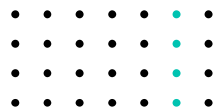# 07

REINFORCE 알고리즘

3. 결과분석

결과분석   로그분석

```python
input_states = Input(shape=(1,self.state_size), name='input_states')
input_action_matrixs = Input(shape=(1,self.action_size), name='input_action_matrixs')
input_action_probs = Input(shape=(1,self.action_size), name='input_action_probs')
input_rewards = Input(shape=(1,self.value_size), name='input_rewards')

x = (input_states)
x = Dense(self.node_num, activation='relu')(x)
out_actions = Dense(self.action_size, activation='softmax', name='output')(x)

model = self.MyModel(inputs=[input_states, input_action_matrixs, input_rewards],
                     outputs=out_actions)

model.compile(optimizer="adam")
```

모델 로그

```
Model: "my_model_7"
_____
Layer (type)                    Output Shape         Param #      Connected to
=========================================================================================
input_states (InputLayer)       [(None, 1, 4)]       0
_____
dense_7 (Dense)                 (None, 1, 12)        60           input_states[0][0]
_____
input_action_matrixs (InputLaye [(None, 1, 2)]       0
_____
input_rewards (InputLayer)      [(None, 1, 1)]       0
_____
output (Dense)                  (None, 1, 2)         26           dense_7[0][0]
=========================================================================================
Total params: 86
Trainable params: 86
Non-trainable params: 0
```

결과분석

로그분석

REINFORCE

```
episode:270, moving_avg:88.75, rewards_avg:68.4870848...
episode:280, moving_avg:99.5, rewards_avg:69.44483985...
episode:290, moving_avg:156.4, rewards_avg:74.46048109965636
episode:300, moving_avg:179.75, rewards_avg:76.70764119601328
episode:310, moving_avg:128.95, rewards_avg:77.90032154340837
episode:320, moving_avg:112.25, rewards_avg:78.85981308411215
episode:330, moving_avg:112.85, rewards_avg:79.95166163141994
episode:340, moving_avg:120.75, rewards_avg:81.25806451612904
episode:350, moving_avg:171.9, rewards_avg:85.13390313390313
episode:360, moving_avg:199.8, rewards_avg:87.77008310249307
episode:370, moving_avg:180.25, rewards_avg:90.20754716981132
episode:380, moving_avg:161.25, rewards_avg:91.5748031496063
episode:390, moving_avg:127.15, rewards_avg:92.0460358056266
episode:400, moving_avg:115.75, rewards_avg:92.73067331670823
episode:410, moving_avg:131.75, rewards_avg:93.9294403892944
episode:420, moving_avg:177.2, rewards_avg:96.6959619952494
episode:430, moving_avg:188.3, rewards_avg:98.26218097447796
episode:440, moving_avg:129.9, rewards_avg:98.15646258503402
episode:450, moving_avg:119.9, rewards_avg:99.17738359201773
episode:460, moving_avg:161.8, rewards_avg:100.87418655097613
episode:470, moving_avg:149.95, rewards_avg:101.29087048832272
episode:480, moving_avg:139.4, rewards_avg:102.43451143451144
episode:490, moving_avg:207.3, rewards_avg:105.56822810590631
INFO:tensorflow:Assets written to: ./model/reinforce\assets
*****end learing
```
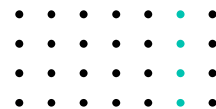
DQN

```
episode:300, moving_avg:99.8, rewards_avg:35.96013289036545
episode:310, moving_avg:92.9, rewards_avg:37.569131832797424
episode:320, moving_avg:89.75, rewards_avg:39.24922118380062
episode:330, moving_avg:92.7, rewards_avg:40.83987915407855
episode:340, moving_avg:92.25, rewards_avg:42.2991202346041
episode:350, moving_avg:101.7, rewards_avg:44.25071225071225
episode:360, moving_avg:103.2, rewards_avg:45.61772853185595
episode:370, moving_avg:101.0, rewards_avg:47.25606469002695
episode:380, moving_avg:100.45, rewards_avg:48.44356955380577
episode:390, moving_avg:101.05, rewards_avg:49.95652173913044
episode:400, moving_avg:110.8, rewards_avg:51.50374064837905
episode:410, moving_avg:113.7, rewards_avg:53.009732360097324
episode:420, moving_avg:119.1, rewards_avg:54.66745843230404
episode:430, moving_avg:128.2, rewards_avg:56.45243619489559
episode:440, moving_avg:124.0, rewards_avg:57.766439909297056
episode:450, moving_avg:104.8, rewards_avg:58.55210643015521
episode:460, moving_avg:92.6, rewards_avg:59.23427331887202
episode:470, moving_avg:80.7, rewards_avg:59.45010615711253
episode:480, moving_avg:70.05, rewards_avg:59.642411642411645
episode:490, moving_avg:59.3, rewards_avg:59.40325865580448
WARNING:tensorflow:From c:\python377\lib\site-
packages\tensorflow\python\ops\resource_variable_ops.py:1817: calling
BaseResourceVariable.__init__ (from tensorflow.python.ops.resource_variable_ops)
with constraint is deprecated and will be removed in a future version.
Instructions for updating:
If using Keras pass *_constraint arguments to layers.
INFO:tensorflow:Assets written to: ./model/dqn\assets
*****end learing
```
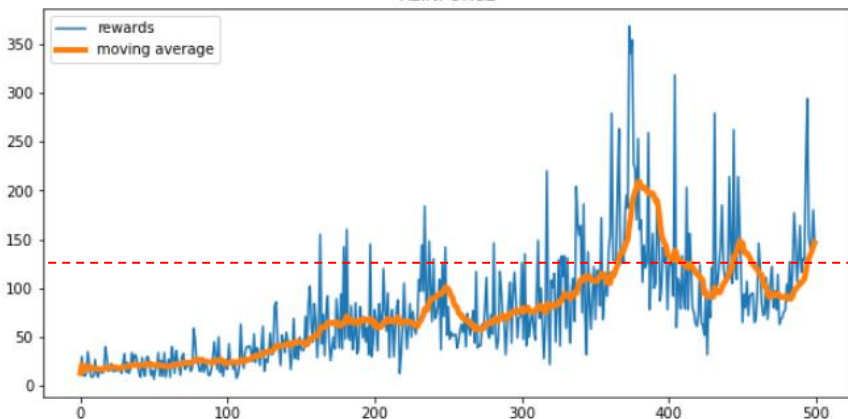
결과분석

시각화