

17
18

```
1 # Timestamp para el nombre de archivo de pesos
2 now = datetime.now()
3 date_time = now.strftime("%Y%m%d%H%M")
4
5 # Nombre del archivo para guardar los pesos
6 weights_filename = 'dqn_{}_weights.h5f'.format(ENV_NAME)
7 weights_filename = date_time + '_' + weights_filename
8
9 # checkpoint por intervalo
10 saved_model= date_time + '_dqn_PartialRunPong_weight_{step}.h5'
11 checkpoint = ModelIntervalCheckpoint(saved_model, interval=50000)
12
13 # logs para TensorBoard
14 tb_log_dir = 'logs/tmp'
15 tb_callback = TensorBoard(log_dir=tb_log_dir, histogram_freq=0, write_graph=True)
16
17 nb_steps = 1750000
18 nb_steps = train_steps
19 # Entrenamiento con diferentes callbacks
20 #dqn.fit(env, nb_steps=nb_steps, visualize=False, callbacks=[checkpoint, tb_callback], verbose=2)
21 #dqn.fit(env, nb_steps=nb_steps, visualize=False, callbacks=[tb_callback], verbose=2)
22 #dqn.fit(env, nb_steps=nb_steps, visualize=False, callbacks=[checkpoint], verbose=2)
23 train_history = dqn.fit(env, nb_steps=nb_steps, visualize=False, callbacks=[checkpoint], verbose=2)
24
25
26 dqn.save_weights(weights_filename, overwrite=True)
27 print("Pesos guardados en:", weights_filename)
28
29 #nb_episodes = 3
30 #dqn.test(env, nb_episodes=nb_episodes, visualize=False)
31
32 grafica_entrenamiento(train_history)
33
34 calcula_media_test(dqn)
35
36 # colab 397.328 segundos
37 # local 754.133 segundos
```



Training for 250000 steps ...

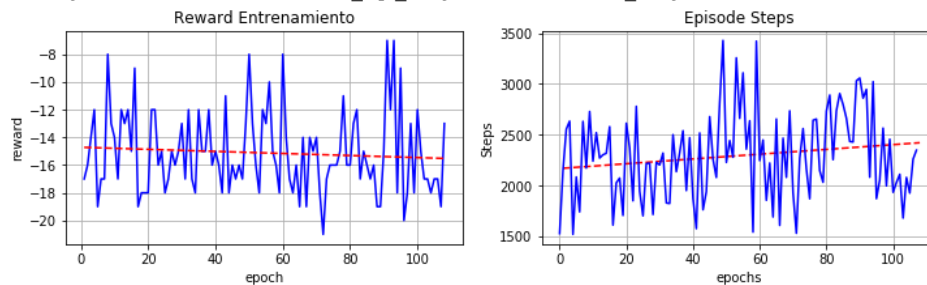
1527/250000: episode: 1, duration: 6.890s, episode steps: 1527, steps per second: 222, episode reward: -17.000, mean re
3670/250000: episode: 2, duration: 11.277s, episode steps: 2143, steps per second: 190, episode reward: -16.000, mean r
6224/250000: episode: 3, duration: 13.436s, episode steps: 2554, steps per second: 190, episode reward: -14.000, mean r
8857/250000: episode: 4, duration: 14.003s, episode steps: 2633, steps per second: 188, episode reward: -12.000, mean r
10376/250000: episode: 5, duration: 8.212s, episode steps: 1519, steps per second: 185, episode reward: -19.000, mean re
12460/250000: episode: 6, duration: 10.978s, episode steps: 2084, steps per second: 190, episode reward: -17.000, mean r
14200/250000: episode: 7, duration: 9.170s, episode steps: 1740, steps per second: 190, episode reward: -17.000, mean re
16831/250000: episode: 8, duration: 13.897s, episode steps: 2631, steps per second: 189, episode reward: -8.000, mean re
19002/250000: episode: 9, duration: 11.503s, episode steps: 2171, steps per second: 189, episode reward: -13.000, mean r
21730/250000: episode: 10, duration: 14.477s, episode steps: 2728, steps per second: 188, episode reward: -14.000, mean
23969/250000: episode: 11, duration: 11.816s, episode steps: 2239, steps per second: 189, episode reward: -17.000, mean
26489/250000: episode: 12, duration: 13.152s, episode steps: 2520, steps per second: 192, episode reward: -12.000, mean
28759/250000: episode: 13, duration: 11.853s, episode steps: 2270, steps per second: 192, episode reward: -13.000, mean
31061/250000: episode: 14, duration: 12.016s, episode steps: 2302, steps per second: 192, episode reward: -12.000, mean
33378/250000: episode: 15, duration: 12.165s, episode steps: 2317, steps per second: 190, episode reward: -15.000, mean
35957/250000: episode: 16, duration: 13.381s, episode steps: 2579, steps per second: 193, episode reward: -9.000, mean r
37567/250000: episode: 17, duration: 8.422s, episode steps: 1610, steps per second: 191, episode reward: -19.000, mean r
39593/250000: episode: 18, duration: 10.704s, episode steps: 2026, steps per second: 189, episode reward: -18.000, mean
41666/250000: episode: 19, duration: 10.824s, episode steps: 2073, steps per second: 192, episode reward: -18.000, mean
43371/250000: episode: 20, duration: 8.912s, episode steps: 1705, steps per second: 191, episode reward: -18.000, mean r
45984/250000: episode: 21, duration: 13.670s, episode steps: 2613, steps per second: 191, episode reward: -12.000, mean
48354/250000: episode: 22, duration: 12.350s, episode steps: 2370, steps per second: 192, episode reward: -12.000, mean
50203/250000: episode: 23, duration: 9.672s, episode steps: 1849, steps per second: 191, episode reward: -16.000, mean r
52982/250000: episode: 24, duration: 14.993s, episode steps: 2779, steps per second: 185, episode reward: -15.000, mean
54926/250000: episode: 25, duration: 10.318s, episode steps: 1944, steps per second: 188, episode reward: -18.000, mean
56627/250000: episode: 26, duration: 8.909s, episode steps: 1701, steps per second: 191, episode reward: -17.000, mean r
58843/250000: episode: 27, duration: 11.623s, episode steps: 2216, steps per second: 191, episode reward: -15.000, mean
61062/250000: episode: 28, duration: 11.643s, episode steps: 2219, steps per second: 191, episode reward: -16.000, mean
62776/250000: episode: 29, duration: 8.999s, episode steps: 1714, steps per second: 190, episode reward: -15.000, mean r
65004/250000: episode: 30, duration: 11.699s, episode steps: 2228, steps per second: 190, episode reward: -13.000, mean
67206/250000: episode: 31, duration: 11.591s, episode steps: 2202, steps per second: 190, episode reward: -17.000, mean
69525/250000: episode: 32, duration: 12.253s, episode steps: 2319, steps per second: 189, episode reward: -12.000, mean
71353/250000: episode: 33, duration: 9.672s, episode steps: 1828, steps per second: 189, episode reward: -17.000, mean r
73176/250000: episode: 34, duration: 9.580s, episode steps: 1823, steps per second: 190, episode reward: -18.000, mean r
75676/250000: episode: 35, duration: 13.142s, episode steps: 2500, steps per second: 190, episode reward: -12.000, mean
77812/250000: episode: 36, duration: 11.379s, episode steps: 2136, steps per second: 188, episode reward: -15.000, mean
80134/250000: episode: 37, duration: 12.206s, episode steps: 2322, steps per second: 190, episode reward: -15.000, mean
82673/250000: episode: 38, duration: 13.355s, episode steps: 2539, steps per second: 190, episode reward: -12.000, mean
84623/250000: episode: 39, duration: 10.289s, episode steps: 1950, steps per second: 190, episode reward: -16.000, mean
87093/250000: episode: 40, duration: 12.963s, episode steps: 2470, steps per second: 191, episode reward: -15.000, mean
88966/250000: episode: 41, duration: 9.843s, episode steps: 1873, steps per second: 190, episode reward: -16.000, mean r
90541/250000: episode: 42, duration: 8.341s, episode steps: 1575, steps per second: 189, episode reward: -18.000, mean r
93058/250000: episode: 43, duration: 13.279s, episode steps: 2517, steps per second: 190, episode reward: -11.000, mean
94819/250000: episode: 44, duration: 9.253s, episode steps: 1761, steps per second: 190, episode reward: -18.000, mean r
96758/250000: episode: 45, duration: 10.233s, episode steps: 1939, steps per second: 189, episode reward: -16.000, mean
99436/250000: episode: 46, duration: 14.225s, episode steps: 2678, steps per second: 188, episode reward: -17.000, mean
101694/250000: episode: 47, duration: 12.053s, episode steps: 2258, steps per second: 187, episode reward: -16.000, mean
103773/250000: episode: 48, duration: 10.937s, episode steps: 2079, steps per second: 190, episode reward: -17.000, mean
106588/250000: episode: 49, duration: 14.834s, episode steps: 2815, steps per second: 190, episode reward: -13.000, mean
110016/250000: episode: 50, duration: 18.018s, episode steps: 3428, steps per second: 190, episode reward: -8.000, mean r
112243/250000: episode: 51, duration: 11.938s, episode steps: 2227, steps per second: 187, episode reward: -13.000, mean
114685/250000: episode: 52, duration: 12.793s, episode steps: 2442, steps per second: 191, episode reward: -16.000, mean
116963/250000: episode: 53, duration: 12.018s, episode steps: 2278, steps per second: 190, episode reward: -18.000, mean
120220/250000: episode: 54, duration: 17.060s, episode steps: 3257, steps per second: 191, episode reward: -12.000, mean
122880/250000: episode: 55, duration: 13.951s, episode steps: 2660, steps per second: 191, episode reward: -13.000, mean
125989/250000: episode: 56, duration: 16.241s, episode steps: 3109, steps per second: 191, episode reward: -10.000, mean
128347/250000: episode: 57, duration: 12.374s, episode steps: 2358, steps per second: 191, episode reward: -15.000, mean
130984/250000: episode: 58, duration: 13.926s, episode steps: 2637, steps per second: 189, episode reward: -16.000, mean
132525/250000: episode: 59, duration: 8.097s, episode steps: 1541, steps per second: 190, episode reward: -18.000, mean r
135946/250000: episode: 60, duration: 18.004s, episode steps: 3421, steps per second: 190, episode reward: -8.000, mean r
138191/250000: episode: 61, duration: 11.680s, episode steps: 2245, steps per second: 192, episode reward: -14.000, mean
140639/250000: episode: 62, duration: 12.840s, episode steps: 2448, steps per second: 191, episode reward: -17.000, mean
142493/250000: episode: 63, duration: 9.706s, episode steps: 1854, steps per second: 191, episode reward: -18.000, mean r
144714/250000: episode: 64, duration: 11.686s, episode steps: 2221, steps per second: 190, episode reward: -16.000, mean
146404/250000: episode: 65, duration: 9.006s, episode steps: 1690, steps per second: 188, episode reward: -19.000, mean r
149059/250000: episode: 66, duration: 13.911s, episode steps: 2655, steps per second: 191, episode reward: -14.000, mean
150666/250000: episode: 67, duration: 8.494s, episode steps: 1607, steps per second: 189, episode reward: -19.000, mean r
153132/250000: episode: 68, duration: 12.921s, episode steps: 2466, steps per second: 191, episode reward: -14.000, mean
155212/250000: episode: 69, duration: 10.899s, episode steps: 2080, steps per second: 191, episode reward: -15.000, mean
157947/250000: episode: 70, duration: 14.294s, episode steps: 2735, steps per second: 191, episode reward: -14.000, mean
159875/250000: episode: 71, duration: 10.132s, episode steps: 1928, steps per second: 190, episode reward: -18.000, mean
161404/250000: episode: 72, duration: 8.109s, episode steps: 1529, steps per second: 189, episode reward: -21.000, mean r
163672/250000: episode: 73, duration: 11.915s, episode steps: 2268, steps per second: 190, episode reward: -17.000, mean
166232/250000: episode: 74, duration: 13.482s, episode steps: 2560, steps per second: 190, episode reward: -16.000, mean
168396/250000: episode: 75, duration: 11.349s, episode steps: 2164, steps per second: 191, episode reward: -16.000, mean
170265/250000: episode: 76, duration: 10.224s, episode steps: 1869, steps per second: 183, episode reward: -16.000, mean
172907/250000: episode: 77, duration: 13.764s, episode steps: 2642, steps per second: 192, episode reward: -15.000, mean
175561/250000: episode: 78, duration: 13.990s, episode steps: 2654, steps per second: 190, episode reward: -11.000, mean
177707/250000: episode: 79, duration: 11.255s, episode steps: 2146, steps per second: 191, episode reward: -16.000, mean
179738/250000: episode: 80, duration: 10.644s, episode steps: 2031, steps per second: 191, episode reward: -16.000, mean
182471/250000: episode: 81, duration: 14.264s, episode steps: 2733, steps per second: 192, episode reward: -13.000, mean
185361/250000: episode: 82, duration: 15.104s, episode steps: 2890, steps per second: 191, episode reward: -12.000, mean
187615/250000: episode: 83, duration: 11.906s, episode steps: 2254, steps per second: 189, episode reward: -17.000, mean
190351/250000: episode: 84, duration: 14.478s, episode steps: 2736, steps per second: 189, episode reward: -15.000, mean
193256/250000: episode: 85, duration: 15.498s, episode steps: 2905, steps per second: 187, episode reward: -16.000, mean
196053/250000: episode: 86, duration: 14.617s, episode steps: 2797, steps per second: 191, episode reward: -17.000, mean
198715/250000: episode: 87, duration: 13.872s, episode steps: 2662, steps per second: 192, episode reward: -16.000, mean
201148/250000: episode: 88, duration: 12.720s, episode steps: 2433, steps per second: 191, episode reward: -19.000, mean
203576/250000: episode: 89, duration: 12.747s, episode steps: 2428, steps per second: 190, episode reward: -19.000, mean
206607/250000: episode: 90, duration: 15.861s, episode steps: 3031, steps per second: 191, episode reward: -15.000, mean
209665/250000: episode: 91, duration: 16.049s, episode steps: 3058, steps per second: 191, episode reward: -7.000, mean r
212525/250000: episode: 92, duration: 14.938s, episode steps: 2860, steps per second: 191, episode reward: -12.000, mean
215471/250000: episode: 93, duration: 15.371s, episode steps: 2946, steps per second: 192, episode reward: -7.000, mean r

```

217552/250000: episode: 94, duration: 10.822s, episode steps: 2081, steps per second: 192, episode reward: -18.000, mean
220574/250000: episode: 95, duration: 15.757s, episode steps: 3022, steps per second: 192, episode reward: -9.000, mean r
222444/250000: episode: 96, duration: 9.912s, episode steps: 1870, steps per second: 189, episode reward: -20.000, mean r
224497/250000: episode: 97, duration: 10.744s, episode steps: 2053, steps per second: 191, episode reward: -18.000, mean
227062/250000: episode: 98, duration: 13.419s, episode steps: 2565, steps per second: 191, episode reward: -13.000, mean
229058/250000: episode: 99, duration: 10.671s, episode steps: 1996, steps per second: 187, episode reward: -18.000, mean
231513/250000: episode: 100, duration: 12.910s, episode steps: 2455, steps per second: 190, episode reward: -12.000, mean
233446/250000: episode: 101, duration: 10.089s, episode steps: 1933, steps per second: 192, episode reward: -15.000, mean
235475/250000: episode: 102, duration: 10.593s, episode steps: 2029, steps per second: 192, episode reward: -17.000, mean
237585/250000: episode: 103, duration: 11.201s, episode steps: 2110, steps per second: 188, episode reward: -17.000, mean
239264/250000: episode: 104, duration: 8.817s, episode steps: 1679, steps per second: 190, episode reward: -18.000, mean
241343/250000: episode: 105, duration: 10.875s, episode steps: 2079, steps per second: 191, episode reward: -17.000, mean
243268/250000: episode: 106, duration: 10.119s, episode steps: 1925, steps per second: 190, episode reward: -17.000, mean
245528/250000: episode: 107, duration: 11.850s, episode steps: 2260, steps per second: 191, episode reward: -19.000, mean
247875/250000: episode: 108, duration: 12.318s, episode steps: 2347, steps per second: 191, episode reward: -13.000, mean
done, took 1319.118 seconds

```

Pesos guardados en: 202001072104_dqn_PongDeterministic-v0_weights.h5f



```

Testing for 10 episodes ...
Episode 1: reward: -11.000, steps: 2928
Episode 2: reward: -12.000, steps: 2900
Episode 3: reward: -17.000, steps: 2044
Episode 4: reward: -15.000, steps: 2338
Episode 5: reward: -17.000, steps: 2239
Episode 6: reward: -18.000, steps: 2440
Episode 7: reward: -20.000, steps: 1866
Episode 8: reward: -15.000, steps: 2270
Episode 9: reward: -14.000, steps: 2282
Episode 10: reward: -16.000, steps: 2366
mean_reward: -15.5
mean_steps: 2367.3
(-15.5, 2367.3)

```

```

1 #my_hist_reward2=[]
2 #my_hist_steps2=[]
3 type(my_hist_reward2)
4 my_hist_reward=(train_history.history['episode_reward'])
5 my_hist_steps=(train_history.history['nb_episode_steps'])
6 my_hist_reward2 = my_hist_reward2 + my_hist_reward
7 my_hist_steps2 = my_hist_steps2 + my_hist_steps
8 type(my_hist_reward2)
9 print(len(my_hist_reward2))

```

551

```

1 #type(my_hist_reward)
2 #sum_history=train_history
3 #import matplotlib.pyplot as plt
4 #print(train_history)
5 #dir(train_history)
6 #mean_reward = mean(pruebas.history.get('episode_reward'))
7 #mean_steps = mean(pruebas.history.get('nb_steps'))
8 #train_history.history.get('episode_reward')
9
10
11 def grafica_entrenamiento2(train_history):
12
13
14
15     # Plot the average line
16
17
18     #epochs = np.arange(1,len(train_history.history['episode_reward'])+1)
19     epochs = np.arange(1,len(my_hist_reward2)+1)
20
21
22     fig = plt.figure(figsize=(12,3))
23     (ax, ax2) = fig.subplots(nrows=1, ncols=2)
24     ax.set_title('Rewards')
25     ax.set_xlabel('epoch')
26     ax.set_ylabel('reward')
27     #ax.plot(epochs,train_history.history['episode_reward'],'b',label='episode_reward')
28     ax.plot(epochs,my_hist_reward2,'b',label='episode_reward')

```