

```

=====
Entrenando ...
Entrenando batch: 1
experimento: Experimento_Arquitectura_9
arquitectura: 9
train_steps: 450000
policy_nb_steps: 100000
agent_nb_steps_warmup: 3000
agent_target_model_update: 10000
agent_batch_size: 32
agent_train_interval: 4
learning_rate: 0.0001
memory_limit: 10000
optimizer_type: Adam
history_reward: []
history_steps: []
*****
Ejecutando experimento ... Experimento_Arquitectura_9
input_shape: (4, 84, 84)
arquitectura: 9
4 (84, 84)
Usando arquitectura: 9

```

Layer (type)	Output Shape	Param #
permute_12 (Permute)	(None, 84, 84, 4)	0
dense_29 (Dense)	(None, 84, 84, 64)	320
dense_30 (Dense)	(None, 84, 84, 32)	2080
dense_31 (Dense)	(None, 84, 84, 16)	528
flatten_9 (Flatten)	(None, 112896)	0
dense_32 (Dense)	(None, 6)	677382
activation_42 (Activation)	(None, 6)	0
Total params: 680,310		
Trainable params: 680,310		
Non-trainable params: 0		

```

Usando: Adam
Modelo inicial guardado en: 202001091630_gs_Experimento_Arquitectura_9_modelo_inicial.h5
Training for 450000 steps ...

```

```

812/450000: episode: 1, duration: 5.899s, episode steps: 812, steps per second: 138, episode reward: -21.000, mean reward: -0.026 [-1.000, 0.000], mean action: 2.
2041/450000: episode: 2, duration: 8.229s, episode steps: 1229, steps per second: 149, episode reward: -18.000, mean reward: -0.015 [-1.000, 1.000], mean action: 2.
2865/450000: episode: 3, duration: 5.469s, episode steps: 824, steps per second: 151, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 2.
3629/450000: episode: 4, duration: 5.309s, episode steps: 764, steps per second: 144, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 2.
4393/450000: episode: 5, duration: 5.335s, episode steps: 764, steps per second: 143, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 2.
5374/450000: episode: 6, duration: 6.775s, episode steps: 981, steps per second: 145, episode reward: -20.000, mean reward: -0.020 [-1.000, 1.000], mean action: 2.
6257/450000: episode: 7, duration: 6.147s, episode steps: 883, steps per second: 144, episode reward: -21.000, mean reward: -0.024 [-1.000, 0.000], mean action: 2.
7118/450000: episode: 8, duration: 5.792s, episode steps: 861, steps per second: 149, episode reward: -20.000, mean reward: -0.023 [-1.000, 1.000], mean action: 2.
7882/450000: episode: 9, duration: 5.365s, episode steps: 764, steps per second: 142, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 2.
8867/450000: episode: 10, duration: 6.617s, episode steps: 985, steps per second: 149, episode reward: -19.000, mean reward: -0.019 [-1.000, 1.000], mean action: 2.
9716/450000: episode: 11, duration: 5.663s, episode steps: 849, steps per second: 150, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 2.
10829/450000: episode: 12, duration: 26.559s, episode steps: 1113, steps per second: 42, episode reward: -20.000, mean reward: -0.018 [-1.000, 1.000], mean action: 2.
11710/450000: episode: 13, duration: 23.627s, episode steps: 881, steps per second: 37, episode reward: -21.000, mean reward: -0.024 [-1.000, 0.000], mean action: 2.
12641/450000: episode: 14, duration: 24.995s, episode steps: 931, steps per second: 37, episode reward: -21.000, mean reward: -0.023 [-1.000, 0.000], mean action: 2.
13527/450000: episode: 15, duration: 23.826s, episode steps: 886, steps per second: 37, episode reward: -21.000, mean reward: -0.024 [-1.000, 0.000], mean action: 2.
14320/450000: episode: 16, duration: 21.680s, episode steps: 793, steps per second: 37, episode reward: -21.000, mean reward: -0.026 [-1.000, 0.000], mean action: 2.
15191/450000: episode: 17, duration: 23.832s, episode steps: 871, steps per second: 37, episode reward: -21.000, mean reward: -0.024 [-1.000, 0.000], mean action: 2.
15955/450000: episode: 18, duration: 21.056s, episode steps: 764, steps per second: 36, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 2.
16719/450000: episode: 19, duration: 20.756s, episode steps: 764, steps per second: 37, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 2.
17592/450000: episode: 20, duration: 23.994s, episode steps: 873, steps per second: 36, episode reward: -21.000, mean reward: -0.024 [-1.000, 0.000], mean action: 2.
18595/450000: episode: 21, duration: 27.296s, episode steps: 1003, steps per second: 37, episode reward: -20.000, mean reward: -0.020 [-1.000, 1.000], mean action: 2.
19724/450000: episode: 22, duration: 31.337s, episode steps: 1129, steps per second: 36, episode reward: -19.000, mean reward: -0.017 [-1.000, 1.000], mean action: 2.
20700/450000: episode: 23, duration: 26.755s, episode steps: 976, steps per second: 36, episode reward: -19.000, mean reward: -0.019 [-1.000, 1.000], mean action: 2.
22017/450000: episode: 24, duration: 35.535s, episode steps: 1317, steps per second: 37, episode reward: -19.000, mean reward: -0.014 [-1.000, 1.000], mean action: 2.
22888/450000: episode: 25, duration: 23.496s, episode steps: 871, steps per second: 37, episode reward: -20.000, mean reward: -0.023 [-1.000, 1.000], mean action: 2.
23791/450000: episode: 26, duration: 24.429s, episode steps: 903, steps per second: 37, episode reward: -20.000, mean reward: -0.022 [-1.000, 1.000], mean action: 2.
24750/450000: episode: 27, duration: 25.933s, episode steps: 959, steps per second: 37, episode reward: -20.000, mean reward: -0.021 [-1.000, 1.000], mean action: 2.
25621/450000: episode: 28, duration: 23.464s, episode steps: 871, steps per second: 37, episode reward: -20.000, mean reward: -0.023 [-1.000, 1.000], mean action: 2.
26462/450000: episode: 29, duration: 22.776s, episode steps: 841, steps per second: 37, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 2.
27520/450000: episode: 30, duration: 28.619s, episode steps: 1058, steps per second: 37, episode reward: -20.000, mean reward: -0.019 [-1.000, 1.000], mean action: 2.
28505/450000: episode: 31, duration: 26.727s, episode steps: 985, steps per second: 37, episode reward: -20.000, mean reward: -0.020 [-1.000, 1.000], mean action: 2.
29348/450000: episode: 32, duration: 22.747s, episode steps: 843, steps per second: 37, episode reward: -20.000, mean reward: -0.024 [-1.000, 1.000], mean action: 2.
30345/450000: episode: 33, duration: 26.922s, episode steps: 997, steps per second: 37, episode reward: -19.000, mean reward: -0.019 [-1.000, 1.000], mean action: 2.
31156/450000: episode: 34, duration: 21.803s, episode steps: 811, steps per second: 37, episode reward: -21.000, mean reward: -0.026 [-1.000, 0.000], mean action: 2.
32146/450000: episode: 35, duration: 26.727s, episode steps: 990, steps per second: 37, episode reward: -21.000, mean reward: -0.021 [-1.000, 0.000], mean action: 2.
33019/450000: episode: 36, duration: 23.469s, episode steps: 873, steps per second: 37, episode reward: -21.000, mean reward: -0.024 [-1.000, 0.000], mean action: 2.
33892/450000: episode: 37, duration: 23.662s, episode steps: 873, steps per second: 37, episode reward: -20.000, mean reward: -0.023 [-1.000, 1.000], mean action: 2.
34810/450000: episode: 38, duration: 24.803s, episode steps: 918, steps per second: 37, episode reward: -21.000, mean reward: -0.023 [-1.000, 0.000], mean action: 2.
35694/450000: episode: 39, duration: 23.831s, episode steps: 884, steps per second: 37, episode reward: -21.000, mean reward: -0.024 [-1.000, 0.000], mean action: 2.
36848/450000: episode: 40, duration: 31.171s, episode steps: 1154, steps per second: 37, episode reward: -20.000, mean reward: -0.017 [-1.000, 1.000], mean action: 2.
37671/450000: episode: 41, duration: 22.378s, episode steps: 823, steps per second: 37, episode reward: -21.000, mean reward: -0.026 [-1.000, 0.000], mean action: 2.
38693/450000: episode: 42, duration: 27.763s, episode steps: 1022, steps per second: 37, episode reward: -20.000, mean reward: -0.020 [-1.000, 1.000], mean action: 2.
39545/450000: episode: 43, duration: 23.080s, episode steps: 852, steps per second: 37, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 2.
40475/450000: episode: 44, duration: 25.249s, episode steps: 930, steps per second: 37, episode reward: -20.000, mean reward: -0.022 [-1.000, 1.000], mean action: 2.
41299/450000: episode: 45, duration: 22.233s, episode steps: 824, steps per second: 37, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 2.
42229/450000: episode: 46, duration: 25.093s, episode steps: 930, steps per second: 37, episode reward: -20.000, mean reward: -0.022 [-1.000, 1.000], mean action: 2.
42993/450000: episode: 47, duration: 20.517s, episode steps: 764, steps per second: 37, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 2.
43937/450000: episode: 48, duration: 25.626s, episode steps: 944, steps per second: 37, episode reward: -21.000, mean reward: -0.022 [-1.000, 0.000], mean action: 2.
44929/450000: episode: 49, duration: 27.009s, episode steps: 992, steps per second: 37, episode reward: -20.000, mean reward: -0.020 [-1.000, 1.000], mean action: 2.
45979/450000: episode: 50, duration: 28.569s, episode steps: 1050, steps per second: 37, episode reward: -19.000, mean reward: -0.018 [-1.000, 1.000], mean action: 2.
47032/450000: episode: 51, duration: 28.831s, episode steps: 1053, steps per second: 37, episode reward: -20.000, mean reward: -0.019 [-1.000, 1.000], mean action: 2.
48024/450000: episode: 52, duration: 26.984s, episode steps: 992, steps per second: 37, episode reward: -19.000, mean reward: -0.019 [-1.000, 1.000], mean action: 2.
48788/450000: episode: 53, duration: 21.029s, episode steps: 764, steps per second: 36, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 2.
49640/450000: episode: 54, duration: 23.520s, episode steps: 852, steps per second: 36, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 2.
50513/450000: episode: 55, duration: 24.083s, episode steps: 873, steps per second: 36, episode reward: -20.000, mean reward: -0.023 [-1.000, 1.000], mean action: 2.
51624/450000: episode: 56, duration: 30.382s, episode steps: 1111, steps per second: 37, episode reward: -19.000, mean reward: -0.017 [-1.000, 1.000], mean action: 2.
52532/450000: episode: 57, duration: 24.990s, episode steps: 908, steps per second: 36, episode reward: -21.000, mean reward: -0.023 [-1.000, 0.000], mean action: 2.
53572/450000: episode: 58, duration: 28.767s, episode steps: 1040, steps per second: 36, episode reward: -19.000, mean reward: -0.018 [-1.000, 1.000], mean action: 2.
54416/450000: episode: 59, duration: 23.002s, episode steps: 844, steps per second: 37, episode reward: -20.000, mean reward: -0.024 [-1.000, 1.000], mean action: 2.
55276/450000: episode: 60, duration: 23.370s, episode steps: 860, steps per second: 37, episode reward: -20.000, mean reward: -0.023 [-1.000, 1.000], mean action: 2.
56215/450000: episode: 61, duration: 25.603s, episode steps: 939, steps per second: 37, episode reward: -19.000, mean reward: -0.020 [-1.000, 1.000], mean action: 2.
57238/450000: episode: 62, duration: 28.190s, episode steps: 1023, steps per second: 36, episode reward: -19.000, mean reward: -0.019 [-1.000, 1.000], mean action: 2.

```





259296/450000:	episode:	281,	duration:	21.366s,	episode steps:	854,	steps per second:	39,	episode reward:	-21.000,	mean reward:	-0.025	[-1.000, 0.000],	mean action:
259240/450000:	episode:	282,	duration:	25.530s,	episode steps:	995,	steps per second:	39,	episode reward:	-20.000,	mean reward:	-0.020	[-1.000, 1.000],	mean action:
254012/450000:	episode:	283,	duration:	28.465s,	episode steps:	1091,	steps per second:	38,	episode reward:	-20.000,	mean reward:	-0.018	[-1.000, 1.000],	mean action:
254926/450000:	episode:	284,	duration:	23.548s,	episode steps:	914,	steps per second:	39,	episode reward:	-21.000,	mean reward:	-0.023	[-1.000, 0.000],	mean action:
256040/450000:	episode:	285,	duration:	28.718s,	episode steps:	1114,	steps per second:	39,	episode reward:	-21.000,	mean reward:	-0.019	[-1.000, 0.000],	mean action:
256927/450000:	episode:	286,	duration:	22.829s,	episode steps:	887,	steps per second:	39,	episode reward:	-21.000,	mean reward:	-0.023	[-1.000, 1.000],	mean action:
257747/450000:	episode:	287,	duration:	21.072s,	episode steps:	820,	steps per second:	39,	episode reward:	-20.000,	mean reward:	-0.026	[-1.000, 0.000],	mean action:
258767/450000:	episode:	288,	duration:	26.178s,	episode steps:	1020,	steps per second:	39,	episode reward:	-20.000,	mean reward:	-0.020	[-1.000, 1.000],	mean action:
259803/450000:	episode:	289,	duration:	27.333s,	episode steps:	1036,	steps per second:	38,	episode reward:	-19.000,	mean reward:	-0.018	[-1.000, 1.000],	mean action:
260567/450000:	episode:	290,	duration:	20.824s,	episode steps:	764,	steps per second:	37,	episode reward:	-21.000,	mean reward:	-0.027	[-1.000, 0.000],	mean action:
261561/450000:	episode:	291,	duration:	27.290s,	episode steps:	994,	steps per second:	36,	episode reward:	-20.000,	mean reward:	-0.020	[-1.000, 1.000],	mean action:
262440/450000:	episode:	292,	duration:	23.740s,	episode steps:	879,	steps per second:	37,	episode reward:	-21.000,	mean reward:	-0.024	[-1.000, 0.000],	mean action:
263284/450000:	episode:	293,	duration:	22.849s,	episode steps:	844,	steps per second:	37,	episode reward:	-20.000,	mean reward:	-0.024	[-1.000, 1.000],	mean action:
264048/450000:	episode:	294,	duration:	20.700s,	episode steps:	764,	steps per second:	37,	episode reward:	-21.000,	mean reward:	-0.027	[-1.000, 0.000],	mean action:
265090/450000:	episode:	295,	duration:	28.287s,	episode steps:	1042,	steps per second:	37,	episode reward:	-20.000,	mean reward:	-0.019	[-1.000, 1.000],	mean action:
265854/450000:	episode:	296,	duration:	20.809s,	episode steps:	764,	steps per second:	37,	episode reward:	-21.000,	mean reward:	-0.027	[-1.000, 0.000],	mean action:
266676/450000:	episode:	297,	duration:	22.281s,	episode steps:	822,	steps per second:	37,	episode reward:	-21.000,	mean reward:	-0.026	[-1.000, 0.000],	mean action:
267440/450000:	episode:	298,	duration:	20.557s,	episode steps:	764,	steps per second:	37,	episode reward:	-21.000,	mean reward:	-0.027	[-1.000, 0.000],	mean action:
268232/450000:	episode:	299,	duration:	21.635s,	episode steps:	792,	steps per second:	37,	episode reward:	-21.000,	mean reward:	-0.027	[-1.000, 0.000],	mean action:
269174/450000:	episode:	300,	duration:	25.841s,	episode steps:	942,	steps per second:	36,	episode reward:	-21.000,	mean reward:	-0.022	[-1.000, 0.000],	mean action:
270045/450000:	episode:	301,	duration:	23.621s,	episode steps:	871,	steps per second:	37,	episode reward:	-20.000,	mean reward:	-0.023	[-1.000, 1.000],	mean action:
270915/450000:	episode:	302,	duration:	23.770s,	episode steps:	870,	steps per second:	37,	episode reward:	-20.000,	mean reward:	-0.023	[-1.000, 1.000],	mean action:
271815/450000:	episode:	303,	duration:	24.450s,	episode steps:	900,	steps per second:	37,	episode reward:	-20.000,	mean reward:	-0.022	[-1.000, 1.000],	mean action:
272834/450000:	episode:	304,	duration:	28.119s,	episode steps:	1019,	steps per second:	36,	episode reward:	-20.000,	mean reward:	-0.020	[-1.000, 1.000],	mean action:
273686/450000:	episode:	305,	duration:	23.561s,	episode steps:	852,	steps per second:	36,	episode reward:	-21.000,	mean reward:	-0.025	[-1.000, 0.000],	mean action:
274594/450000:	episode:	306,	duration:	24.843s,	episode steps:	908,	steps per second:	37,	episode reward:	-21.000,	mean reward:	-0.023	[-1.000, 0.000],	mean action:
275418/450000:	episode:	307,	duration:	22.820s,	episode steps:	824,	steps per second:	36,	episode reward:	-21.000,	mean reward:	-0.025	[-1.000, 0.000],	mean action:
276378/450000:	episode:	308,	duration:	26.416s,	episode steps:	960,	steps per second:	36,	episode reward:	-19.000,	mean reward:	-0.020	[-1.000, 1.000],	mean action:
277198/450000:	episode:													

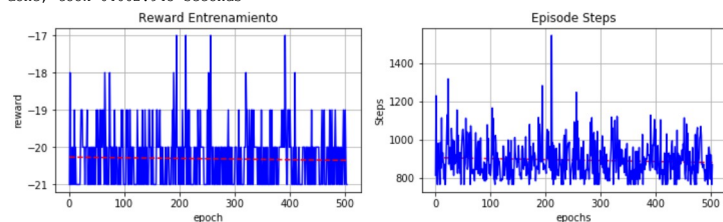
```

352251/450000: episode: 390, duration: 23.49s, episode steps: 872, steps per second: 36, episode reward: -20.000, mean reward: -0.023 [-1.000, 1.000], mean action: 352048/450000: episode: 391, duration: 22.65s, episode steps: 837, steps per second: 37, episode reward: -20.000, mean reward: -0.024 [-1.000, 1.000], mean action: 350215/450000: episode: 392, duration: 30.610s, episode steps: 1127, steps per second: 37, episode reward: -17.000, mean reward: -0.015 [-1.000, 1.000], mean action: 351164/450000: episode: 393, duration: 25.608s, episode steps: 949, steps per second: 37, episode reward: -19.000, mean reward: -0.020 [-1.000, 1.000], mean action: 352048/450000: episode: 394, duration: 24.018s, episode steps: 884, steps per second: 37, episode reward: -20.000, mean reward: -0.023 [-1.000, 1.000], mean action: 352900/450000: episode: 395, duration: 22.932s, episode steps: 852, steps per second: 37, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 353872/450000: episode: 396, duration: 26.287s, episode steps: 972, steps per second: 37, episode reward: -21.000, mean reward: -0.022 [-1.000, 0.000], mean action: 354839/450000: episode: 397, duration: 26.066s, episode steps: 967, steps per second: 37, episode reward: -19.000, mean reward: -0.020 [-1.000, 1.000], mean action: 355863/450000: episode: 398, duration: 27.315s, episode steps: 1024, steps per second: 37, episode reward: -19.000, mean reward: -0.019 [-1.000, 1.000], mean action: 356967/450000: episode: 399, duration: 29.743s, episode steps: 1104, steps per second: 37, episode reward: -19.000, mean reward: -0.017 [-1.000, 1.000], mean action: 357809/450000: episode: 400, duration: 23.075s, episode steps: 842, steps per second: 36, episode reward: -20.000, mean reward: -0.024 [-1.000, 1.000], mean action: 358862/450000: episode: 401, duration: 28.958s, episode steps: 1053, steps per second: 36, episode reward: -19.000, mean reward: -0.018 [-1.000, 1.000], mean action: 359704/450000: episode: 402, duration: 23.073s, episode steps: 842, steps per second: 36, episode reward: -20.000, mean reward: -0.024 [-1.000, 1.000], mean action: 360546/450000: episode: 403, duration: 23.318s, episode steps: 842, steps per second: 36, episode reward: -20.000, mean reward: -0.024 [-1.000, 1.000], mean action: 361366/450000: episode: 404, duration: 22.366s, episode steps: 820, steps per second: 37, episode reward: -21.000, mean reward: -0.026 [-1.000, 0.000], mean action: 362334/450000: episode: 405, duration: 26.500s, episode steps: 968, steps per second: 37, episode reward: -21.000, mean reward: -0.022 [-1.000, 0.000], mean action: 363207/450000: episode: 406, duration: 23.402s, episode steps: 873, steps per second: 37, episode reward: -21.000, mean reward: -0.024 [-1.000, 0.000], mean action: 364077/450000: episode: 407, duration: 22.831s, episode steps: 870, steps per second: 38, episode reward: -20.000, mean reward: -0.023 [-1.000, 1.000], mean action: 364869/450000: episode: 408, duration: 20.895s, episode steps: 792, steps per second: 38, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 365758/450000: episode: 409, duration: 22.879s, episode steps: 889, steps per second: 39, episode reward: -21.000, mean reward: -0.024 [-1.000, 0.000], mean action: 366861/450000: episode: 410, duration: 28.381s, episode steps: 1103, steps per second: 39, episode reward: -18.000, mean reward: -0.016 [-1.000, 1.000], mean action: 367681/450000: episode: 411, duration: 21.091s, episode steps: 820, steps per second: 39, episode reward: -21.000, mean reward: -0.026 [-1.000, 0.000], mean action: 368523/450000: episode: 412, duration: 21.557s, episode steps: 842, steps per second: 39, episode reward: -20.000, mean reward: -0.024 [-1.000, 1.000], mean action: 369287/450000: episode: 413, duration: 19.590s, episode steps: 764, steps per second: 39, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 370051/450000: episode: 414, duration: 19.625s, episode steps: 764, steps per second: 39, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 370891/450000: episode: 415, duration: 21.524s, episode steps: 840, steps per second: 39, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 371711/450000: episode: 416, duration: 21.101s, episode steps: 820, steps per second: 39, episode reward: -21.000, mean reward: -0.026 [-1.000, 0.000], mean action: 372640/450000: episode: 417, duration: 23.927s, episode steps: 929, steps per second: 39, episode reward: -20.000, mean reward: -0.022 [-1.000, 1.000], mean action: 373432/450000: episode: 418, duration: 20.371s, episode steps: 792, steps per second: 39, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 374196/450000: episode: 419, duration: 19.729s, episode steps: 764, steps per second: 39, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 375072/450000: episode: 420, duration: 23.270s, episode steps: 876, steps per second: 38, episode reward: -21.000, mean reward: -0.024 [-1.000, 0.000], mean action: 376101/450000: episode: 421, duration: 27.044s, episode steps: 1029, steps per second: 38, episode reward: -21.000, mean reward: -0.020 [-1.000, 0.000], mean action: 376865/450000: episode: 422, duration: 19.961s, episode steps: 764, steps per second: 38, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 377657/450000: episode: 423, duration: 20.315s, episode steps: 792, steps per second: 39, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 378527/450000: episode: 424, duration: 22.301s, episode steps: 870, steps per second: 39, episode reward: -20.000, mean reward: -0.023 [-1.000, 1.000], mean action: 379456/450000: episode: 425, duration: 23.908s, episode steps: 929, steps per second: 39, episode reward: -20.000, mean reward: -0.022 [-1.000, 1.000], mean action: 380358/450000: episode: 426, duration: 23.220s, episode steps: 902, steps per second: 39, episode reward: -20.000, mean reward: -0.022 [-1.000, 1.000], mean action: 381276/450000: episode: 427, duration: 23.493s, episode steps: 916, steps per second: 39, episode reward: -20.000, mean reward: -0.022 [-1.000, 1.000], mean action: 382066/450000: episode: 428, duration: 20.340s, episode steps: 792, steps per second: 39, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 382974/450000: episode: 429, duration: 23.337s, episode steps: 908, steps per second: 39, episode reward: -21.000, mean reward: -0.023 [-1.000, 0.000], mean action: 383937/450000: episode: 430, duration: 24.771s, episode steps: 963, steps per second: 39, episode reward: -20.000, mean reward: -0.021 [-1.000, 1.000], mean action: 384930/450000: episode: 431, duration: 25.495s, episode steps: 993, steps per second: 39, episode reward: -20.000, mean reward: -0.020 [-1.000, 1.000], mean action: 385773/450000: episode: 432, duration: 21.689s, episode steps: 843, steps per second: 39, episode reward: -20.000, mean reward: -0.024 [-1.000, 1.000], mean action: 386627/450000: episode: 433, duration: 22.500s, episode steps: 854, steps per second: 38, episode reward: -21.000, mean reward: -0.025 [-1.000, 0.000], mean action: 387553/450000: episode: 434, duration: 24.391s, episode steps: 926, steps per second: 38, episode reward: -20.000, mean reward: -0.022 [-1.000, 1.000], mean action: 388317/450000: episode: 435, duration: 19.977s, episode steps: 764, steps per second: 38, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action: 389279/450000: episode: 436, duration: 24.669s, episode steps: 962, steps per second: 39, episode reward: -21.000, mean reward: -0.022 [-1.000, 0.000], mean action: 390149/450000: episode: 437, duration: 22.366s, episode steps: 870, steps per second: 39
```

```

445181/450000: episode: 499, duration: 23.794s, episode steps: 926, steps per second: 39, episode reward: -20.000, mean reward: -0.022 [-1.000, 1.000], mean action:
445973/450000: episode: 500, duration: 20.343s, episode steps: 792, steps per second: 39, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action:
446893/450000: episode: 501, duration: 23.605s, episode steps: 920, steps per second: 39, episode reward: -19.000, mean reward: -0.021 [-1.000, 1.000], mean action:
447809/450000: episode: 502, duration: 23.519s, episode steps: 916, steps per second: 39, episode reward: -20.000, mean reward: -0.022 [-1.000, 1.000], mean action:
448573/450000: episode: 503, duration: 19.611s, episode steps: 764, steps per second: 39, episode reward: -21.000, mean reward: -0.027 [-1.000, 0.000], mean action:
449440/450000: episode: 504, duration: 22.303s, episode steps: 867, steps per second: 39, episode reward: -21.000, mean reward: -0.024 [-1.000, 0.000], mean action:
done, took 64662.943 seconds

```



```

Testing for 10 episodes ...
Episode 1: reward: -21.000, steps: 764
Episode 2: reward: -21.000, steps: 764
Episode 3: reward: -21.000, steps: 792
Episode 4: reward: -21.000, steps: 764
Episode 5: reward: -21.000, steps: 764
Episode 6: reward: -21.000, steps: 764
Episode 7: reward: -21.000, steps: 764
Episode 8: reward: -21.000, steps: 764
Episode 9: reward: -21.000, steps: 764
Episode 10: reward: -21.000, steps: 824
mean_reward: -21.0
mean_steps: 772.8
mean_reward: -21.0
mean_steps 772.8
Pesos guardados en: 202001101028_gs_Experimento_Arquitectura_9--21.0_NOT_BEST_.hdf5W
mean_reward: -21.0 mean_steps: 772.8
-----
La recompensa no mejoró, fin del entrenamiento ...
Tiempo de ejecución del experimento: 64714.0

```