

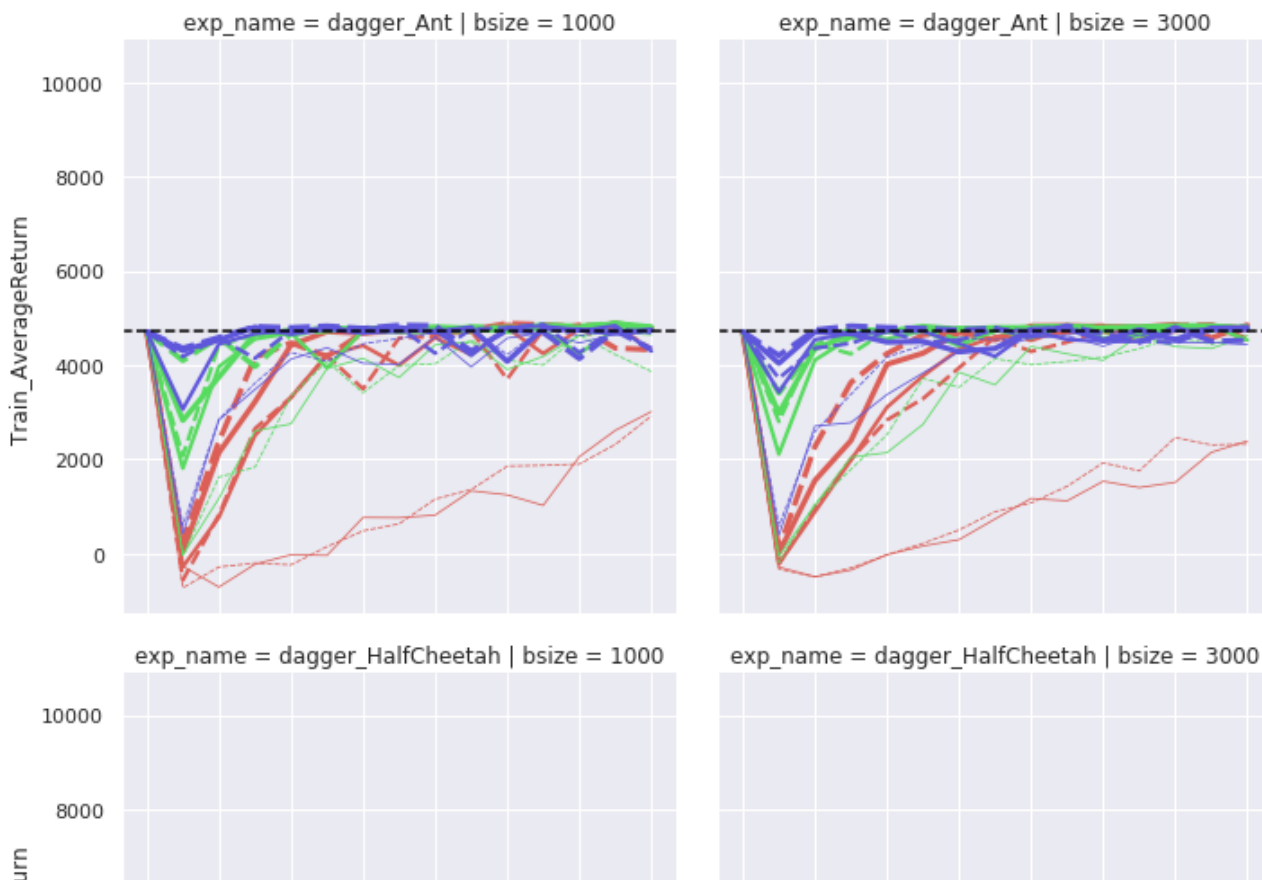
# cs285 hw1

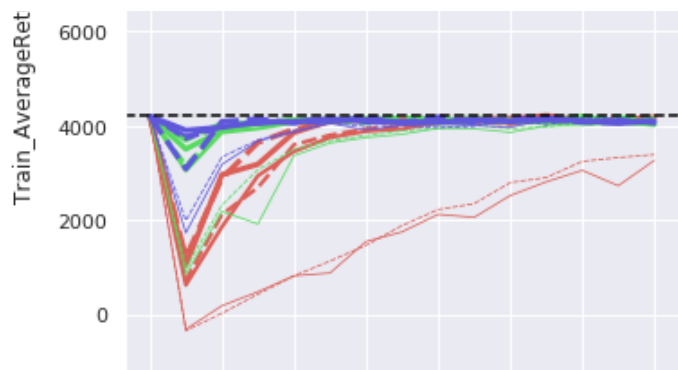
- [cs285 hw1](#)
  - [Config](#)
  - [Result](#)
  - [Analysis](#)
    - [Sensitive hyper-params](#)
    - [Envs](#)

## Config

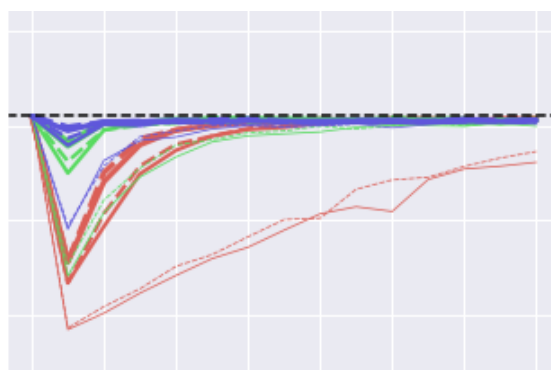
```
exp_name_list = [dagger_Humanoid, dagger_HalfCheetah, dagger_Hopper, dagger_Ant, dagger_Walker2d]
num_agent_train_steps_per_iter_list = [256, 1000, 1500]
batch_size_list = [1000, 3000]
train_batch_size_list = [100, 256]
learning_rate_list = [1e-3, 5e-3, 1e-2]
```

## Result





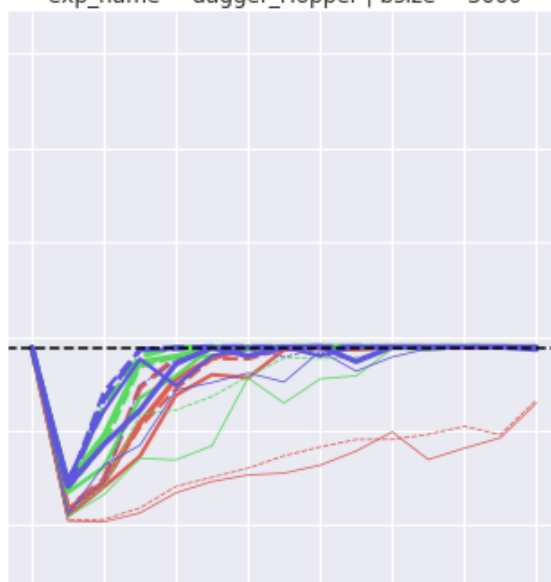
exp\_name = dagger\_Hopper | bsize = 1000



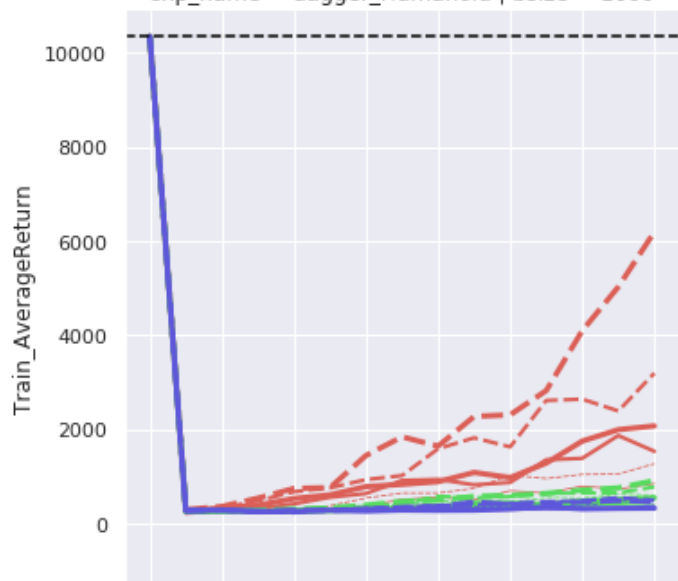
exp\_name = dagger\_Hopper | bsize = 3000



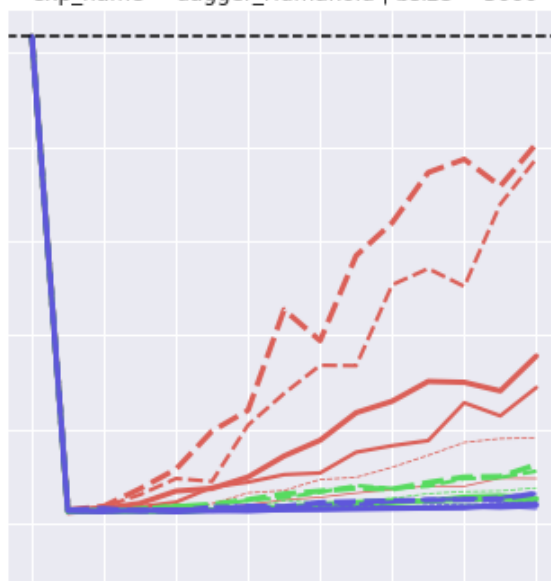
exp\_name = dagger\_Humanoid | bsize = 1000



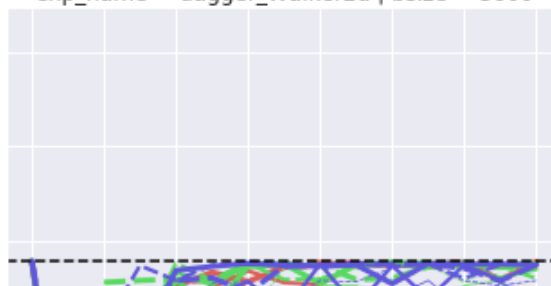
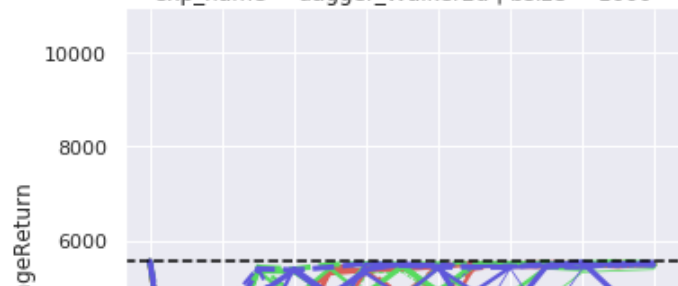
exp\_name = dagger\_Humanoid | bsize = 3000

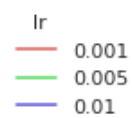
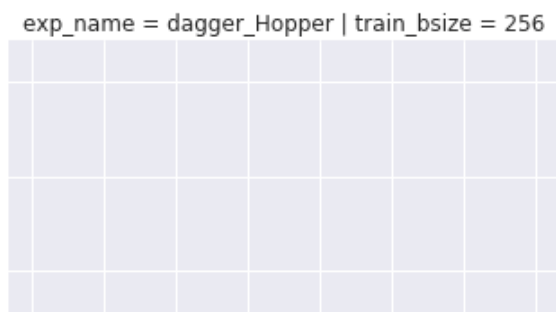
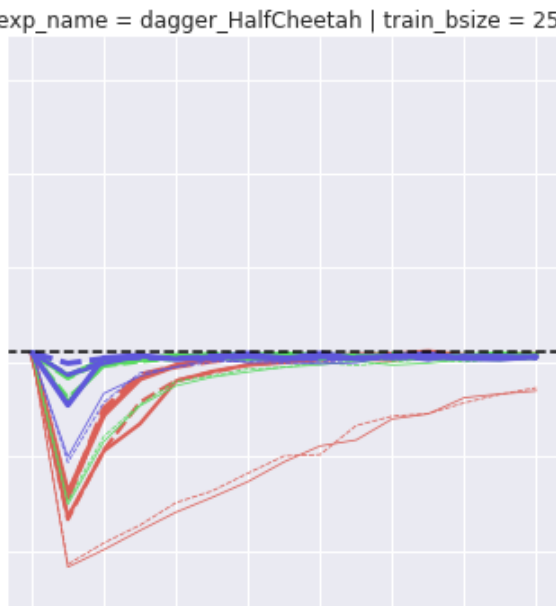
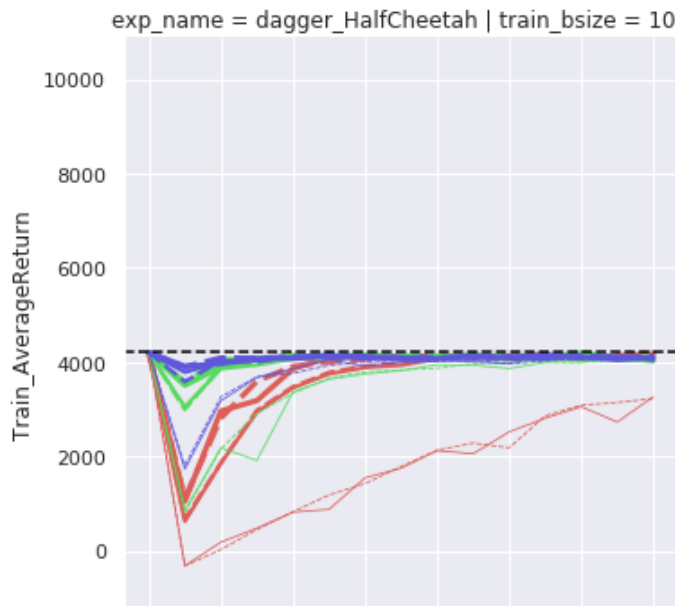
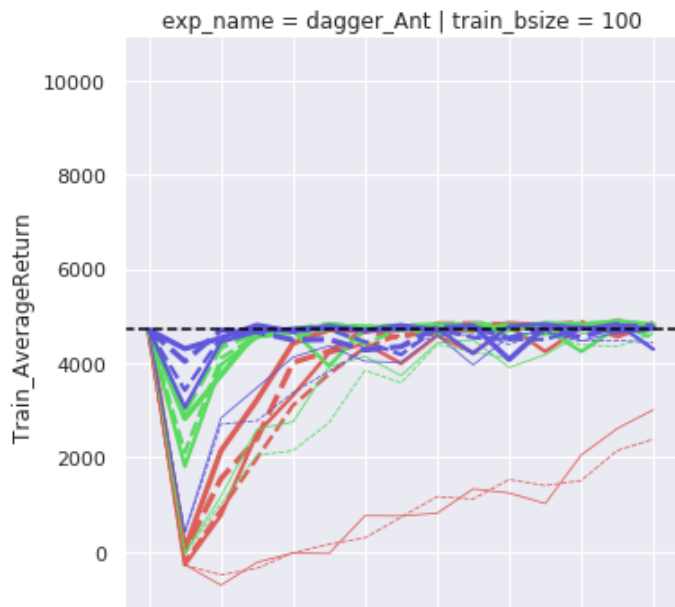
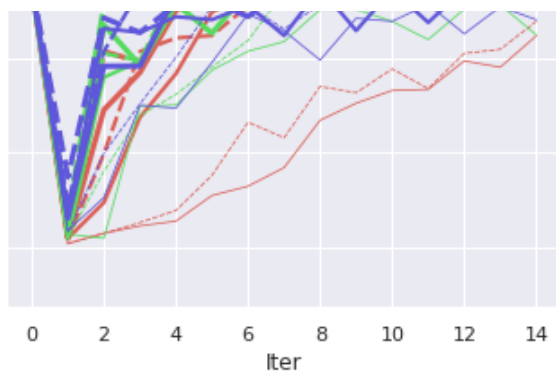
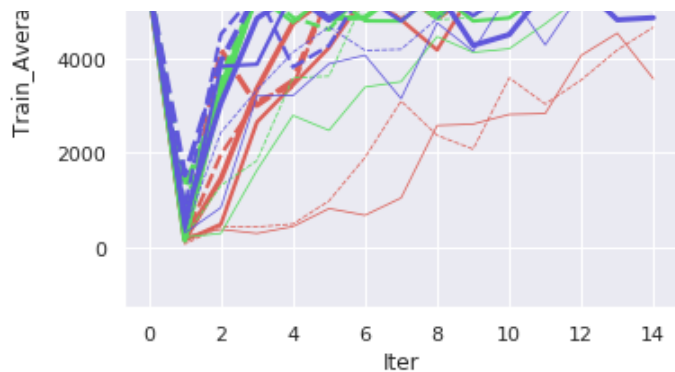


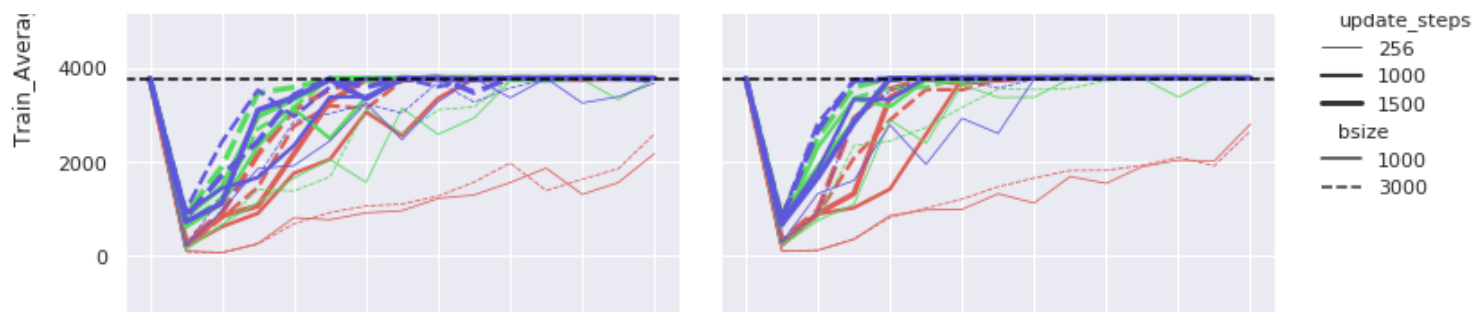
exp\_name = dagger\_Walker2d | bsize = 1000



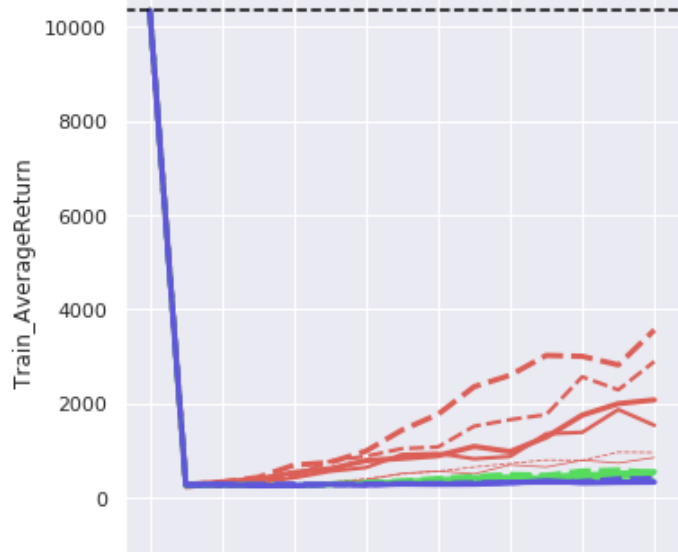
exp\_name = dagger\_Walker2d | bsize = 3000



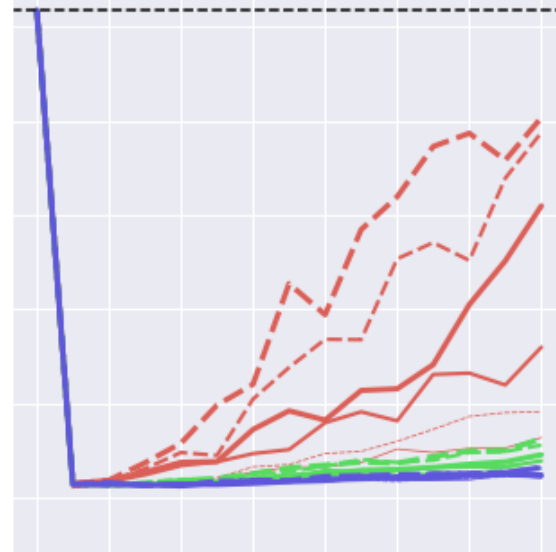




exp\_name = dagger\_Humanoid | train\_bsize = 100



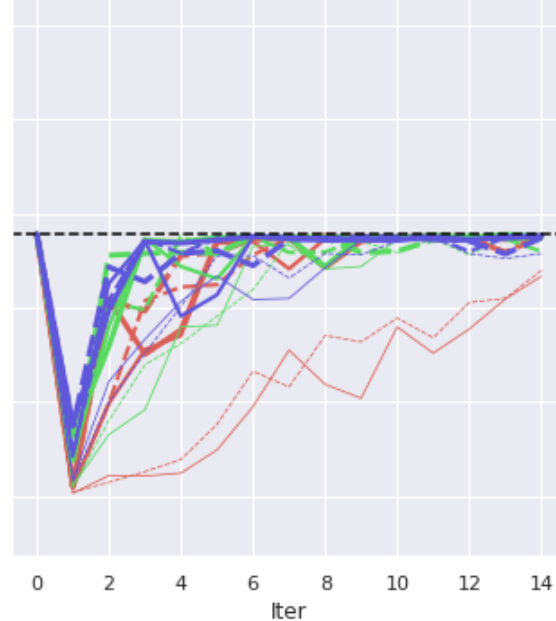
exp\_name = dagger\_Humanoid | train\_bsize = 256



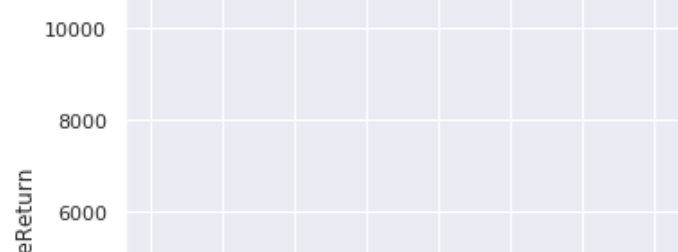
exp\_name = dagger\_Walker2d | train\_bsize = 100



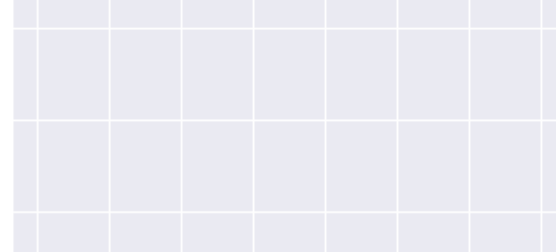
exp\_name = dagger\_Walker2d | train\_bsize = 256

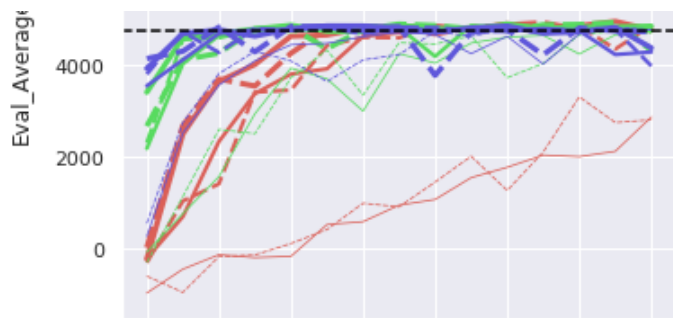


exp\_name = dagger\_Ant | bsize = 1000

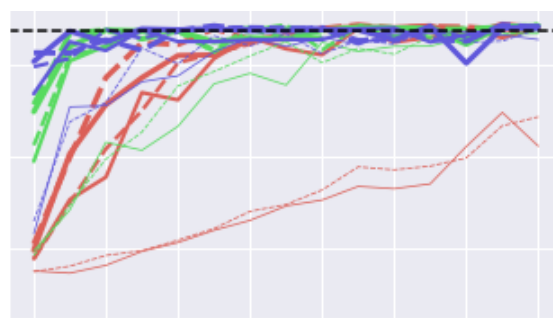


exp\_name = dagger\_Ant | bsize = 3000

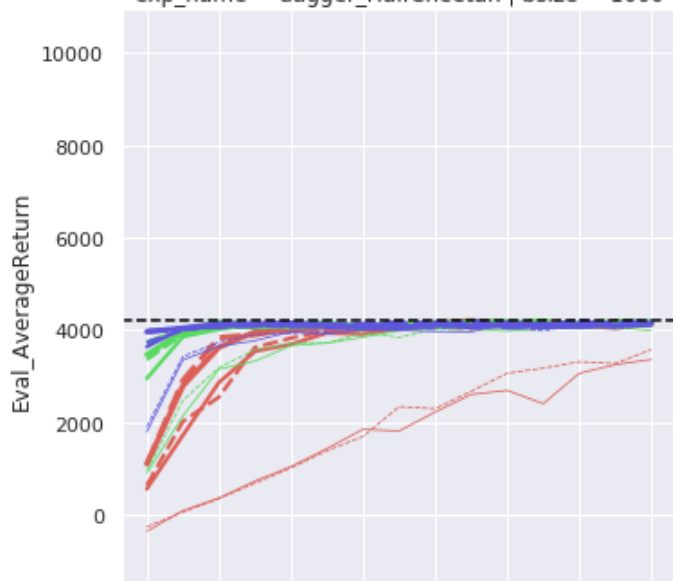




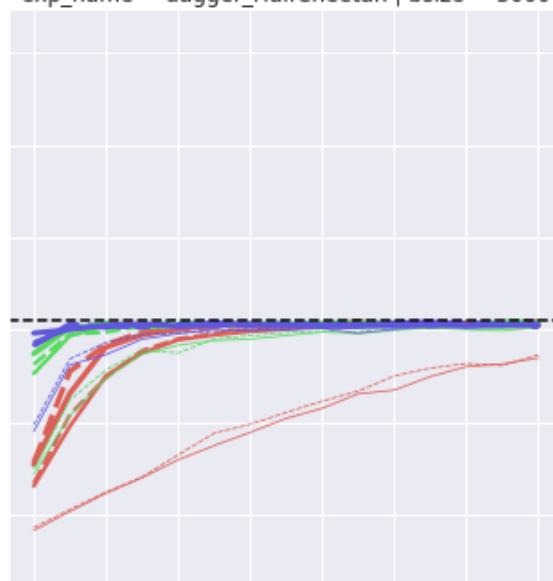
exp\_name = dagger\_HalfCheetah | bsize = 1000



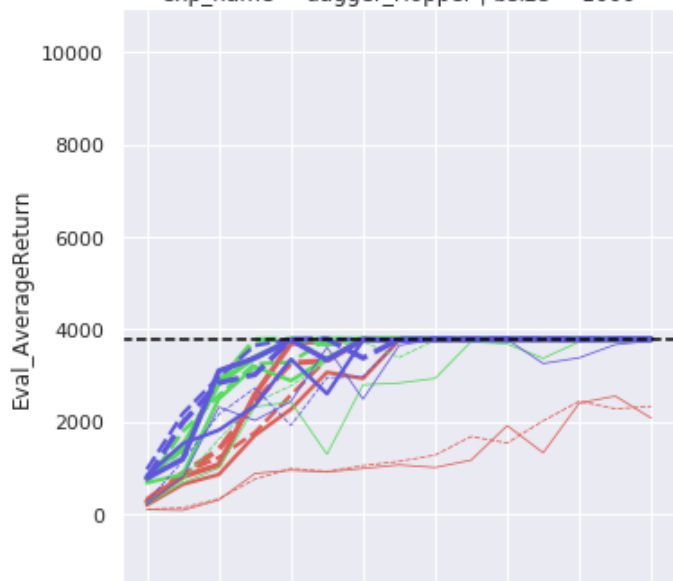
exp\_name = dagger\_HalfCheetah | bsize = 3000



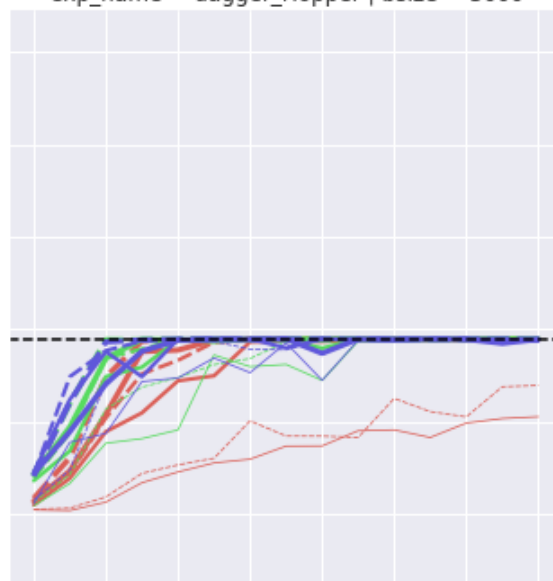
exp\_name = dagger\_Hopper | bsize = 1000



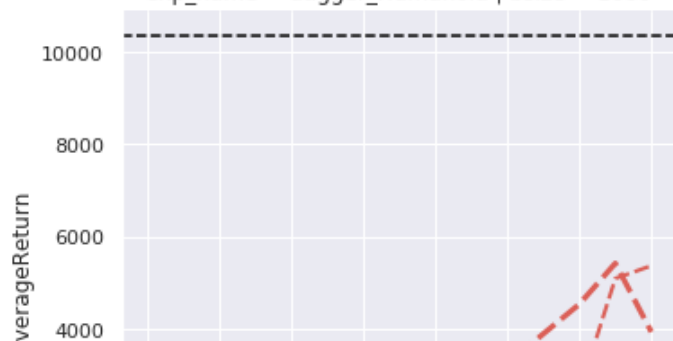
exp\_name = dagger\_Hopper | bsize = 3000



exp\_name = dagger\_Humanoid | bsize = 1000

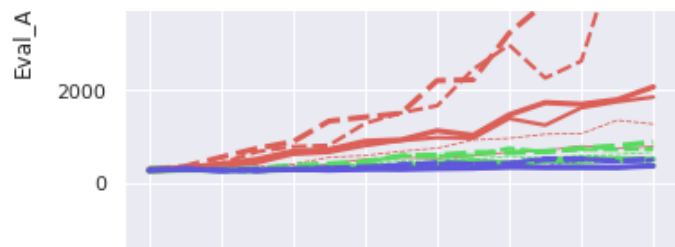


exp\_name = dagger\_Humanoid | bsize = 3000

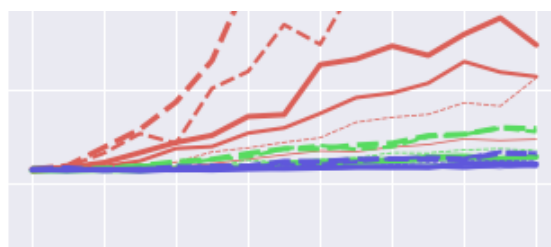


verageReturn

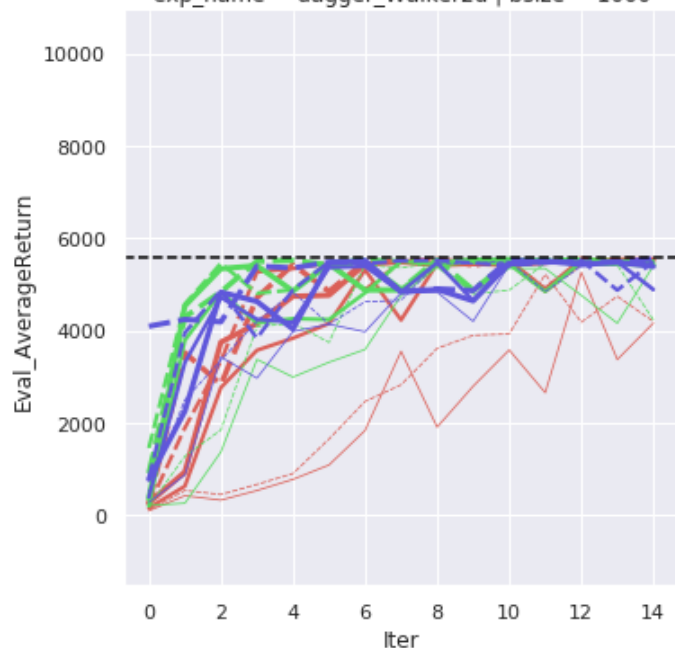




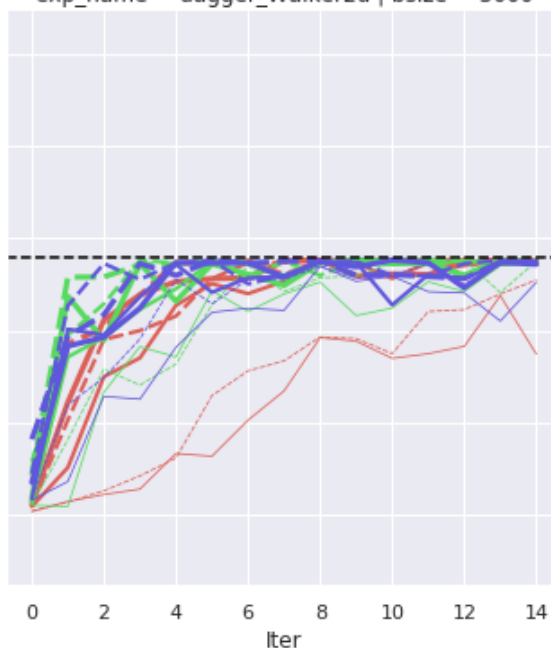
exp\_name = dagger\_Walker2d | bsize = 1000



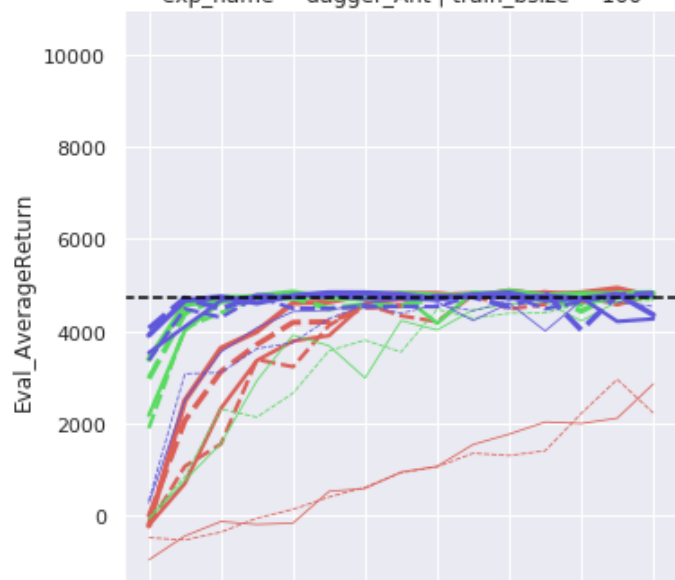
exp\_name = dagger\_Walker2d | bsize = 3000



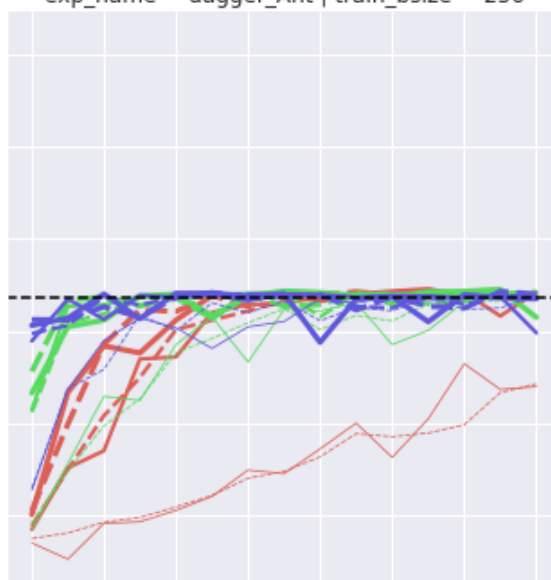
exp\_name = dagger\_Ant | train\_bsize = 100



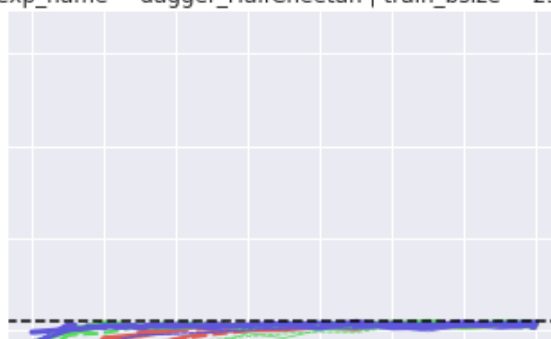
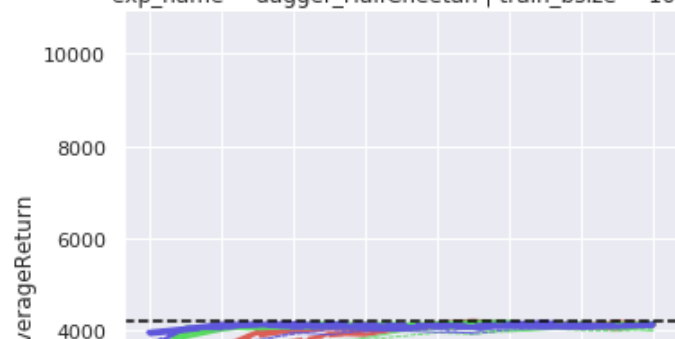
exp\_name = dagger\_Ant | train\_bsize = 256

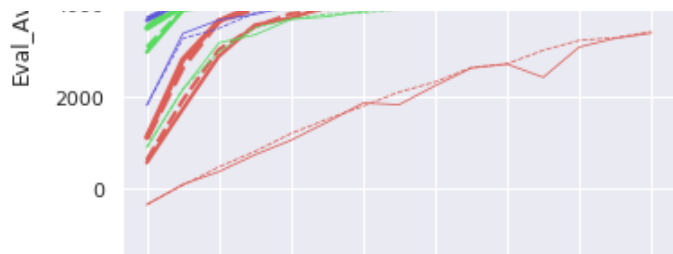


exp\_name = dagger\_HalfCheetah | train\_bsize = 100

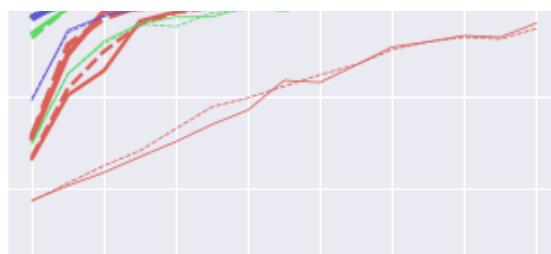


exp\_name = dagger\_HalfCheetah | train\_bsize = 256

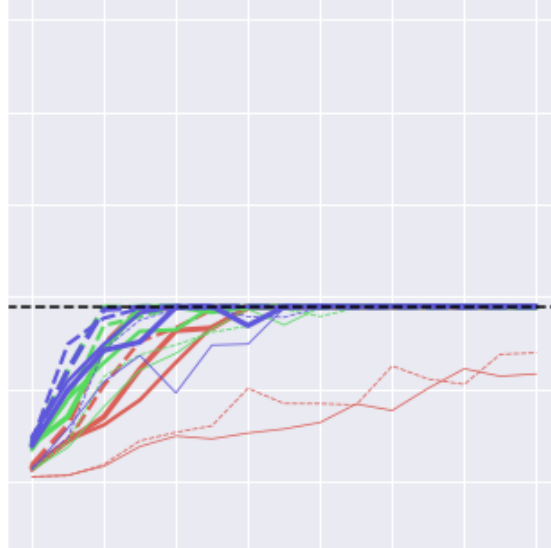
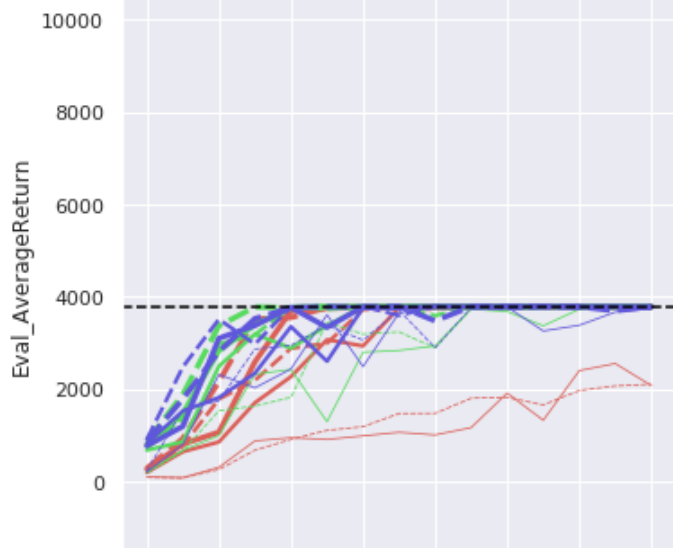




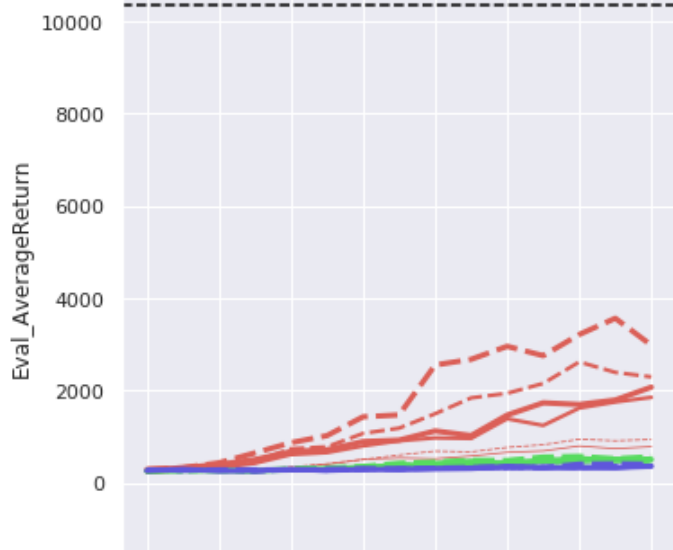
exp\_name = dagger\_Hopper | train\_bsize = 100



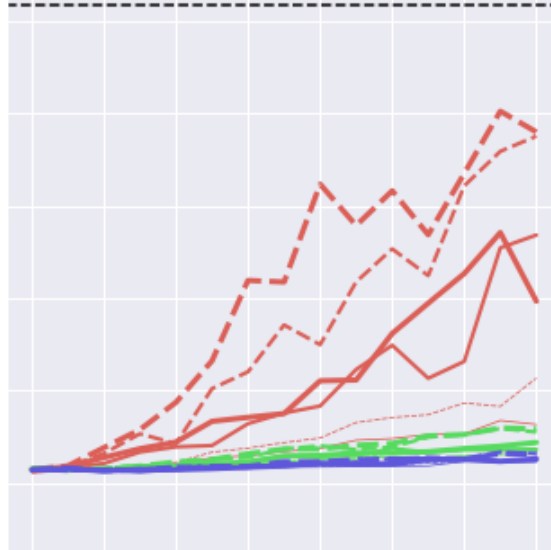
exp\_name = dagger\_Hopper | train\_bsize = 256



exp\_name = dagger\_Humanoid | train\_bsize = 100



exp\_name = dagger\_Humanoid | train\_bsize = 256



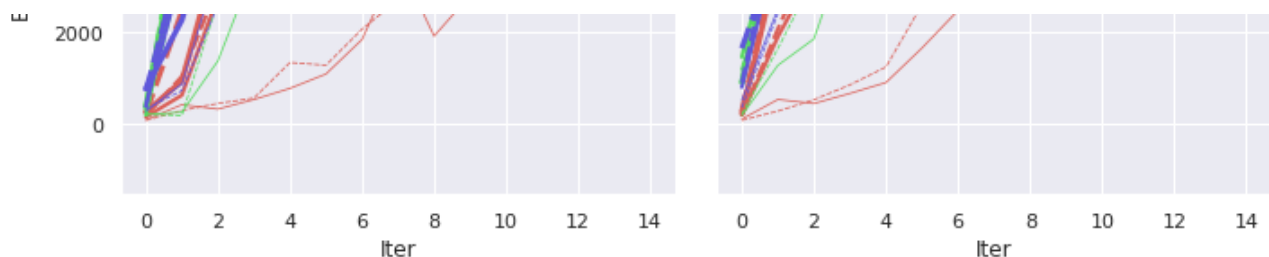
exp\_name = dagger\_Walker2d | train\_bsize = 100



exp\_name = dagger\_Walker2d | train\_bsize = 256







## Analysis

### Sensitive hyper-params

- **lr**: 不仅会影响收敛的速度, 也会影响最终收敛后的结果. 5个实验中, dagger\_Humanoid中lr=1e-3明显好于5e-3和0.01, 其他4个都是5e-3和0.01明显好于1e-3. lr=1e-3时, 虽然相对来说小了一些,但是如果update\_steps设置大一些的话最终收敛的结果会差不多.
- **update\_steps**: 影响收敛的速度, 不影响最终结果, 更大更容易收敛.
- **batch\_size**(collectin size in training): 1000与3000差别不大, 但是一般3000会稳定好于1000一点(在大部分实验中, 除了dagger\_Humanoid, 在该实验中3000会明显好于1000).
- **train\_batch\_size**: 除了dagger\_Humanoid(256明显好于100), 其他都无明显差别

## Envs

最优参数配置:

- dagger\_Humanoid
  - lr: 1e-3 >> 5e-3, 1e-2
  - update\_steps: 1000~1500
  - batch\_size: 3000 >> 1000
  - train\_batch\_size: 256 >> 100
  - dagger agent距离expert还有不小的差距
- dagger\_HalfCheetah, dagger\_Hopper, dagger\_Ant, dagger\_Walker2d
  - lr: 5e-3
  - update\_steps: 1000
  - batch\_size: 1000
  - train\_batch\_size: 100
  - dagger agent基本能达到expert的性能.

总的来说, bc agent都很差.