

Exercise no 2.a description

in this exercise, I have use the CartPole-v0 environment , and use the neural network to to see the behavior of the pendulum , there is four intial input to the nural network, and we have use the 1 hidden layer with five neuran, there is two neuran in output layer , when we used this setup with 10 episode and 200 step, we get the reward value , the value of the reward for one episode should be 200 for excellent result, but when we run the program the value of the reward is in range 10-50, so these value is not satisfy, for stability of the CartPole-v0, so we will use the algorithm to improve the stability in next excrise , the result of the reward are given in below figure ,

```
user@45a87cfbedaf: /opt/evorobotpy
reward for th episode 6is 200
reward for th episode 7is 200
reward for th episode 8is 200
reward for th episode 9is 200
user@45a87cfbedaf:/opt/evorobotpy$ python3.5 nu2.py
/usr/local/lib/python3.5/dist-packages/gym/logger.py:30: UserWarning: WARN: Box
bound precision lowered by casting to float32
  warnings.warn(colorize('%s: %s'%( 'WARN', msg % args), 'yellow'))
/usr/local/lib/python3.5/dist-packages/gym/logger.py:30: UserWarning: WARN: You
are calling 'step()' even though this environment has already returned done = Tr
ue. You should always call 'reset()' once you receive 'done = True' -- any furth
er steps are undefined behavior.
  warnings.warn(colorize('%s: %s'%( 'WARN', msg % args), 'yellow'))
reward for th episode 0is 44
reward for th episode 1is 25
reward for th episode 2is 20
reward for th episode 3is 27
reward for th episode 4is 13
reward for th episode 5is 44
reward for th episode 6is 19
reward for th episode 7is 41
reward for th episode 8is 34
reward for th episode 9is 17
user@45a87cfbedaf:/opt/evorobotpy$
```