policy\_kwargs = dict(activation\_fn=th.nn.Tanh, net\_arch=[dict(pi=[128, 128], vf=[128, 128])])

Hyperparameter 조합

이솔

n\_steps: 4096

batch\_size: 64

learning\_rate: lin(0.00025)

n\_epochs =10

n\_steps: 4096

batch\_size: 64

learning\_rate: lin(0.00025)

clip\_range=0.1

정엽

n\_steps: 4096

batch\_size: 64

learning\_rate: lin(0.0007)

ent\_coef: 0.0

n\_steps: 4096

batch\_size: 64

learning\_rate: lin(0.0007)

clip\_range = 0.1

n\_epochs =10

진성

n\_steps: 4096

batch\_size: 64

learning\_rate: lin(0.00025)

n\_epochs =20

n\_steps: 4096

batch\_size: 64

learning\_rate: lin(0.00025)

verbose: 2

n\_epochs =20

target\_kl=0.01