```
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 # Entrenamento 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 ...
    870/250000: episode: 1, duration: 1.943s, episode steps: 870, steps per second: 448, episode reward: -20.000, mean rew
   1923/250000: episode: 2, duration: 2.258s, episode steps: 1053, steps per second: 466, episode reward: -19.000, mean re
   2870/250000: episode: 3, duration: 2.050s, episode steps: 947, steps per second: 462, episode reward: -20.000, mean rew
   3752/250000: episode: 4, duration: 1.917s, episode steps: 882, steps per second: 460, episode reward: -21.000, mean rew
   4901/250000: episode: 5, duration: 2.553s, episode steps: 1149, steps per second: 450, episode reward: -18.000, mean re
   5881/250000: episode: 6, duration: 2.148s, episode steps: 980, steps per second: 456, episode reward: -19.000, mean rew
   6967/250000: episode: 7, duration: 2.355s, episode steps: 1086, steps per second: 461, episode reward: -19.000, mean re
   7991/250000: episode: 8, duration: 2.221s, episode steps: 1024, steps per second: 461, episode reward: -21.000, mean re
   8813/250000: episode: 9, duration: 1.810s, episode steps: 822, steps per second: 454, episode reward: -21.000, mean rew
   9910/250000: episode: 10, duration: 2.371s, episode steps: 1097, steps per second: 463, episode reward: -20.000, mean r
  10734/250000: episode: 11, duration: 1.798s, episode steps: 824, steps per second: 458, episode reward: -21.000, mean re
  11660/250000: episode: 12, duration: 1.984s, episode steps: 926, steps per second: 467, episode reward: -20.000, mean re
  12687/250000: episode: 13, duration: 2.259s, episode steps: 1027, steps per second: 455, episode reward: -20.000, mean r
  13509/250000: episode: 14, duration: 1.769s, episode steps: 822, steps per second: 465, episode reward: -21.000, mean re
  14511/250000: episode: 15, duration: 2.099s, episode steps: 1002, steps per second: 477, episode reward: -21.000, mean re
  15441/250000: episode: 16, duration: 1.953s, episode steps: 930, steps per second: 476, episode reward: -20.000, mean re
  16429/250000: episode: 17, duration: 2.077s, episode steps: 988, steps per second: 476, episode reward: -21.000, mean re
  17652/250000: episode: 18, duration: 2.689s, episode steps: 1223, steps per second: 455, episode reward: -19.000, mean r
  18694/250000: episode: 19, duration: 2.261s, episode steps: 1042, steps per second: 461, episode reward: -19.000, mean r
  19714/250000: episode: 20, duration: 2.205s, episode steps: 1020, steps per second: 462, episode reward: -20.000, mean r
  20599/250000: episode: 21, duration: 1.890s, episode steps: 885, steps per second: 468, episode reward: -21.000, mean re
  21604/250000: episode: 22, duration: 2.147s, episode steps: 1005, steps per second: 468, episode reward: -20.000, mean r
  22704/250000: episode: 23, duration: 2.456s, episode steps: 1100, steps per second: 448, episode reward: -20.000, mean r
  23695/250000: episode: 24, duration: 2.291s, episode steps: 991, steps per second: 433, episode reward: -20.000, mean re
  24837/250000: episode: 25, duration: 2.537s, episode steps: 1142, steps per second: 450, episode reward: -20.000, mean r
  26035/250000: episode: 26, duration: 2.684s, episode steps: 1198, steps per second: 446, episode reward: -18.000, mean r
  27003/250000: episode: 27, duration: 2.227s, episode steps: 968, steps per second: 435, episode reward: -21.000, mean re
  28111/250000: episode: 28, duration: 2.346s, episode steps: 1108, steps per second: 472, episode reward: -20.000, mean r
  28977/250000: episode: 29, duration: 1.869s, episode steps: 866, steps per second: 463, episode reward: -21.000, mean re
  29978/250000: episode: 30, duration: 2.108s, episode steps: 1001, steps per second: 475, episode reward: -19.000, mean r
  31120/250000: episode: 31, duration: 2.475s, episode steps: 1142, steps per second: 461, episode reward: -20.000, mean r
  32484/250000: episode: 32, duration: 2.970s, episode steps: 1364, steps per second: 459, episode reward: -16.000, mean r
  33649/250000: episode: 33, duration: 2.540s, episode steps: 1165, steps per second: 459, episode reward: -19.000, mean r
  34815/250000: episode: 34, duration: 2.549s, episode steps: 1166, steps per second: 457, episode reward: -20.000, mean r
  36105/250000: episode: 35, duration: 2.874s, episode steps: 1290, steps per second: 449, episode reward: -20.000, mean r
  37192/250000: episode: 36, duration: 2.347s, episode steps: 1087, steps per second: 463, episode reward: -20.000, mean r
  38281/250000: episode: 37, duration: 2.352s, episode steps: 1089, steps per second: 463, episode reward: -19.000, mean r
  39400/250000: episode: 38, duration: 2.390s, episode steps: 1119, steps per second: 468, episode reward: -18.000, mean r
  41078/250000: episode: 39, duration: 3.596s, episode steps: 1678, steps per second: 467, episode reward: -17.000, mean r
  42767/250000: episode: 40, duration: 3.625s, episode steps: 1689, steps per second: 466, episode reward: -15.000, mean r
  44138/250000: episode: 41, duration: 2.886s, episode steps: 1371, steps per second: 475, episode reward: -20.000, mean r
  45325/250000: episode: 42, duration: 2.505s, episode steps: 1187, steps per second: 474, episode reward: -20.000, mean r
  46550/250000: episode: 43, duration: 2.574s, episode steps: 1225, steps per second: 476, episode reward: -21.000, mean r
  47915/250000: episode: 44, duration: 2.905s, episode steps: 1365, steps per second: 470, episode reward: -20.000, mean r
  48814/250000: episode: 45, duration: 1.929s, episode steps: 899, steps per second: 466, episode reward: -21.000, mean re
  50132/250000: episode: 46, duration: 3.243s, episode steps: 1318, steps per second: 406, episode reward: -19.000, mean r
  51605/250000: episode: 47, duration: 7.755s, episode steps: 1473, steps per second: 190, episode reward: -20.000, mean r
  53127/250000: episode: 48, duration: 7.921s, episode steps: 1522, steps per second: 192, episode reward: -18.000, mean r
  54612/250000: episode: 49, duration: 7.720s, episode steps: 1485, steps per second: 192, episode reward: -18.000, mean r
  56186/250000: episode: 50, duration: 8.191s, episode steps: 1574, steps per second: 192, episode reward: -18.000, mean r
  57734/250000: episode: 51, duration: 8.030s, episode steps: 1548, steps per second: 193, episode reward: -16.000, mean r
  59158/250000: episode: 52, duration: 7.375s, episode steps: 1424, steps per second: 193, episode reward: -18.000, mean r
  60379/250000: episode: 53, duration: 6.380s, episode steps: 1221, steps per second: 191, episode reward: -20.000, mean r
  62015/250000: episode: 54, duration: 8.476s, episode steps: 1636, steps per second: 193, episode reward: -18.000, mean r
  63473/250000: episode: 55, duration: 7.561s, episode steps: 1458, steps per second: 193, episode reward: -18.000, mean r
  64479/250000: episode: 56, duration: 5.221s, episode steps: 1006, steps per second: 193, episode reward: -21.000, mean r
  65983/250000: episode: 57, duration: 7.831s, episode steps: 1504, steps per second: 192, episode reward: -19.000, mean r
  67532/250000: episode: 58, duration: 8.051s, episode steps: 1549, steps per second: 192, episode reward: -19.000, mean r
  68991/250000: episode: 59, duration: 7.557s, episode steps: 1459, steps per second: 193, episode reward: -18.000, mean r
  70395/250000: episode: 60, duration: 7.271s, episode steps: 1404, steps per second: 193, episode reward: -21.000, mean r
  71891/250000: episode: 61, duration: 7.927s, episode steps: 1496, steps per second: 189, episode reward: -17.000, mean r
  73605/250000: episode: 62, duration: 8.933s, episode steps: 1714, steps per second: 192, episode reward: -19.000, mean r
  75439/250000: episode: 63, duration: 9.589s, episode steps: 1834, steps per second: 191, episode reward: -20.000, mean r
  77428/250000: episode: 64, duration: 10.309s, episode steps: 1989, steps per second: 193, episode reward: -18.000, mean
  79140/250000: episode: 65, duration: 8.916s, episode steps: 1712, steps per second: 192, episode reward: -16.000, mean r
  80803/250000: episode: 66, duration: 8.623s, episode steps: 1663, steps per second: 193, episode reward: -18.000, mean r
  83090/250000: episode: 67, duration: 11.801s, episode steps: 2287, steps per second: 194, episode reward: -15.000, mean
  84777/250000: episode: 68, duration: 8.705s, episode steps: 1687, steps per second: 194, episode reward: -19.000, mean r
  86724/250000: episode: 69, duration: 10.035s, episode steps: 1947, steps per second: 194, episode reward: -17.000, mean
  88779/250000: episode: 70, duration: 10.636s, episode steps: 2055, steps per second: 193, episode reward: -15.000, mean
  90973/250000: episode: 71, duration: 11.467s, episode steps: 2194, steps per second: 191, episode reward: -17.000, mean
  92494/250000: episode: 72, duration: 7.826s, episode steps: 1521, steps per second: 194, episode reward: -19.000, mean r
  94805/250000: episode: 73, duration: 11.899s, episode steps: 2311, steps per second: 194, episode reward: -15.000, mean
  97123/250000: episode: 74, duration: 11.965s, episode steps: 2318, steps per second: 194, episode reward: -16.000, mean
  99384/250000: episode: 75, duration: 11.979s, episode steps: 2261, steps per second: 189, episode reward: -13.000, mean
 101635/250000: episode: 76, duration: 11.698s, episode steps: 2251, steps per second: 192, episode reward: -17.000, mean
 103470/250000: episode: 77, duration: 9.429s, episode steps: 1835, steps per second: 195, episode reward: -20.000, mean r
 106185/250000: episode: 78, duration: 14.067s, episode steps: 2715, steps per second: 193, episode reward: -13.000, mean
 108992/250000: episode: 79, duration: 14.463s, episode steps: 2807, steps per second: 194, episode reward: -12.000, mean
 111369/250000: episode: 80, duration: 12.180s, episode steps: 2377, steps per second: 195, episode reward: -14.000, mean
 113041/250000: episode: 81, duration: 8.707s, episode steps: 1672, steps per second: 192, episode reward: -17.000, mean r
 115051/250000: episode: 82, duration: 10.406s, episode steps: 2010, steps per second: 193, episode reward: -17.000, mean
 117115/250000: episode: 83, duration: 10.637s, episode steps: 2064, steps per second: 194, episode reward: -19.000, mean
 119166/250000: episode: 84, duration: 10.646s, episode steps: 2051, steps per second: 193, episode reward: -18.000, mean
 121911/250000: episode: 85, duration: 14.245s, episode steps: 2745, steps per second: 193, episode reward: -10.000, mean
 123682/250000: episode: 86, duration: 9.179s, episode steps: 1771, steps per second: 193, episode reward: -20.000, mean r
 125278/250000: episode: 87, duration: 8.288s, episode steps: 1596, steps per second: 193, episode reward: -20.000, mean r
 127727/250000: episode: 88, duration: 12.613s, episode steps: 2449, steps per second: 194, episode reward: -11.000, mean
 129978/250000: episode: 89, duration: 11.740s, episode steps: 2251, steps per second: 192, episode reward: -15.000, mean
 131692/250000: episode: 90, duration: 8.876s, episode steps: 1714, steps per second: 193, episode reward: -18.000, mean r
 133892/250000: episode: 91, duration: 11.287s, episode steps: 2200, steps per second: 195, episode reward: -16.000, mean
 135726/250000: episode: 92, duration: 9.450s, episode steps: 1834, steps per second: 194, episode reward: -18.000, mean r
 137786/250000: episode: 93, duration: 10.696s, episode steps: 2060, steps per second: 193, episode reward: -17.000, mean
```
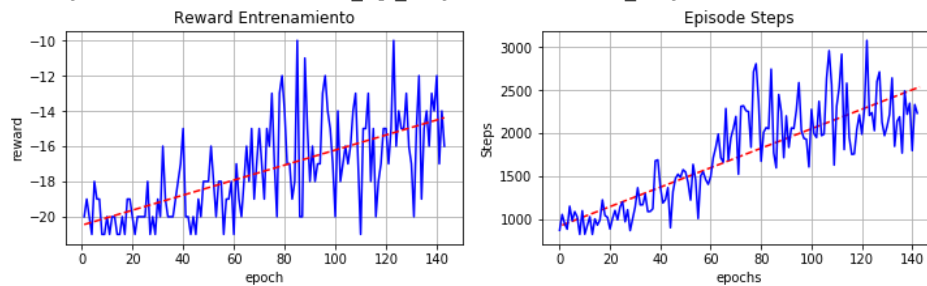
```
139839/250000: episode: 94, duration: 10.562s, episode steps: 2053, steps per second: 194, episode reward: -17.000, mean
142105/250000: episode: 95, duration: 11.740s, episode steps: 2266, steps per second: 193, episode reward: -13.000, mean
144693/250000: episode: 96, duration: 13.323s, episode steps: 2588, steps per second: 194, episode reward: -12.000, mean
146724/250000: episode: 97, duration: 10.460s, episode steps: 2031, steps per second: 194, episode reward: -14.000, mean
148663/250000: episode: 98, duration: 10.041s, episode steps: 1939, steps per second: 193, episode reward: -15.000, mean
150591/250000: episode: 99, duration: 9.945s, episode steps: 1928, steps per second: 194, episode reward: -17.000, mean r
152197/250000: episode: 100, duration: 8.327s, episode steps: 1606, steps per second: 193, episode reward: -20.000, mean
154473/250000: episode: 101, duration: 11.731s, episode steps: 2276, steps per second: 194, episode reward: -14.000, mean
156467/250000: episode: 102, duration: 10.299s, episode steps: 1994, steps per second: 194, episode reward: -18.000, mean
158416/250000: episode: 103, duration: 10.246s, episode steps: 1949, steps per second: 190, episode reward: -17.000, mean
160786/250000: episode: 104, duration: 12.582s, episode steps: 2370, steps per second: 188, episode reward: -16.000, mean
162757/250000: episode: 105, duration: 10.202s, episode steps: 1971, steps per second: 193, episode reward: -17.000, mean
164746/250000: episode: 106, duration: 10.376s, episode steps: 1989, steps per second: 192, episode reward: -16.000, mean
167384/250000: episode: 107, duration: 13.633s, episode steps: 2638, steps per second: 193, episode reward: -14.000, mean
170344/250000: episode: 108, duration: 15.185s, episode steps: 2960, steps per second: 195, episode reward: -13.000, mean
172908/250000: episode: 109, duration: 13.173s, episode steps: 2564, steps per second: 195, episode reward: -17.000, mean
174536/250000: episode: 110, duration: 8.363s, episode steps: 1628, steps per second: 195, episode reward: -21.000, mean
176850/250000: episode: 111, duration: 11.855s, episode steps: 2314, steps per second: 195, episode reward: -15.000, mean
179379/250000: episode: 112, duration: 12.972s, episode steps: 2529, steps per second: 195, episode reward: -15.000, mean
182298/250000: episode: 113, duration: 14.998s, episode steps: 2919, steps per second: 195, episode reward: -13.000, mean
184107/250000: episode: 114, duration: 9.358s, episode steps: 1809, steps per second: 193, episode reward: -18.000, mean
186688/250000: episode: 115, duration: 13.248s, episode steps: 2581, steps per second: 195, episode reward: -15.000, mean
188638/250000: episode: 116, duration: 10.091s, episode steps: 1950, steps per second: 193, episode reward: -20.000, mean
190391/250000: episode: 117, duration: 9.026s, episode steps: 1753, steps per second: 194, episode reward: -18.000, mean
192149/250000: episode: 118, duration: 9.006s, episode steps: 1758, steps per second: 195, episode reward: -17.000, mean
194190/250000: episode: 119, duration: 10.461s, episode steps: 2041, steps per second: 195, episode reward: -15.000, mean
196404/250000: episode: 120, duration: 11.398s, episode steps: 2214, steps per second: 194, episode reward: -15.000, mean
198391/250000: episode: 121, duration: 10.290s, episode steps: 1987, steps per second: 193, episode reward: -17.000, mean
200662/250000: episode: 122, duration: 11.752s, episode steps: 2271, steps per second: 193, episode reward: -15.000, mean
203741/250000: episode: 123, duration: 15.793s, episode steps: 3079, steps per second: 195, episode reward: -10.000, mean
205947/250000: episode: 124, duration: 11.293s, episode steps: 2206, steps per second: 195, episode reward: -16.000, mean
208186/250000: episode: 125, duration: 11.454s, episode steps: 2239, steps per second: 195, episode reward: -14.000, mean
210215/250000: episode: 126, duration: 10.465s, episode steps: 2029, steps per second: 194, episode reward: -15.000, mean
212814/250000: episode: 127, duration: 13.340s, episode steps: 2599, steps per second: 195, episode reward: -15.000, mean
215527/250000: episode: 128, duration: 14.069s, episode steps: 2713, steps per second: 193, episode reward: -13.000, mean
217656/250000: episode: 129, duration: 10.957s, episode steps: 2129, steps per second: 194, episode reward: -16.000, mean
219627/250000: episode: 130, duration: 10.556s, episode steps: 1971, steps per second: 187, episode reward: -17.000, mean
221703/250000: episode: 131, duration: 10.702s, episode steps: 2076, steps per second: 194, episode reward: -20.000, mean
223920/250000: episode: 132, duration: 11.338s, episode steps: 2217, steps per second: 196, episode reward: -15.000, mean
226562/250000: episode: 133, duration: 13.499s, episode steps: 2642, steps per second: 196, episode reward: -12.000, mean
228408/250000: episode: 134, duration: 9.496s, episode steps: 1846, steps per second: 194, episode reward: -19.000, mean
230545/250000: episode: 135, duration: 11.052s, episode steps: 2137, steps per second: 193, episode reward: -15.000, mean
232735/250000: episode: 136, duration: 11.235s, episode steps: 2190, steps per second: 195, episode reward: -14.000, mean
234503/250000: episode: 137, duration: 9.109s, episode steps: 1768, steps per second: 194, episode reward: -16.000, mean
236994/250000: episode: 138, duration: 12.775s, episode steps: 2491, steps per second: 195, episode reward: -13.000, mean
239213/250000: episode: 139, duration: 11.468s, episode steps: 2219, steps per second: 193, episode reward: -14.000, mean
241561/250000: episode: 140, duration: 12.098s, episode steps: 2348, steps per second: 194, episode reward: -12.000, mean
243358/250000: episode: 141, duration: 9.226s, episode steps: 1797, steps per second: 195, episode reward: -17.000, mean
245690/250000: episode: 142, duration: 12.070s, episode steps: 2332, steps per second: 193, episode reward: -14.000, mean
247922/250000: episode: 143, duration: 11.633s, episode steps: 2232, steps per second: 192, episode reward: -16.000, mean
done, took 1147.839 seconds
Pesos guardados en: 202001072037_dqn_PongDeterministic-v0_weights.h5f
```


Reward Entrenamiento | Episode Steps

```
Testing for 10 episodes ...
Episode 1: reward: -14.000, steps: 2507
Episode 2: reward: -11.000, steps: 2697
Episode 3: reward: -15.000, steps: 2399
Episode 4: reward: -12.000, steps: 2349
Episode 5: reward: -12.000, steps: 2558
Episode 6: reward: -10.000, steps: 2956
Episode 7: reward: -12.000, steps: 2781
Episode 8: reward: -10.000, steps: 2730
Episode 9: reward: -19.000, steps: 2261
Episode 10: reward: -18.000, steps: 2473
mean reward: -13.3
mean_steps: 2571.1
(-13.3, 2571.1)
```

```
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)
```