

# Answers

Kevin Huang

November 2022

## 1 4.5 Results

These values were for models trained using the arguments shared in [Ed post 274](#).

- **CartPole, BC:** Mean reward incurred is 61.5.
- **CartPole, DAGGER:** Mean reward incurred is 79.3.
- **MountainCar, BC:** Mean reward incurred is -105.1.
- **MountainCar, DAGGER:** Mean reward incurred is -97.6.

## 2 4.6 Explanations

- **How does BC perform with a smaller dataset? Is this expected? This can be done by setting the argument `num_dataset_samples` in the arguments.**

When we perform BC with a smaller dataset, we expect it to improve in terms of performance because BC is a form of supervised learning where models perform better with more data because it has more training points to work with. When looking at CartPole

- With `bc_epochs` set to 10:
  - \* With `num_dataset_samples` set to 10, reward is 11.
  - \* With `num_dataset_samples` set to 100, reward is 69.9.
  - \* With `num_dataset_samples` set to 1000, reward is 137.4.
- With `bc_epochs` set to 100:
  - \* With `num_dataset_samples` set to 10, reward is 39.5.
  - \* With `num_dataset_samples` set to 100, reward is 61.5.
  - \* With `num_dataset_samples` set to 1000, reward is 200.
- With `batch_size` set to 128:
  - \* With `num_dataset_samples` set to 10, reward is 17.5.

- \* With `num_dataset_samples` set to 100, reward is 57.
- \* With `num_dataset_samples` set to 1000, reward is 200.

As we can see, regardless of other parameter values, increasing the `num_dataset_samples` resulted in higher rewards and lower loss.

- **How do BC and DAGGER perform with the same starting dataset? How quickly does each converge loss-wise?**

I noticed that in all cases, increasing `num_dataset_samples` resulted in improved rewards incurred during testing. However, something strange that I noticed and thought would be useful to take note of is that for the MountainCar training, if the dataset samples were too low, the models would actually converge to 0 loss (which is an example of over fitting). Additionally, MountainCar loss plots in general looked very interesting where increasing `num_dataset_samples` did NOT extend the trend seen in lower counts. I believe also that generally BC and DAGGER converged at similar rates (perhaps DAGGER converged slightly faster) and both performed very similarly.

I noticed that in all cases, `lr` at very low values and very high values did not perform well. This is an example of overfitting/underfitting the data where because the learning rate was too low or too high, the model either over compensated or did not sufficiently learn the trends in the data. However, there is once again a slight difference in reporting between MountainCar and CartPole. Specifically, when overfit, the loss plot shows that the model converges to 0. However, for MountainCar, the loss plots show the models converging to about 1. Finally, for `lr`, BC converged more quickly.