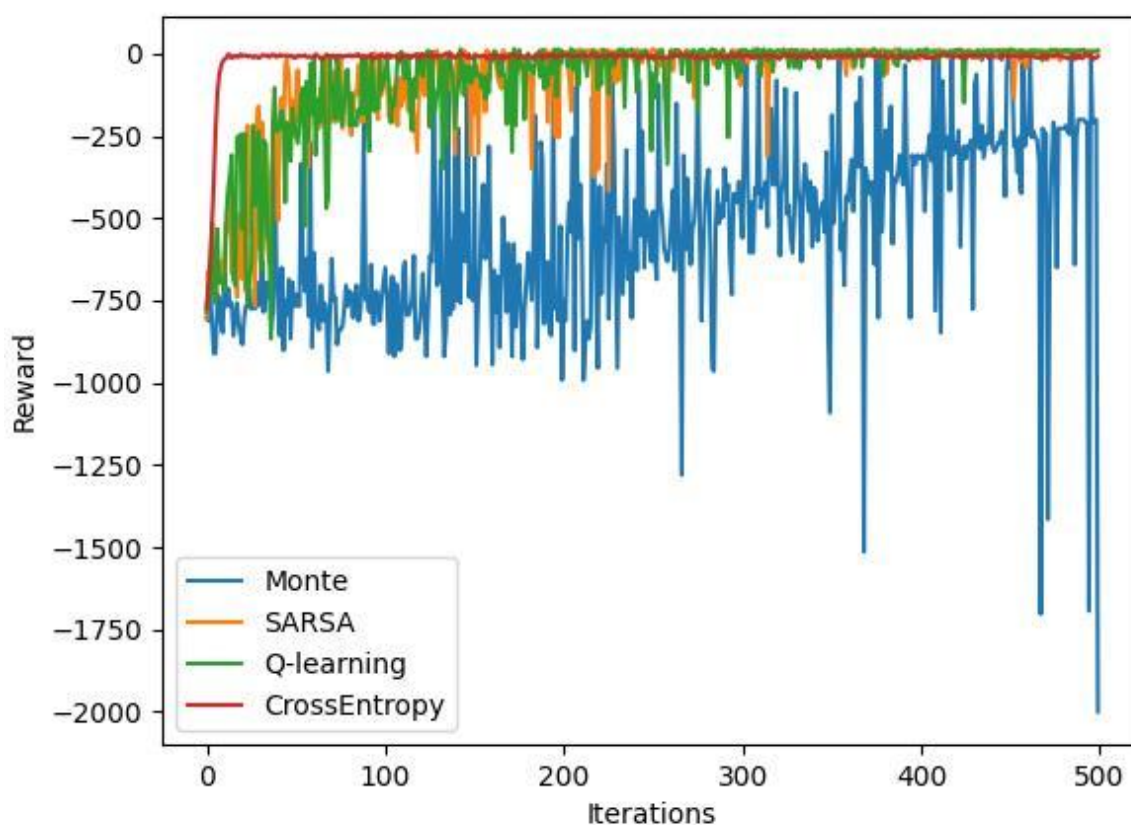


## Отчет по домашнему заданию.

1. Реализовать Q-Learning и сравнить его результаты с реализованными ранее алгоритмами: Cross-Entropy, Monte Carlo, SARSA в задаче Taxi-v3. Для сравнения как минимум нужно использовать графики обучения. Причем графики лучше делать относительно количества сгенерированных траекторий.

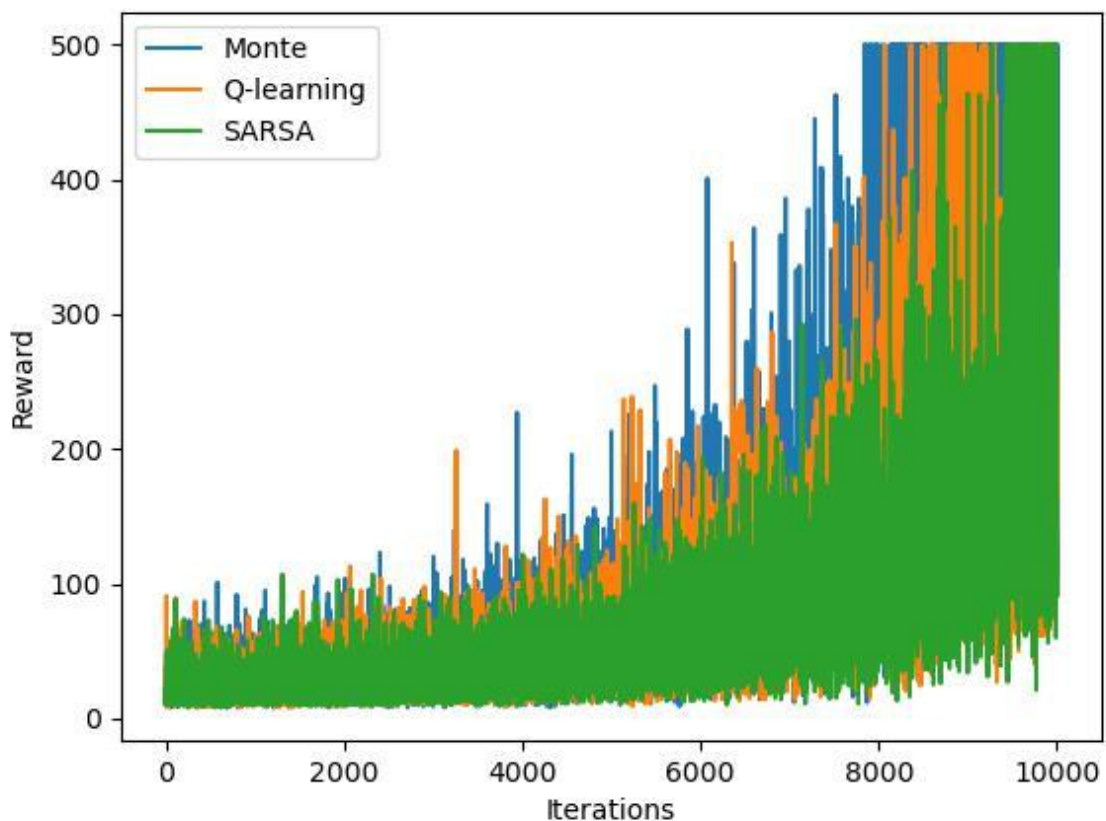
В первом задании, надо было дописать метод Q-Learning. По сути, он отличается одной строкой кода от SARSA. Дальнейший тест в задаче Taxi-v3 отображен на графике:

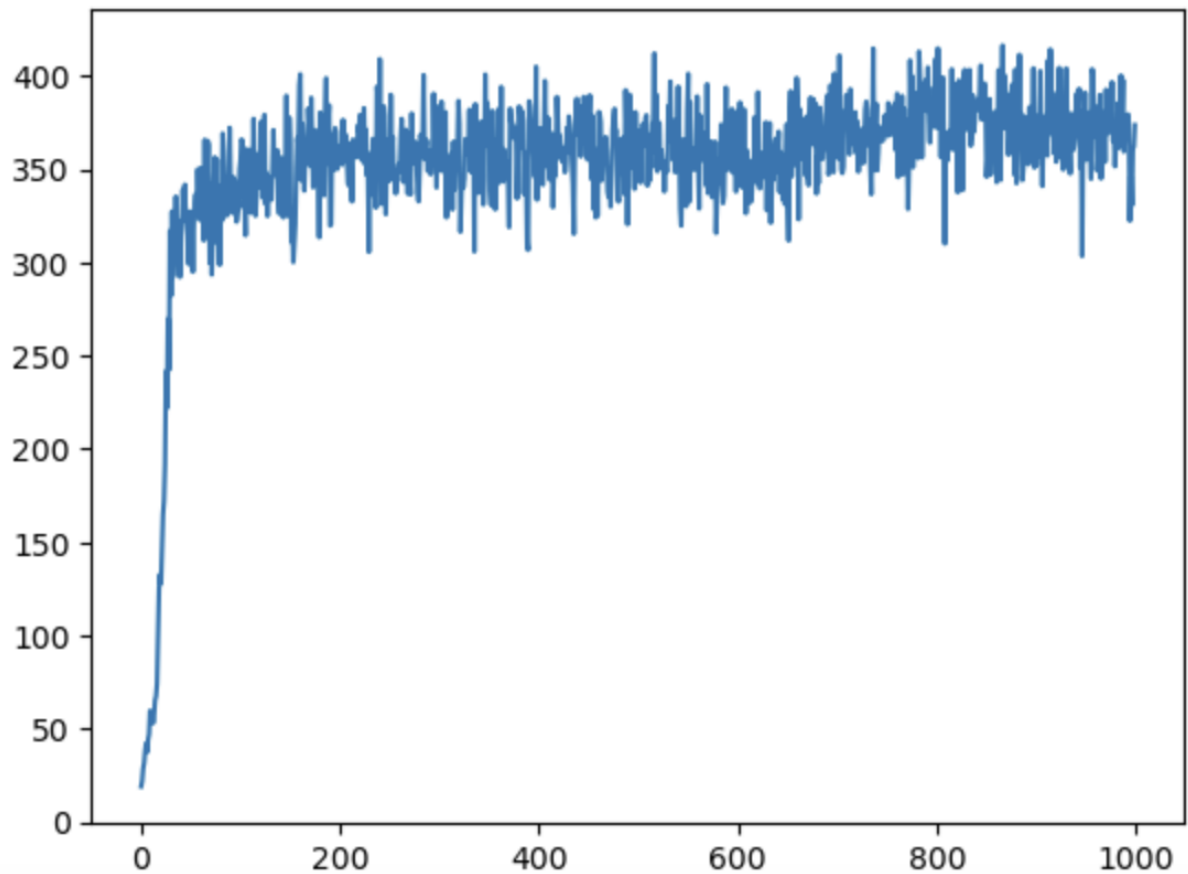


Итог: Видно, что CrossEntropy обучилась лучше, она быстрее всех вышла на плато.

2. Дискретизировать (можно использовать `numpy.round()`) пространство состояний и обучить Агента решать CartPole-v1, Acrobot-v1, MountainCar-v0, или LunarLander-v2 (одну на выбор) методами Monte Carlo, SARSA и Q-Learning. Сравнить результаты этих алгоритмов и реализованного ранее алгоритма Deep Cross-Entropy на графиках.

В этом задании, нужно было проверить, как методы справляются в задаче с НЕ дискретными состояниями. Эта проблема решается дискретизацией состояний, по сути, у каждого параметра состояний есть свои ограничения по значениям. Я эти значения разбил на бины. Решал задачу CartPole-v1, на первом графике будут результаты обучения методов: Monte Carlo, SARSA и Q-Learning. На втором Deep Cross-Entropy.





Итог: В этом задании лучше себя показала модель CrossEntropy.  
Признаки те же, она быстрее всех вышла на плато, + имеет меньшую дисперсию в значениях.