

Notes about the results

Car Environment

NFQ (offline)

- *Setup 1* :
 - with 10 episodes and 1 data gathering episode per episode, $lr=0.001$, $\gamma=0.99$ -> first positive reward when simulating in 17 steps
 - model : `nfq_car_model_10_ep.pkl`

DQN (online)

- *Setup 1* :
 - with 10 episodes, internal max steps of 200 to compute the TD-error, $lr=0.001$, $\gamma=0.99$
-> first positive reward when simulating in 23 steps
 - model : `dqn_car_model_10_ep.pth`

Pendulum Environment
