

Question 2 - behavioral cloning

Question 2.2 - behavioral cloning

Architecture of the neural net

- Normalization of inputs (aka Z-score) (rmk: std+=1e-6 to avoid divide by 0)
- 2 `Dense` hidden layer, `tanh` activation
- `Dense` output layer
- `Adam` optimizer, `batch_size=256`, `validation_split=10%`, `verbose=2`
- `Adam` hyperparams to be learnt and the default value `learning_rate=0.001`, `epochs=10`
- each rollout stop until `max_steps = env.spec.timestep_limit`

Question 2.3 - BC sensitivity analysis to hyperparameters

learning rate

training epochs

Question 3 - DAgger

Question 3.2 - performance against behavioral cloning

Architecture of the neural net

- Normalization of inputs (aka Z-score) (rmk: std+=1e-6 to avoid divide by 0)
- 2 `Dense` hidden layer, `tanh` activation
- `Dense` output layer
- `Adam` optimizer, `batch_size=256`, `validation_split=10%`, `verbose=2`, `learning_rate=0.0001`, `epochs=5`
- each rollout stop until `max_steps = env.spec.timestep_limit`