

```
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='dqn_PartialRunPong_weight_{step}.h5'
11 checkpoint = ModelIntervalCheckpoint(saved_model, interval=500000)
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, 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
```

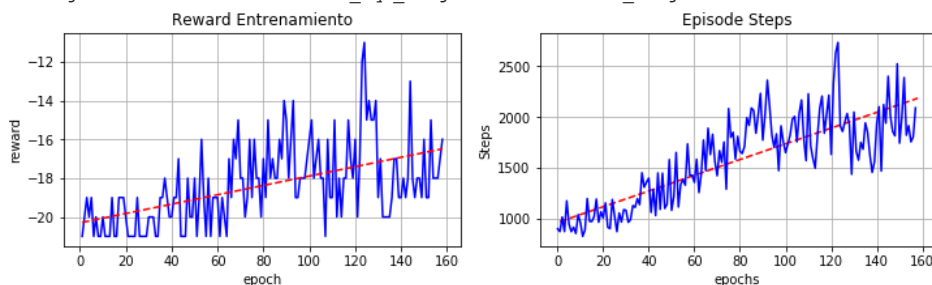


Training for 250000 steps ...

898/250000: episode: 1, duration: 1.978s, episode steps: 898, steps per second: 454, episode reward: -21.000, mean rew
1769/250000: episode: 2, duration: 1.933s, episode steps: 871, steps per second: 451, episode reward: -20.000, mean rew
2783/250000: episode: 3, duration: 2.229s, episode steps: 1014, steps per second: 455, episode reward: -19.000, mean re
3654/250000: episode: 4, duration: 1.908s, episode steps: 871, steps per second: 456, episode reward: -20.000, mean rew
4828/250000: episode: 5, duration: 2.565s, episode steps: 1174, steps per second: 458, episode reward: -19.000, mean re
5772/250000: episode: 6, duration: 2.064s, episode steps: 944, steps per second: 457, episode reward: -21.000, mean rew
6642/250000: episode: 7, duration: 1.897s, episode steps: 870, steps per second: 459, episode reward: -20.000, mean rew
7556/250000: episode: 8, duration: 2.068s, episode steps: 914, steps per second: 442, episode reward: -21.000, mean rew
8408/250000: episode: 9, duration: 1.868s, episode steps: 852, steps per second: 456, episode reward: -21.000, mean rew
9446/250000: episode: 10, duration: 2.256s, episode steps: 1038, steps per second: 460, episode reward: -20.000, mean r
10405/250000: episode: 11, duration: 2.098s, episode steps: 959, steps per second: 457, episode reward: -21.000, mean re
11229/250000: episode: 12, duration: 1.781s, episode steps: 824, steps per second: 463, episode reward: -21.000, mean re
12113/250000: episode: 13, duration: 1.938s, episode steps: 884, steps per second: 456, episode reward: -21.000, mean re
13309/250000: episode: 14, duration: 2.585s, episode steps: 1196, steps per second: 463, episode reward: -19.000, mean r
14282/250000: episode: 15, duration: 2.110s, episode steps: 973, steps per second: 461, episode reward: -21.000, mean re
15252/250000: episode: 16, duration: 2.134s, episode steps: 970, steps per second: 455, episode reward: -21.000, mean re
16270/250000: episode: 17, duration: 2.280s, episode steps: 1018, steps per second: 447, episode reward: -19.000, mean r
17461/250000: episode: 18, duration: 2.550s, episode steps: 1191, steps per second: 467, episode reward: -19.000, mean r
18424/250000: episode: 19, duration: 2.050s, episode steps: 963, steps per second: 470, episode reward: -19.000, mean re
19490/250000: episode: 20, duration: 2.279s, episode steps: 1066, steps per second: 468, episode reward: -20.000, mean r
20495/250000: episode: 21, duration: 2.172s, episode steps: 1005, steps per second: 463, episode reward: -21.000, mean r
21648/250000: episode: 22, duration: 2.555s, episode steps: 1153, steps per second: 451, episode reward: -21.000, mean r
22562/250000: episode: 23, duration: 1.986s, episode steps: 914, steps per second: 460, episode reward: -21.000, mean re
23462/250000: episode: 24, duration: 2.007s, episode steps: 900, steps per second: 448, episode reward: -21.000, mean re
24649/250000: episode: 25, duration: 2.682s, episode steps: 1187, steps per second: 443, episode reward: -19.000, mean r
25672/250000: episode: 26, duration: 2.286s, episode steps: 1023, steps per second: 448, episode reward: -21.000, mean r
26542/250000: episode: 27, duration: 1.841s, episode steps: 870, steps per second: 472, episode reward: -21.000, mean re
27596/250000: episode: 28, duration: 2.287s, episode steps: 1054, steps per second: 461, episode reward: -21.000, mean r
28561/250000: episode: 29, duration: 2.107s, episode steps: 965, steps per second: 458, episode reward: -21.000, mean re
29648/250000: episode: 30, duration: 2.317s, episode steps: 1087, steps per second: 469, episode reward: -20.000, mean r
30728/250000: episode: 31, duration: 2.379s, episode steps: 1080, steps per second: 454, episode reward: -20.000, mean r
31690/250000: episode: 32, duration: 2.186s, episode steps: 962, steps per second: 440, episode reward: -20.000, mean re
32679/250000: episode: 33, duration: 2.147s, episode steps: 989, steps per second: 461, episode reward: -21.000, mean re
33806/250000: episode: 34, duration: 2.419s, episode steps: 1127, steps per second: 466, episode reward: -21.000, mean r
34916/250000: episode: 35, duration: 2.406s, episode steps: 1110, steps per second: 461, episode reward: -19.000, mean r
36111/250000: episode: 36, duration: 2.541s, episode steps: 1195, steps per second: 470, episode reward: -19.000, mean r
37247/250000: episode: 37, duration: 2.400s, episode steps: 1136, steps per second: 473, episode reward: -18.000, mean r
38697/250000: episode: 38, duration: 3.044s, episode steps: 1450, steps per second: 476, episode reward: -19.000, mean r
40005/250000: episode: 39, duration: 2.840s, episode steps: 1308, steps per second: 461, episode reward: -20.000, mean r
41364/250000: episode: 40, duration: 2.951s, episode steps: 1359, steps per second: 461, episode reward: -20.000, mean r
42761/250000: episode: 41, duration: 2.993s, episode steps: 1397, steps per second: 467, episode reward: -19.000, mean r
43821/250000: episode: 42, duration: 2.245s, episode steps: 1060, steps per second: 472, episode reward: -19.000, mean r
45106/250000: episode: 43, duration: 2.712s, episode steps: 1285, steps per second: 474, episode reward: -17.000, mean r
46135/250000: episode: 44, duration: 2.118s, episode steps: 1029, steps per second: 486, episode reward: -21.000, mean r
47583/250000: episode: 45, duration: 3.021s, episode steps: 1448, steps per second: 479, episode reward: -21.000, mean r
48672/250000: episode: 46, duration: 2.378s, episode steps: 1089, steps per second: 458, episode reward: -21.000, mean r
50118/250000: episode: 47, duration: 3.490s, episode steps: 1446, steps per second: 414, episode reward: -18.000, mean r
51217/250000: episode: 48, duration: 5.720s, episode steps: 1099, steps per second: 192, episode reward: -20.000, mean r
52357/250000: episode: 49, duration: 5.971s, episode steps: 1140, steps per second: 191, episode reward: -20.000, mean r
53937/250000: episode: 50, duration: 8.223s, episode steps: 1580, steps per second: 192, episode reward: -18.000, mean r
55017/250000: episode: 51, duration: 5.557s, episode steps: 1080, steps per second: 194, episode reward: -21.000, mean r
56234/250000: episode: 52, duration: 6.281s, episode steps: 1217, steps per second: 194, episode reward: -19.000, mean r
57884/250000: episode: 53, duration: 8.547s, episode steps: 1650, steps per second: 193, episode reward: -16.000, mean r
58997/250000: episode: 54, duration: 5.796s, episode steps: 1113, steps per second: 192, episode reward: -19.000, mean r
60335/250000: episode: 55, duration: 6.883s, episode steps: 1338, steps per second: 194, episode reward: -21.000, mean r
61717/250000: episode: 56, duration: 7.269s, episode steps: 1382, steps per second: 190, episode reward: -18.000, mean r
63045/250000: episode: 57, duration: 7.071s, episode steps: 1328, steps per second: 188, episode reward: -21.000, mean r
64779/250000: episode: 58, duration: 8.984s, episode steps: 1734, steps per second: 193, episode reward: -19.000, mean r
66202/250000: episode: 59, duration: 7.309s, episode steps: 1423, steps per second: 195, episode reward: -19.000, mean r
67628/250000: episode: 60, duration: 7.309s, episode steps: 1426, steps per second: 195, episode reward: -19.000, mean r
68984/250000: episode: 61, duration: 6.958s, episode steps: 1356, steps per second: 195, episode reward: -21.000, mean r
70567/250000: episode: 62, duration: 8.146s, episode steps: 1583, steps per second: 194, episode reward: -19.000, mean r
71822/250000: episode: 63, duration: 6.439s, episode steps: 1255, steps per second: 195, episode reward: -20.000, mean r
73258/250000: episode: 64, duration: 7.407s, episode steps: 1436, steps per second: 194, episode reward: -21.000, mean r
75030/250000: episode: 65, duration: 9.145s, episode steps: 1772, steps per second: 194, episode reward: -17.000, mean r
76501/250000: episode: 66, duration: 7.558s, episode steps: 1471, steps per second: 195, episode reward: -19.000, mean r
78390/250000: episode: 67, duration: 9.619s, episode steps: 1889, steps per second: 196, episode reward: -16.000, mean r
80019/250000: episode: 68, duration: 8.330s, episode steps: 1629, steps per second: 196, episode reward: -17.000, mean r
81848/250000: episode: 69, duration: 9.381s, episode steps: 1829, steps per second: 195, episode reward: -15.000, mean r
83409/250000: episode: 70, duration: 7.988s, episode steps: 1561, steps per second: 195, episode reward: -18.000, mean r
84830/250000: episode: 71, duration: 7.312s, episode steps: 1421, steps per second: 194, episode reward: -18.000, mean r
86498/250000: episode: 72, duration: 8.585s, episode steps: 1668, steps per second: 194, episode reward: -20.000, mean r
88051/250000: episode: 73, duration: 8.002s, episode steps: 1553, steps per second: 194, episode reward: -19.000, mean r
89804/250000: episode: 74, duration: 9.029s, episode steps: 1753, steps per second: 194, episode reward: -16.000, mean r
91093/250000: episode: 75, duration: 6.657s, episode steps: 1289, steps per second: 194, episode reward: -19.000, mean r
93176/250000: episode: 76, duration: 10.700s, episode steps: 2083, steps per second: 195, episode reward: -16.000, mean
94975/250000: episode: 77, duration: 9.252s, episode steps: 1799, steps per second: 194, episode reward: -18.000, mean r
96824/250000: episode: 78, duration: 9.466s, episode steps: 1849, steps per second: 195, episode reward: -18.000, mean r
98409/250000: episode: 79, duration: 8.232s, episode steps: 1585, steps per second: 193, episode reward: -20.000, mean r
100218/250000: episode: 80, duration: 9.253s, episode steps: 1809, steps per second: 196, episode reward: -17.000, mean r
101879/250000: episode: 81, duration: 8.487s, episode steps: 1661, steps per second: 196, episode reward: -19.000, mean r
103519/250000: episode: 82, duration: 8.393s, episode steps: 1640, steps per second: 195, episode reward: -15.000, mean r
105229/250000: episode: 83, duration: 8.765s, episode steps: 1710, steps per second: 195, episode reward: -18.000, mean r
107221/250000: episode: 84, duration: 10.228s, episode steps: 1992, steps per second: 195, episode reward: -17.000, mean
109137/250000: episode: 85, duration: 9.836s, episode steps: 1916, steps per second: 195, episode reward: -18.000, mean r
111225/250000: episode: 86, duration: 10.676s, episode steps: 2088, steps per second: 196, episode reward: -18.000, mean
113289/250000: episode: 87, duration: 10.520s, episode steps: 2064, steps per second: 196, episode reward: -16.000, mean
115128/250000: episode: 88, duration: 9.412s, episode steps: 1839, steps per second: 195, episode reward: -17.000, mean r
117149/250000: episode: 89, duration: 10.353s, episode steps: 2021, steps per second: 195, episode reward: -14.000, mean
119381/250000: episode: 90, duration: 11.390s, episode steps: 2232, steps per second: 196, episode reward: -15.000, mean
121153/250000: episode: 91, duration: 9.129s, episode steps: 1772, steps per second: 194, episode reward: -18.000, mean r
123175/250000: episode: 92, duration: 10.706s, episode steps: 2022, steps per second: 189, episode reward: -16.000, mean
125537/250000: episode: 93, duration: 12.153s, episode steps: 2362, steps per second: 194, episode reward: -14.000, mean

127639/250000: episode: 94, duration: 10.787s, episode steps: 2102, steps per second: 195, episode reward: -19.000, mean
129460/250000: episode: 95, duration: 9.270s, episode steps: 1821, steps per second: 196, episode reward: -19.000, mean r
131160/250000: episode: 96, duration: 8.678s, episode steps: 1700, steps per second: 196, episode reward: -18.000, mean r
132993/250000: episode: 97, duration: 9.430s, episode steps: 1833, steps per second: 194, episode reward: -18.000, mean r
134464/250000: episode: 98, duration: 7.546s, episode steps: 1471, steps per second: 195, episode reward: -18.000, mean r
136376/250000: episode: 99, duration: 9.823s, episode steps: 1912, steps per second: 195, episode reward: -17.000, mean r
138131/250000: episode: 100, duration: 9.006s, episode steps: 1755, steps per second: 195, episode reward: -16.000, mean
139779/250000: episode: 101, duration: 8.438s, episode steps: 1648, steps per second: 195, episode reward: -15.000, mean
141511/250000: episode: 102, duration: 8.841s, episode steps: 1732, steps per second: 196, episode reward: -18.000, mean
143333/250000: episode: 103, duration: 9.321s, episode steps: 1822, steps per second: 195, episode reward: -17.000, mean
145311/250000: episode: 104, duration: 10.091s, episode steps: 1978, steps per second: 196, episode reward: -16.000, mean
147316/250000: episode: 105, duration: 10.170s, episode steps: 2005, steps per second: 197, episode reward: -18.000, mean
149071/250000: episode: 106, duration: 8.958s, episode steps: 1755, steps per second: 196, episode reward: -18.000, mean
151108/250000: episode: 107, duration: 10.434s, episode steps: 2037, steps per second: 195, episode reward: -21.000, mean
153275/250000: episode: 108, duration: 11.107s, episode steps: 2167, steps per second: 195, episode reward: -16.000, mean
155061/250000: episode: 109, duration: 9.124s, episode steps: 1786, steps per second: 196, episode reward: -19.000, mean
156631/250000: episode: 110, duration: 8.041s, episode steps: 1570, steps per second: 195, episode reward: -19.000, mean
158859/250000: episode: 111, duration: 11.476s, episode steps: 2228, steps per second: 194, episode reward: -15.000, mean
160573/250000: episode: 112, duration: 8.759s, episode steps: 1714, steps per second: 196, episode reward: -20.000, mean
162165/250000: episode: 113, duration: 8.159s, episode steps: 1592, steps per second: 195, episode reward: -18.000, mean
163661/250000: episode: 114, duration: 7.646s, episode steps: 1496, steps per second: 196, episode reward: -20.000, mean
165475/250000: episode: 115, duration: 9.243s, episode steps: 1814, steps per second: 196, episode reward: -18.000, mean
167561/250000: episode: 116, duration: 10.692s, episode steps: 2086, steps per second: 195, episode reward: -18.000, mean
169766/250000: episode: 117, duration: 11.301s, episode steps: 2205, steps per second: 195, episode reward: -15.000, mean
171605/250000: episode: 118, duration: 9.435s, episode steps: 1839, steps per second: 195, episode reward: -17.000, mean
173626/250000: episode: 119, duration: 10.391s, episode steps: 2021, steps per second: 194, episode reward: -18.000, mean
175840/250000: episode: 120, duration: 11.323s, episode steps: 2214, steps per second: 196, episode reward: -16.000, mean
177471/250000: episode: 121, duration: 8.382s, episode steps: 1631, steps per second: 195, episode reward: -20.000, mean
179795/250000: episode: 122, duration: 11.855s, episode steps: 2324, steps per second: 196, episode reward: -17.000, mean
182427/250000: episode: 123, duration: 13.625s, episode steps: 2632, steps per second: 193, episode reward: -12.000, mean
185161/250000: episode: 124, duration: 14.296s, episode steps: 2734, steps per second: 191, episode reward: -11.000, mean
187072/250000: episode: 125, duration: 9.777s, episode steps: 1911, steps per second: 195, episode reward: -15.000, mean
188929/250000: episode: 126, duration: 9.477s, episode steps: 1857, steps per second: 196, episode reward: -14.000, mean
190880/250000: episode: 127, duration: 9.988s, episode steps: 1951, steps per second: 195, episode reward: -15.000, mean
192913/250000: episode: 128, duration: 10.421s, episode steps: 2033, steps per second: 195, episode reward: -15.000, mean
194837/250000: episode: 129, duration: 9.831s, episode steps: 1924, steps per second: 196, episode reward: -14.000, mean
196272/250000: episode: 130, duration: 7.396s, episode steps: 1435, steps per second: 194, episode reward: -19.000, mean
198320/250000: episode: 131, duration: 10.528s, episode steps: 2048, steps per second: 195, episode reward: -17.000, mean
199980/250000: episode: 132, duration: 8.540s, episode steps: 1660, steps per second: 194, episode reward: -20.000, mean
201555/250000: episode: 133, duration: 8.089s, episode steps: 1575, steps per second: 195, episode reward: -20.000, mean
203304/250000: episode: 134, duration: 8.899s, episode steps: 1749, steps per second: 197, episode reward: -20.000, mean
204986/250000: episode: 135, duration: 8.651s, episode steps: 1682, steps per second: 194, episode reward: -20.000, mean
206927/250000: episode: 136, duration: 9.963s, episode steps: 1941, steps per second: 195, episode reward: -19.000, mean
208774/250000: episode: 137, duration: 9.459s, episode steps: 1847, steps per second: 195, episode reward: -17.000, mean
210399/250000: episode: 138, duration: 8.341s, episode steps: 1625, steps per second: 195, episode reward: -17.000, mean
211853/250000: episode: 139, duration: 7.487s, episode steps: 1454, steps per second: 194, episode reward: -19.000, mean
213412/250000: episode: 140, duration: 8.008s, episode steps: 1559, steps per second: 195, episode reward: -19.000, mean
215345/250000: episode: 141, duration: 9.969s, episode steps: 1933, steps per second: 194, episode reward: -18.000, mean
217445/250000: episode: 142, duration: 10.861s, episode steps: 2100, steps per second: 193, episode reward: -19.000, mean
218912/250000: episode: 143, duration: 7.568s, episode steps: 1467, steps per second: 194, episode reward: -18.000, mean
221034/250000: episode: 144, duration: 10.907s, episode steps: 2122, steps per second: 195, episode reward: -13.000, mean
222973/250000: episode: 145, duration: 9.936s, episode steps: 1939, steps per second: 195, episode reward: -18.000, mean
225374/250000: episode: 146, duration: 12.276s, episode steps: 2401, steps per second: 196, episode reward: -19.000, mean
227381/250000: episode: 147, duration: 10.318s, episode steps: 2007, steps per second: 195, episode reward: -18.000, mean
229237/250000: episode: 148, duration: 9.491s, episode steps: 1856, steps per second: 196, episode reward: -18.000, mean
231051/250000: episode: 149, duration: 9.333s, episode steps: 1814, steps per second: 194, episode reward: -19.000, mean
233575/250000: episode: 150, duration: 12.930s, episode steps: 2524, steps per second: 195, episode reward: -16.000, mean
235317/250000: episode: 151, duration: 8.909s, episode steps: 1742, steps per second: 196, episode reward: -19.000, mean
237312/250000: episode: 152, duration: 10.188s, episode steps: 1995, steps per second: 196, episode reward: -19.000, mean
239700/250000: episode: 153, duration: 12.205s, episode steps: 2388, steps per second: 196, episode reward: -15.000, mean
241520/250000: episode: 154, duration: 9.304s, episode steps: 1820, steps per second: 196, episode reward: -18.000, mean
243434/250000: episode: 155, duration: 10.155s, episode steps: 1914, steps per second: 188, episode reward: -18.000, mean
245189/250000: episode: 156, duration: 9.113s, episode steps: 1755, steps per second: 193, episode reward: -18.000, mean
246993/250000: episode: 157, duration: 9.297s, episode steps: 1804, steps per second: 194, episode reward: -17.000, mean
249082/250000: episode: 158, duration: 10.672s, episode steps: 2089, steps per second: 196, episode reward: -16.000, mean
done, took 1141.620 seconds

Pesos guardados en: 202001072014_dqn_PongDeterministic-v0_weights.h5f



Testing for 10 episodes ...

Episode 1: reward: -17.000, steps: 2450
Episode 2: reward: -14.000, steps: 2603
Episode 3: reward: -20.000, steps: 2454
Episode 4: reward: -16.000, steps: 2169
Episode 5: reward: -17.000, steps: 2311
Episode 6: reward: -15.000, steps: 2960
Episode 7: reward: -17.000, steps: 2496
Episode 8: reward: -18.000, steps: 2428
Episode 9: reward: -18.000, steps: 2359
Episode 10: reward: -18.000, steps: 2072
mean reward: -17.0
mean_steps: 2430.2
(-17.0, 2430.2)