Нестерова М. И.

Код приведен в тетрадке hw1/cds\_rl/scripts/run\_hw1. В этом файле практическая часть и теория.

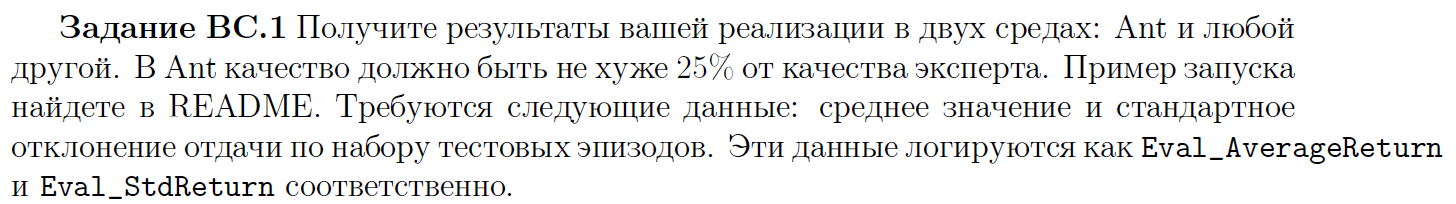
Practical assignment: Imitation learning

Исследуемые среды: Ant и HalfCheetah.

Используются параметры по умолчанию. При этом были изменены следующие параметры: параметры:

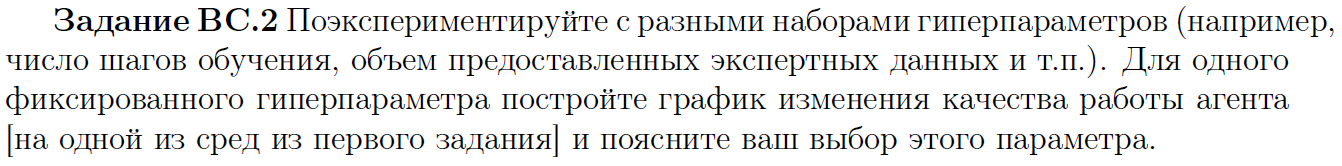
* args.eval\_batch\_size = 10000
* args.ep\_len = 1000
* args.size = 96

# BC1



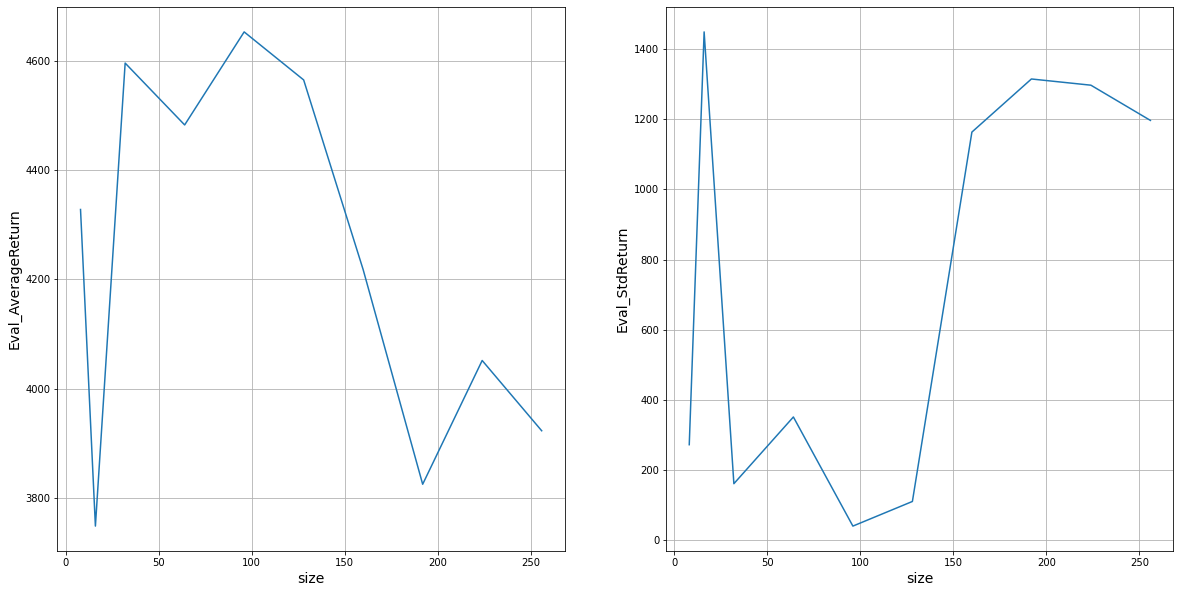
|  |  |  |
| --- | --- | --- |
|  | Ant | HalfCheetah |
| Eval\_AverageReturn | 4653.38 | 3952.64 |
| Eval\_StdReturn | 39.93 | 84.59 |
| Expert \_AverageReturn | 4713.65 | 4205.77 |

# BC2

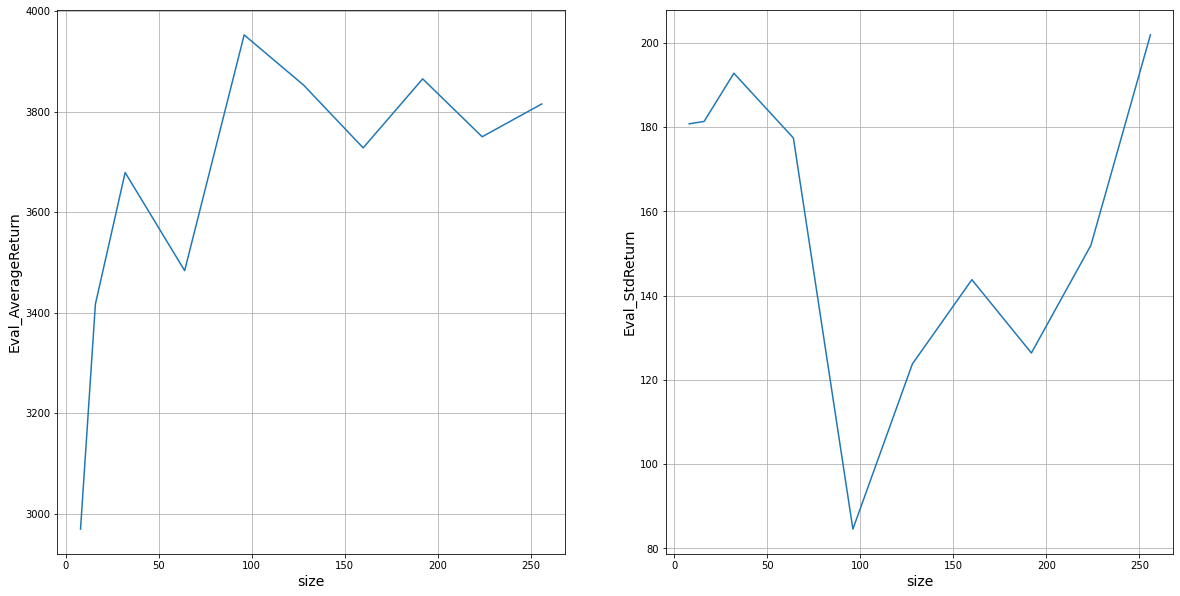


Проверялись гиперпараметры: eval\_batch\_size, ep\_len, size, learning\_rate, train\_batch\_size.

## Ant

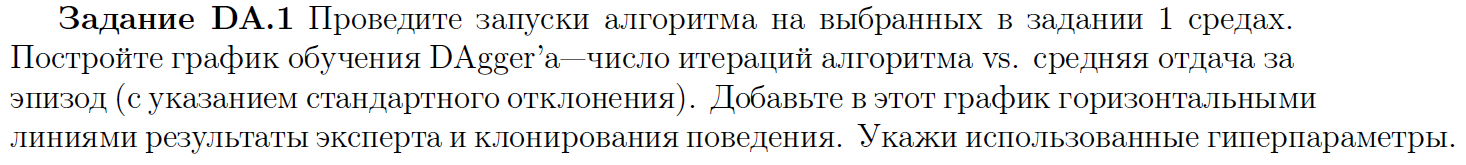


## HalfCheetah

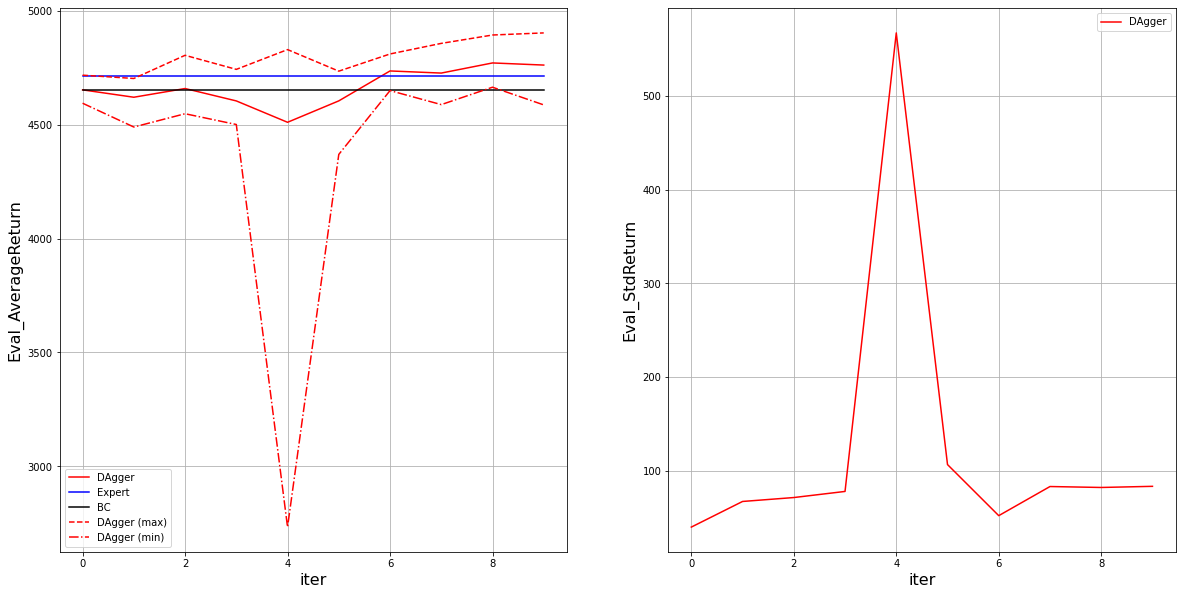


Для Ant и HalfCheetah наилучший результат был при size=96. Был выбран этот параметр, так как он наиболее явно влиял на Eval\_AverageReturn. В блокноте представлены графики с другими параметрами для Ant.

# DAgger



## Ant



## HalfCheetah

