RL OpenAI gym assignment
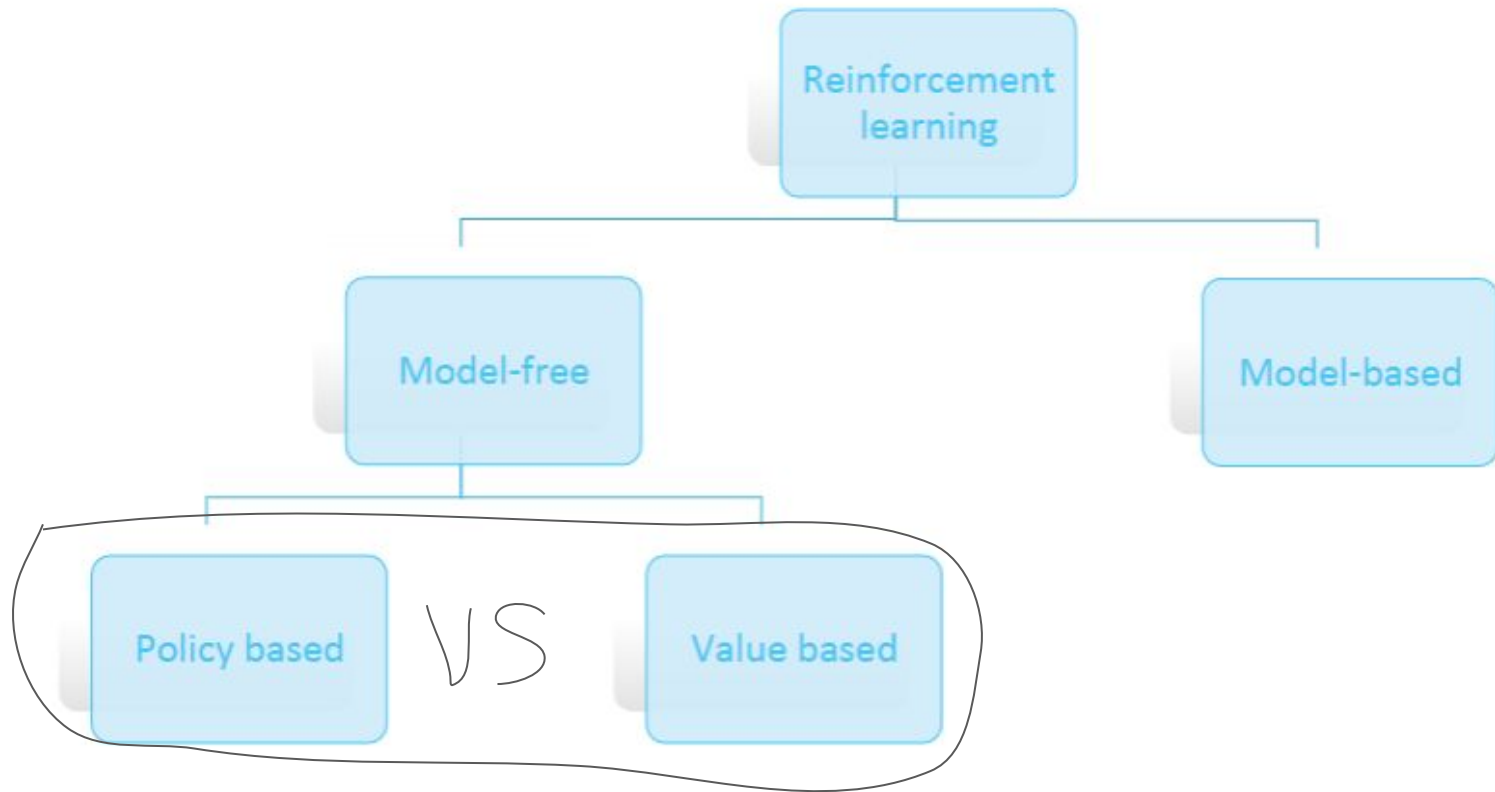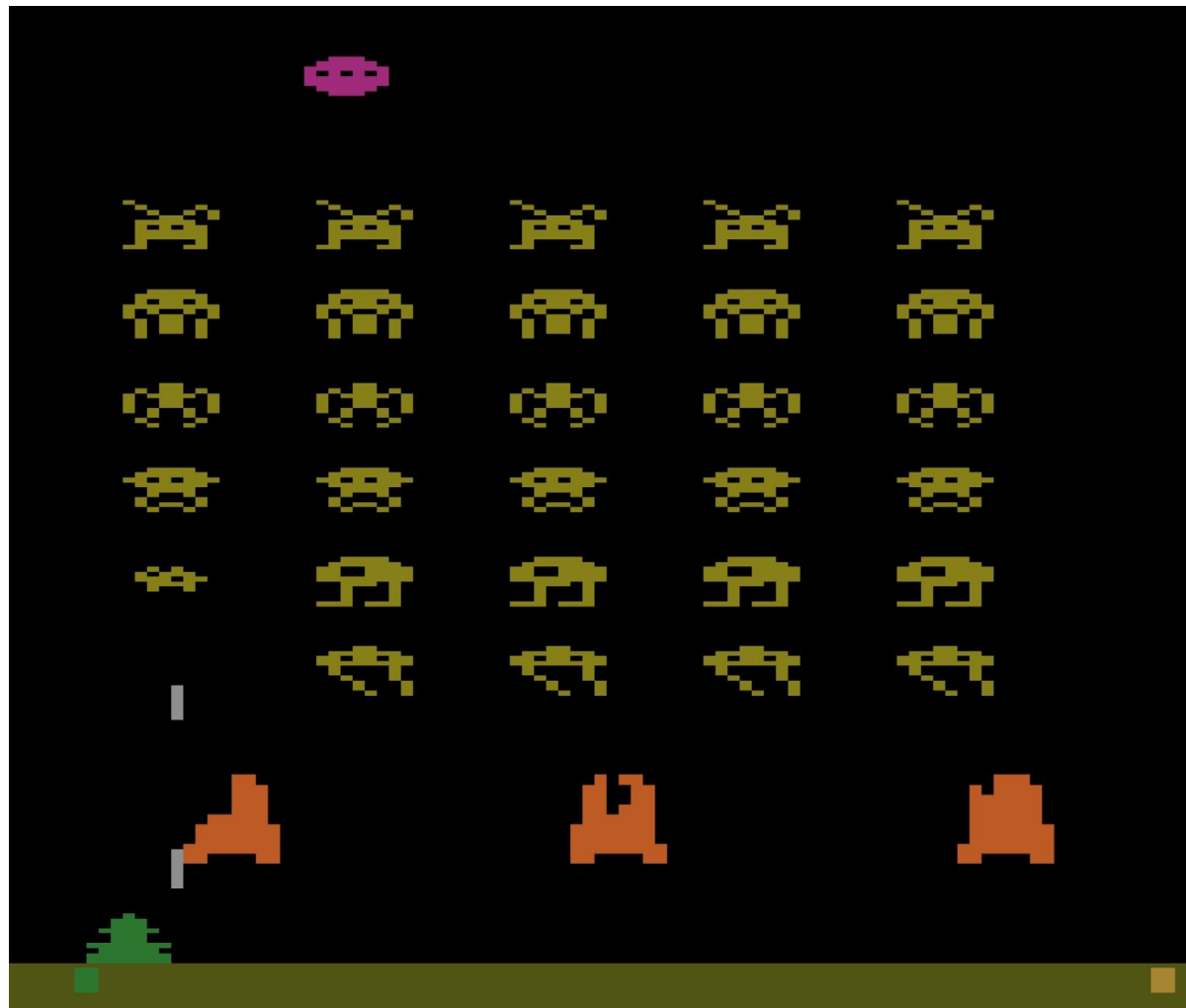
Policy vs Value based (sort off)

Environment

Custom score
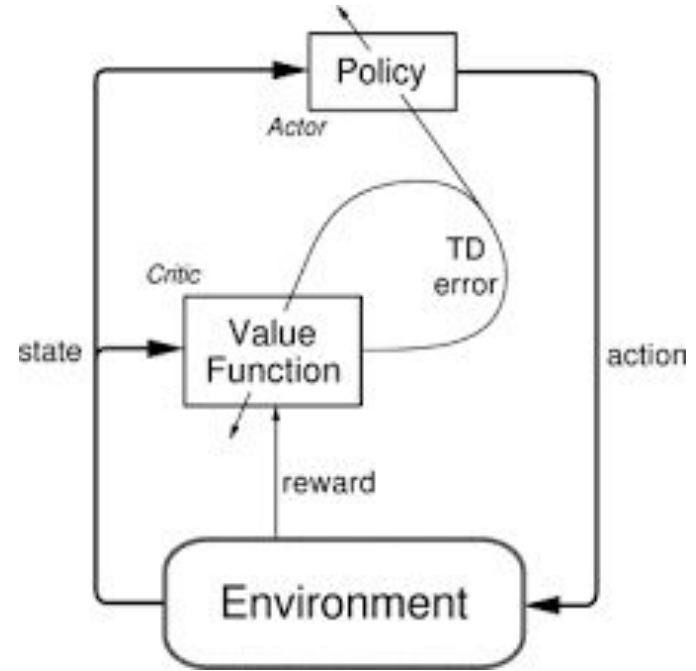
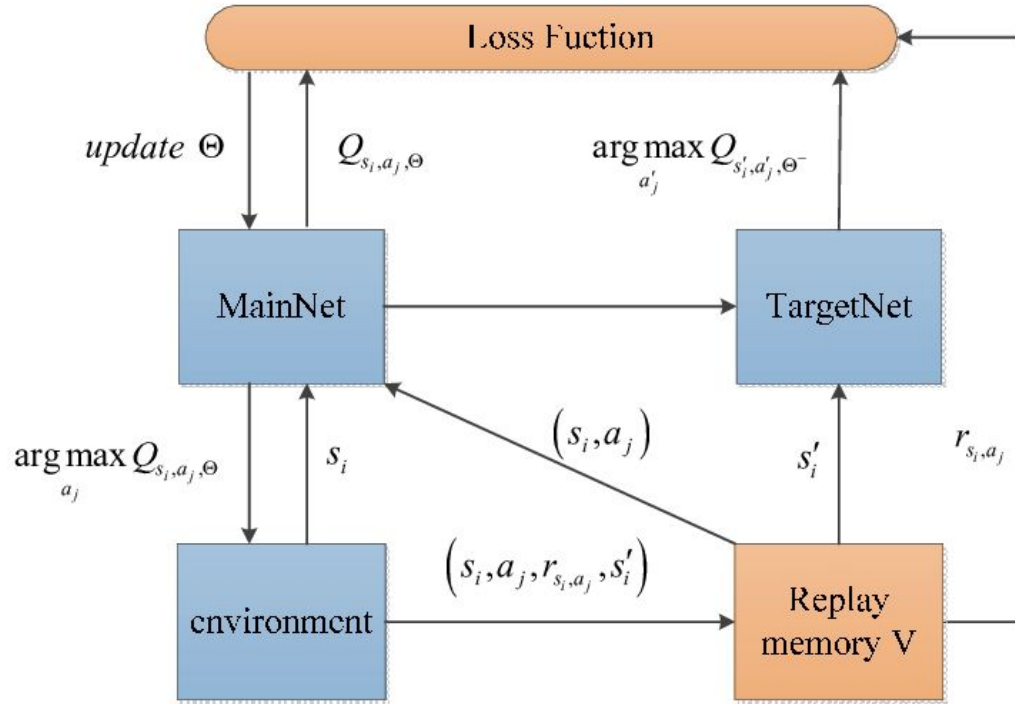100

Wrappers

```python
env = gym.make("SpaceInvaders-v4")
env = GrayScaleObservation(env,keep_dim=True)
env = ResizeObservation(env, (84,84))
env = FrameStack(env, 4)
env = Monitor(env, "./recordings/", force=True)
```
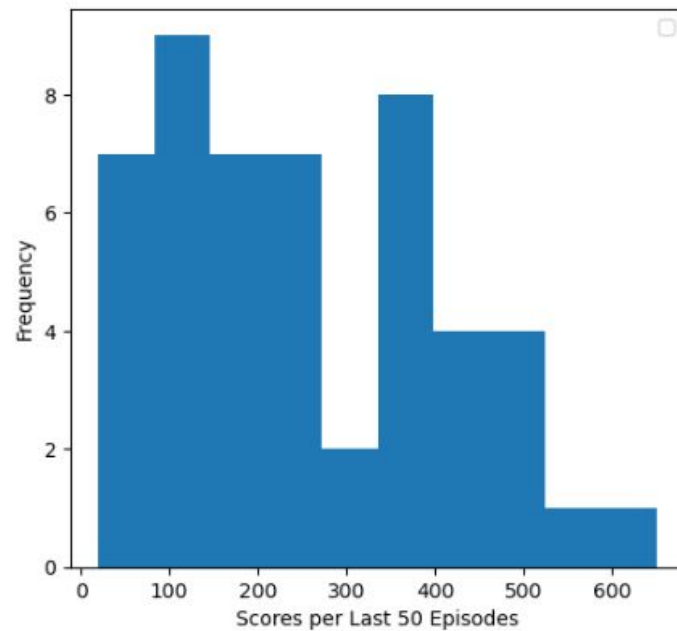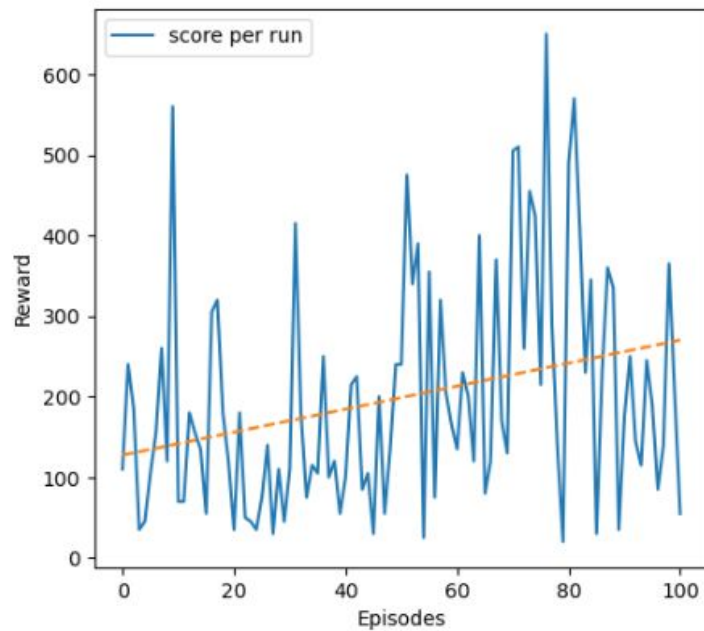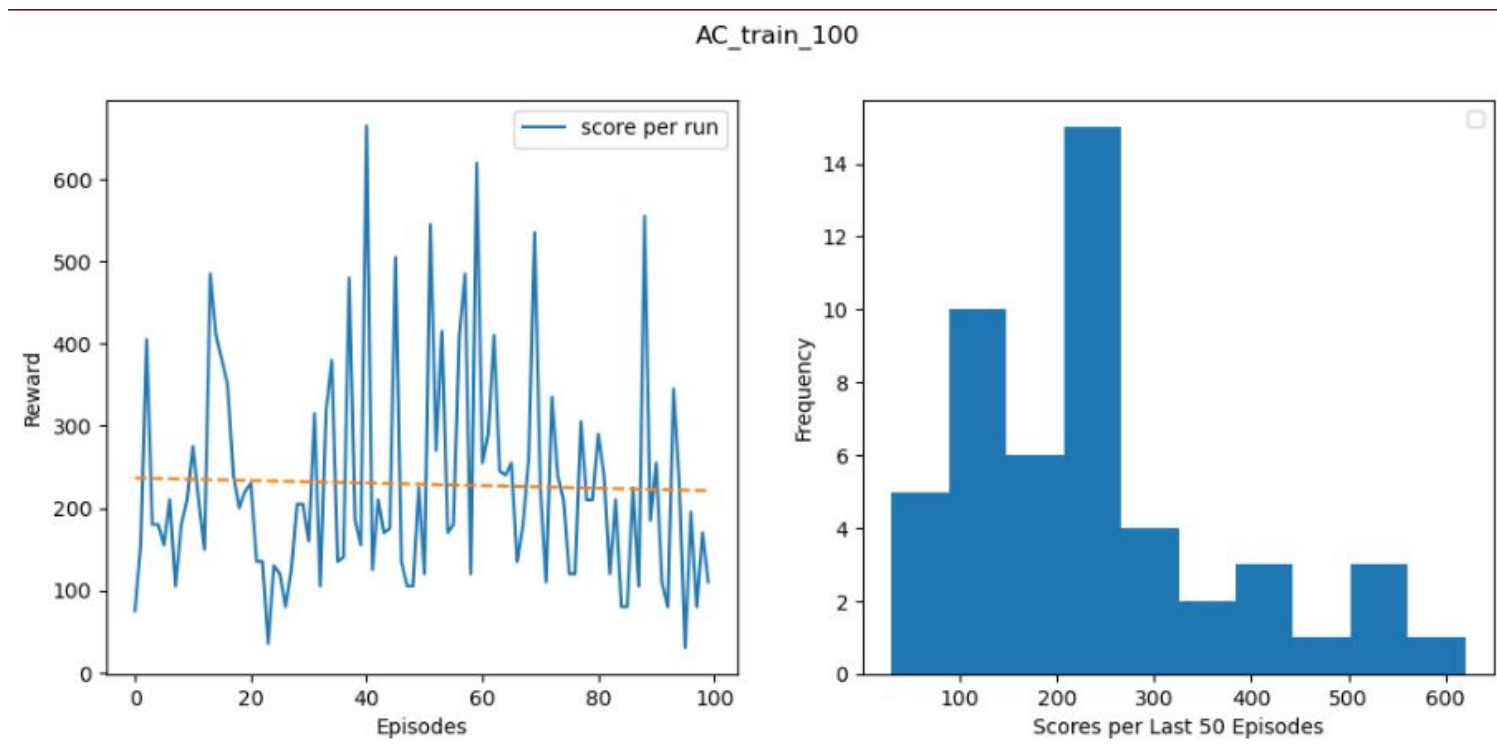
Algorithmes

DQN - AC

Layer (type)	Output Shape	Param #	Connected to
input_1 (InputLayer)	[(None, 84, 84, 1)]	0	[]
conv2d (Conv2D)	(None, 20, 20, 32)	2080	['input_1[0][0]']
conv2d_1 (Conv2D)	(None, 9, 9, 64)	32832	['conv2d[0][0]']
conv2d_2 (Conv2D)	(None, 7, 7, 128)	73856	['conv2d_1[0][0]']
flatten (Flatten)	(None, 6272)	0	['conv2d_2[0][0]']
dense (Dense)	(None, 512)	3211776	['flatten[0][0]']
dense_1 (Dense)	(None, 256)	131328	['dense[0][0]']
dense_2 (Dense)	(None, 6)	1542	['dense_1[0][0]']
dense_3 (Dense)	(None, 1)	257	['dense_1[0][0]']

Total params: 3,453,671
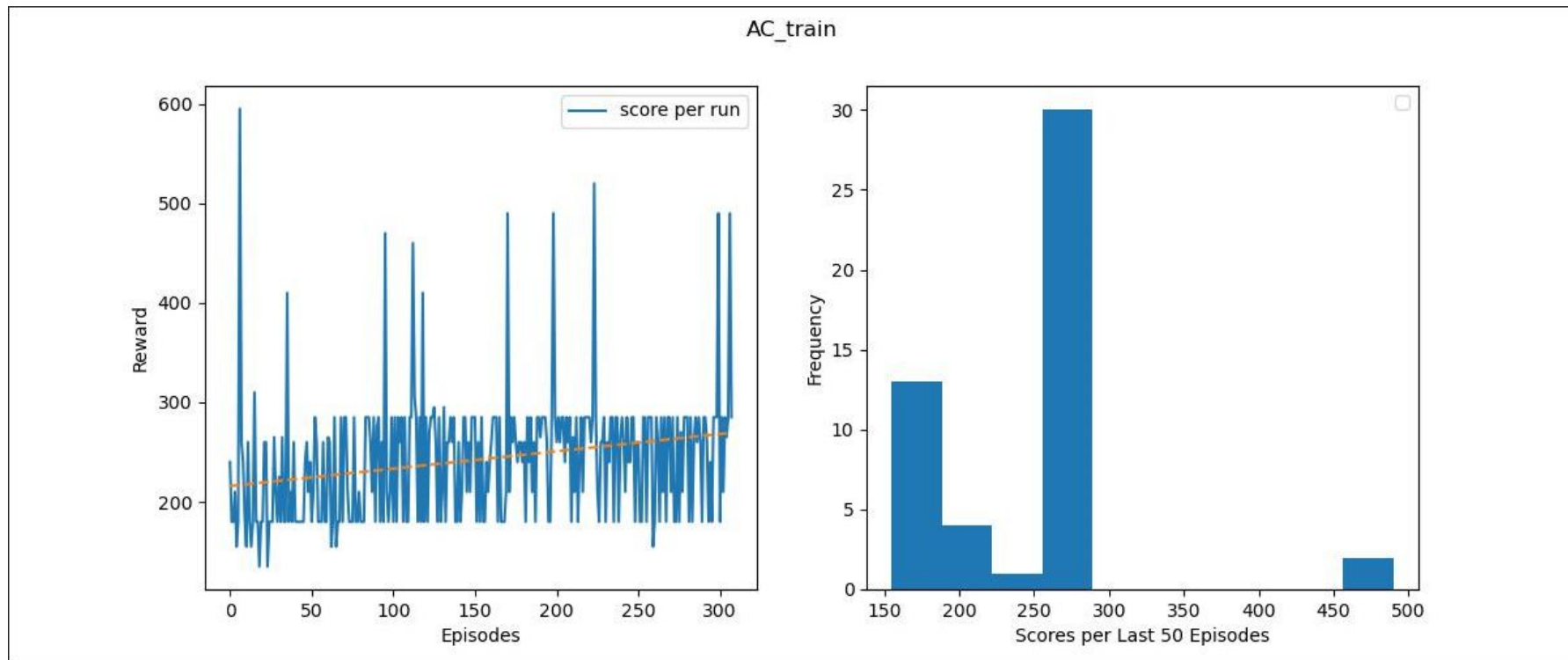Trainable params: 3,453,671
Non-trainable params: 0

DQN

DQN_train2
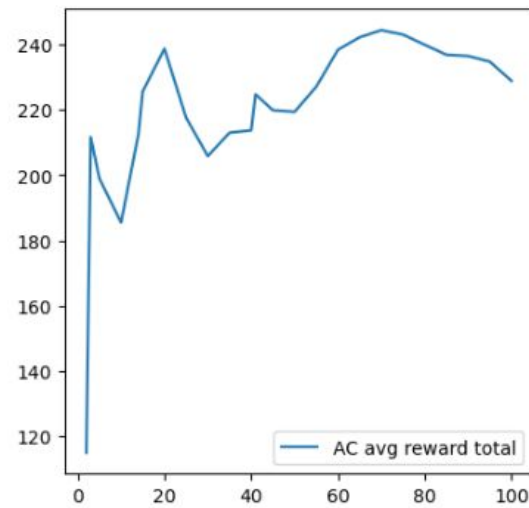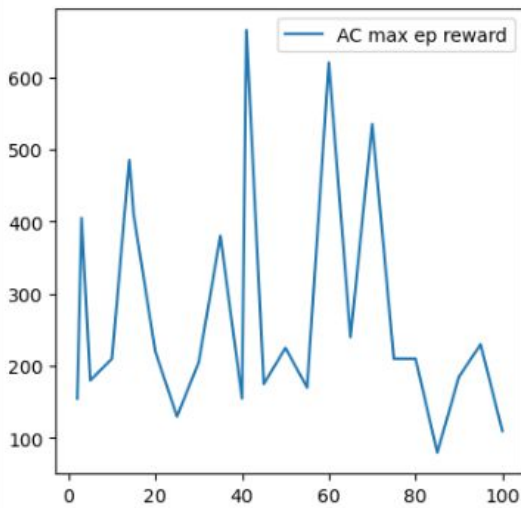
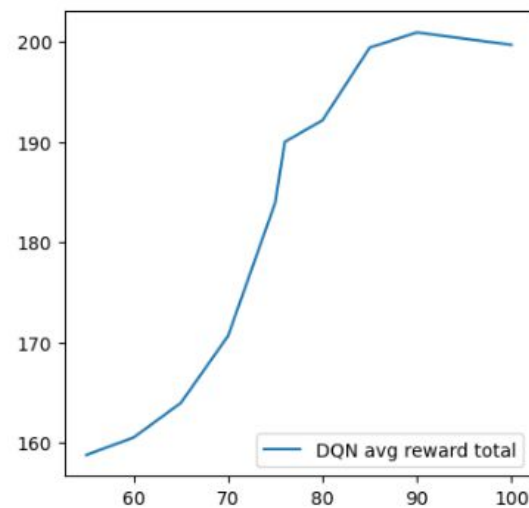# AC
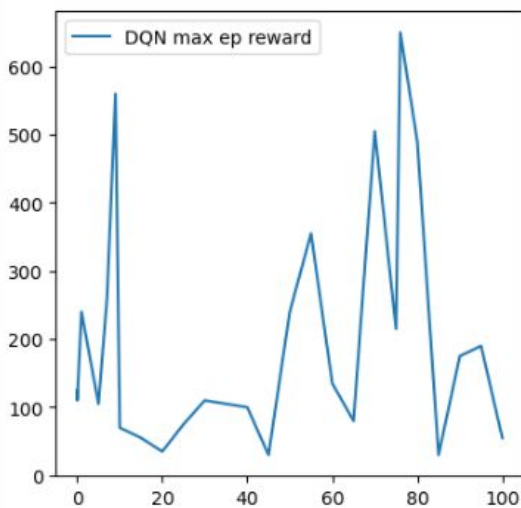

AC_train_100

The curse of 285

Episode rewards

DQN avg: 201
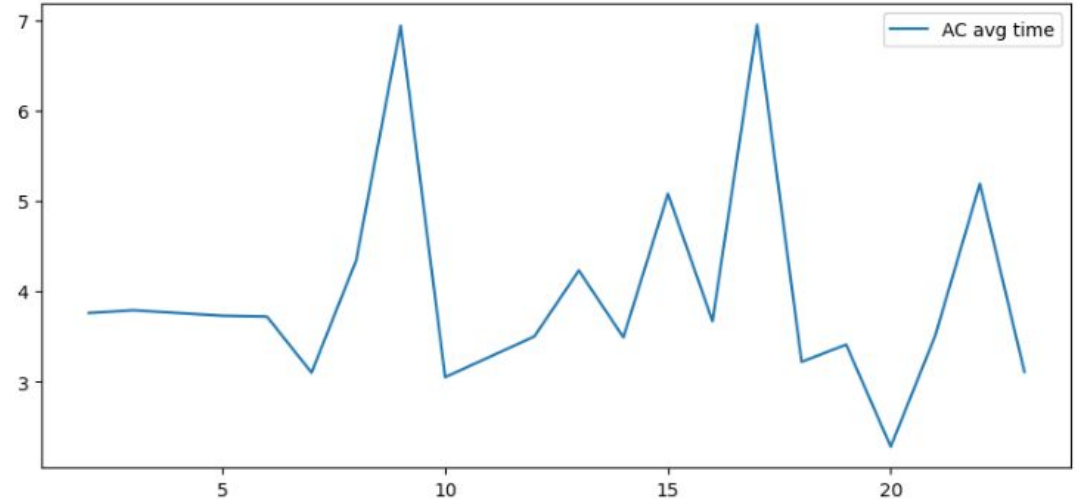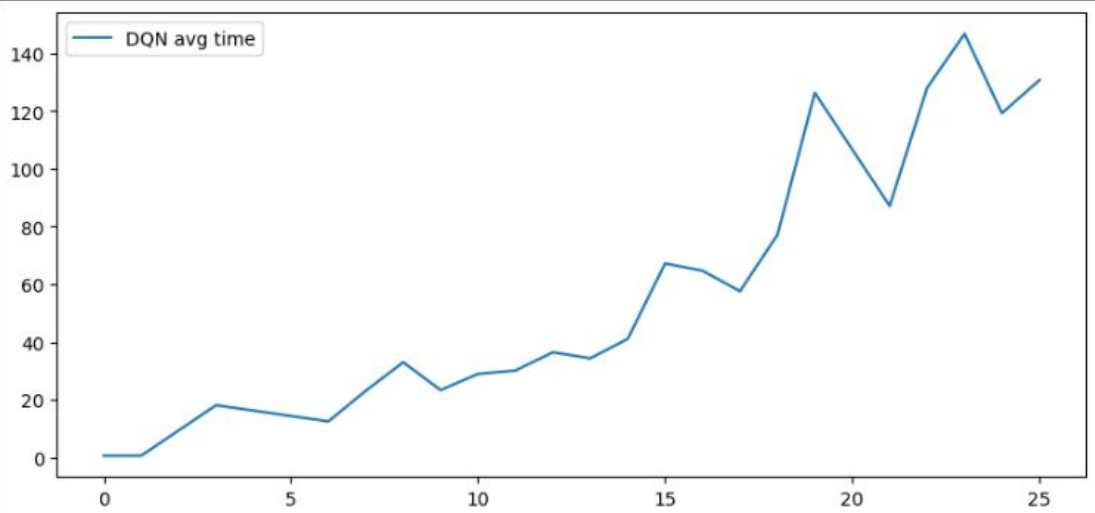AC average: 244

DQN highest: 650
AC highest: 665

Average time

DQN: 0,7 - 58,5 - 146,7

AC : 2,3 - 4 - 7

DQN total ~ 2 u

AC total ~ 15 min

Results

	DQN	AC
High Score	660	665
Consistency	201	244
Time	120 min	15 min

Random actions (50) Best AC result (425) Best dqn result (520)