Aproximación de la Ecuación de Movimiento de un Péndulo Simple utilizando Redes Neuronales

Juan José González
Juan Camilo Ramírez
Juan David Osorio

Junio 2024

Abstract

Este informe presenta el proceso y los resultados del Proyecto II del curso de Sistemas Inteligentes II, cuyo objetivo fue entrenar y comparar diferentes modelos DQN (Deep Q-Network) para resolver los problemas de control clásicos en el entorno OpenAI Gym CartPole y MountainCar. Se experimentó con diversas configuraciones de hiperparámetros para evaluar su impacto en el rendimiento del modelo.

1 Introducción

El control de sistemas dinámicos es un área crucial en la inteligencia artificial y la robótica. Este proyecto se centra en el uso del algoritmo DQN para resolver problemas de control en entornos simulados proporcionados por OpenAI Gym. El objetivo es entrenar y comparar diferentes configuraciones de modelos DQN para evaluar su eficacia en los entornos CartPole y MountainCar.

2 Metodología

2.1 Configuración del entorno OpenAI Gym

Se utilizaron los entornos CartPole y MountainCar de OpenAI Gym para entrenar los modelos DQN. Estos entornos proporcionan desafíos clásicos de

control que son ampliamente utilizados para evaluar algoritmos de aprendizaje por refuerzo.

2.2 Arquitectura del modelo DQN

El mejor modelo DQN para MountainCar con 2 capas con 64 neuronas y una función de activación ReLU. La capa de salida es una función de activación lineal.

2.3 Hiperparámetros utilizados

Los hiperparámetros clave que se experimentaron incluyen la tasa de aprendizaje, el tamaño del batch, el factor de descuento (gamma), la tasa de exploración-explotación (epsilon), el número de épocas, el número de capas y el número de neuronas en cada capa. En el caso específico de MountainCar se añadió un hiperparámetro para controlar el peso que tiene la velocidad en la recompensa.

2.4 Estrategia de entrenamiento

Los modelos se entrenaron utilizando la técnica de aprendizaje por refuerzo con experiencia de repetición, donde se almacenan transiciones en un buffer de experiencia y se realizan actualizaciones periódicas de la red objetivo. Además se utilizó la política de epsilon-greedy para controlar la exploración del modelo.

3 Resultados

Los modelos finales muestran una buena capacidad de acción y precisión en los entornos de control de OpenAI Gym. Se evaluaron las recompensas promedio y la tasa de éxito en cada entorno. Los resultados se presentan en forma de gráficos y tablas comparativas.

3.1 Mountain

model	evaluate_max_reward	evaluate_rewards_mean	evaluate_rewards_std
2024-06-15_12-31-16_dqn_mountaincar	-87.0	-109.1	20.34
2024-06-15_13-02-52_dqn_mountaincar	-86.0	-112.9	22.10
2024-06-15_13-39-22_dqn_mountaincar	-146.0	-150.3	2.96
2024-06-15_13-45-28_dqn_mountaincar	-116.0	-183.5	33.01
2024-06-15_12-43-21_dqn_mountaincar	-84.0	-121.7	23.66
2024-06-15_13-00-05_dqn_mountaincar	-85.0	-115.1	26.20
2024-06-15_08-13-32_dqn_mountaincar	-200.0	-200.0	0.00
2024-06-15_13-20-27_dqn_mountaincar	-140.0	-147.8	8.68
2024-06-15_10-54-10_dqn_mountaincar	-108.0	-141.2	11.74
2024-06-15_13-41-53_dqn_mountaincar	-92.0	-130.9	23.40
2024-06-15_12-23-27_dqn_mountaincar	-87.0	-126.7	23.06
2024-06-15_11-21-25_dqn_mountaincar	-110.0	-137.7	19.36
2024-06-15_13-45-47_dqn_mountaincar	-96.0	-134.2	24.63
2024-06-15_11-01-53_dqn_mountaincar	-85.0	-109.1	20.54
2024-06-15_12-35-51_dqn_mountaincar	-89.0	-127.1	22.23
2024-06-15_16-46-16_dqn_mountaincar	-111.0	-128.5	14.23
2024-06-15_12-16-56_dqn_mountaincar	-108.0	-126.6	17.48
2024-06-15_13-48-51_dqn_mountaincar	-93.0	-113.8	14.22
2024-06-15_12-43-26_dqn_mountaincar	-144.0	-154.9	17.40
2024-06-15_13-36-42_dqn_mountaincar	-84.0	-104.4	18.46
2024-06-15_11-57-24_dqn_mountaincar	-118.0	-136.7	16.19
2024-06-15_13-56-21_dqn_mountaincar	-85.0	-114.8	26.32

model	hidden_layers	units_per_layer	activation_per_layer	output_layer_activation
2024-06-15_12-31-16_dqn_mountaincar	2	[64, 64]	['relu', 'relu']	linear
2024-06-15_13-02-52_dqn_mountaincar	2	[64, 64]	['relu', 'relu']	linear
2024-06-15_13-39-22_dqn_mountaincar	2	[64, 64]	['relu', 'relu']	linear
2024-06-15_13-45-28_dqn_mountaincar	2	[64, 64]	['relu', 'relu']	linear
2024-06-15_12-43-21_dqn_mountaincar	2	[64, 64]	['relu', 'relu']	linear
2024-06-15_13-00-05_dqn_mountaincar	2	[64, 64]	['relu', 'relu']	linear
2024-06-15_08-13-32_dqn_mountaincar	2	[64, 64]	['relu', 'relu']	linear
2024-06-15_13-20-27_dqn_mountaincar	2	[64, 64]	['relu', 'relu']	linear
2024-06-15_10-54-10_dqn_mountaincar	2	[64, 64]	['relu', 'relu']	linear
2024-06-15_13-41-53_dqn_mountaincar	2	[64, 64]	['relu', 'relu']	linear
2024-06-15_12-23-27_dqn_mountaincar	2	[64, 64]	['relu', 'relu']	linear
2024-06-15_11-21-25_dqn_mountaincar	2	[64, 64]	['relu', 'relu']	linear
2024-06-15_13-45-47_dqn_mountaincar	2	[64, 64]	['relu', 'relu']	linear
2024-06-15_11-01-53_dqn_mountaincar	2	[64, 64]	['relu', 'relu']	linear
2024-06-15_12-35-51_dqn_mountaincar	2	[64, 64]	['relu', 'relu']	linear
2024-06-15_16-46-16_dqn_mountaincar	2	[64, 64]	['relu', 'relu']	linear
2024-06-15_12-16-56_dqn_mountaincar	2	[64, 64]	['relu', 'relu']	linear
2024-06-15_13-48-51_dqn_mountaincar	2	[64, 64]	['relu', 'relu']	linear
2024-06-15_12-43-26_dqn_mountaincar	2	[64, 64]	['relu', 'relu']	linear
2024-06-15_13-36-42_dqn_mountaincar	2	[64, 64]	['relu', 'relu']	linear
2024-06-15_11-57-24_dqn_mountaincar	2	[64, 64]	['relu', 'relu']	linear
2024-06-15_13-56-21_dqn_mountaincar	2	[64, 64]	['relu', 'relu']	linear

model	num_episodes	learning_rate	epsilon
2024-06-15_12-31-16_dqn_mountaincar	500	0.005	0.6063789448611848
2024-06-15_13-02-52_dqn_mountaincar	500	0.001	0.0009550049507968265
2024-06-15_13-39-22_dqn_mountaincar	200	0.0001	0.0009833015279105794
2024-06-15_13-45-28_dqn_mountaincar	200	0.01	0.13397967485796175
2024-06-15_12-43-21_dqn_mountaincar	500	0.001	0.07341467529625044
2024-06-15_13-00-05_dqn_mountaincar	500	0.001	0.07341467529625044
2024-06-15_08-13-32_dqn_mountaincar	500	0.005	0.6063789448611848
2024-06-15_13-20-27_dqn_mountaincar	500	0.01	0.006570483042414605
2024-06-15_10-54-10_dqn_mountaincar	1000	0.005	0.3676954247709635
2024-06-15_13-41-53_dqn_mountaincar	200	0.005	0.0009833015279105794
2024-06-15_12-23-27_dqn_mountaincar	1000	0.005	0.000993147795920851
2024-06-15_11-21-25_dqn_mountaincar	1000	0.005	0.3676954247709635
2024-06-15_13-45-47_dqn_mountaincar	200	0.005	0.0009833015279105794
2024-06-15_11-01-53_dqn_mountaincar	1000	0.005	0.9048328935585562
2024-06-15_12-35-51_dqn_mountaincar	500	0.001	0.006570483042414605
2024-06-15_16-46-16_dqn_mountaincar	200	0.001	0.13397967485796175
2024-06-15_12-16-56_dqn_mountaincar	10000	0.005	0.3678610464329692
2024-06-15_13-48-51_dqn_mountaincar	200	0.001	0.13397967485796175
2024-06-15_12-43-26_dqn_mountaincar	1000	0.01	0.3676954247709635
2024-06-15_13-36-42_dqn_mountaincar	200	0.001	0.0009833015279105794
2024-06-15_11-57-24_dqn_mountaincar	1000	0.01	0.3676954247709635
2024-06-15_13-56-21_dqn_mountaincar	1000	0.001	0.000993147795920851

model	epsilon_min	epsilon_decay	batch_size
2024-06-15_12-31-16_dqn_mountaincar	0.001	0.999	128
2024-06-15_13-02-52_dqn_mountaincar	0.001	0.9	64
2024-06-15_13-39-22_dqn_mountaincar	0.001	0.95	64
2024-06-15_13-45-28_dqn_mountaincar	0.001	0.99	32
2024-06-15_12-43-21_dqn_mountaincar	0.001	0.995	256
2024-06-15_13-00-05_dqn_mountaincar	0.001	0.995	256
2024-06-15_08-13-32_dqn_mountaincar	0.005	0.999	256
2024-06-15_13-20-27_dqn_mountaincar	0.001	0.99	256
2024-06-15_10-54-10_dqn_mountaincar	0.001	0.999	128
2024-06-15_13-41-53_dqn_mountaincar	0.001	0.95	64
2024-06-15_12-23-27_dqn_mountaincar	0.001	0.99	128
2024-06-15_11-21-25_dqn_mountaincar	0.001	0.999	128
2024-06-15_13-45-47_dqn_mountaincar	0.001	0.95	64
2024-06-15_11-01-53_dqn_mountaincar	0.001	0.9999	128
2024-06-15_12-35-51_dqn_mountaincar	0.001	0.99	128
2024-06-15_16-46-16_dqn_mountaincar	0.001	0.99	32
2024-06-15_12-16-56_dqn_mountaincar	0.001	0.9999	128
2024-06-15_13-48-51_dqn_mountaincar	0.001	0.99	512
2024-06-15_12-43-26_dqn_mountaincar	0.001	0.999	128
2024-06-15_13-36-42_dqn_mountaincar	0.001	0.95	64
2024-06-15_11-57-24_dqn_mountaincar	0.001	0.999	128
2024-06-15_13-56-21_dqn_mountaincar	0.001	0.99	512

model	discount	update_target_every	speed_weight
2024-06-15_12-31-16_dqn_mountaincar	0.9	40	100.0
2024-06-15_13-02-52_dqn_mountaincar	0.9	20	50.0
2024-06-15_13-39-22_dqn_mountaincar	0.9	20	20.0
2024-06-15_13-45-28_dqn_mountaincar	0.9	10	20.0
2024-06-15_12-43-21_dqn_mountaincar	0.9	20	50.0
2024-06-15_13-00-05_dqn_mountaincar	0.9	20	10.0
2024-06-15_08-13-32_dqn_mountaincar	0.95	40	
2024-06-15_13-20-27_dqn_mountaincar	0.9	20	100.0
2024-06-15_10-54-10_dqn_mountaincar	0.9	40	1000.0
2024-06-15_13-41-53_dqn_mountaincar	0.9	20	20.0
2024-06-15_12-23-27_dqn_mountaincar	0.9	40	100.0
2024-06-15_11-21-25_dqn_mountaincar	0.9	40	100.0
2024-06-15_13-45-47_dqn_mountaincar	0.9	20	20.0
2024-06-15_11-01-53_dqn_mountaincar	0.9	40	1000.0
2024-06-15_12-35-51_dqn_mountaincar	0.9	40	100.0
2024-06-15_16-46-16_dqn_mountaincar	0.9	10	100.0
2024-06-15_12-16-56_dqn_mountaincar	0.9	40	1000.0
2024-06-15_13-48-51_dqn_mountaincar	0.7	40	100.0
2024-06-15_12-43-26_dqn_mountaincar	0.9	40	100.0
2024-06-15_13-36-42_dqn_mountaincar	0.9	20	20.0
2024-06-15_11-57-24_dqn_mountaincar	0.9	40	100.0
2024-06-15_13-56-21_dqn_mountaincar	0.7	40	100.0

3.2 CartPole

model	evaluate_max_reward	evaluate_rewards_mean	evaluate_rewards_std
2024-06-13_17-34-57_dqn_cartpole	10.0	10.0	0.0
2024-06-14_00-20-45_dqn_cartpole	86.0	86.0	0.0
2024-06-13_21-08-24_dqn_cartpole	92.0	92.0	0.0
2024-06-14_09-34-50_dqn_cartpole	201.0	201.0	0.0
2024-06-13_18-59-17_dqn_cartpole	86.0	86.0	0.0
2024-06-14_17-31-42_dqn_cartpole	214.0	214.0	0.0
2024-06-14_18-54-27_dqn_cartpole	9.0	9.0	0.0
2024-06-14_18-17-10_dqn_cartpole	95.0	95.0	0.0
2024-06-13_18-42-33_dqn_cartpole	9.0	9.0	0.0
2024-06-13_23-08-40_dqn_cartpole	89.0	89.0	0.0
2024-06-14_16-42-26_dqn_cartpole	119.0	119.0	0.0
2024-06-15_00-29-02_dqn_cartpole	10.0	10.0	0.0
2024-06-13_18-50-05_dqn_cartpole	336.0	336.0	0.0
2024-06-14_18-50-42_dqn_cartpole	500.0	500.0	0.0
2024-06-13_19-11-27_dqn_cartpole	88.0	88.0	0.0
2024-06-14_13-15-56_dqn_cartpole	9.0	9.0	0.0
2024-06-14_15-38-10_dqn_cartpole	214.0	214.0	0.0
2024-06-13_18-56-41_dqn_cartpole	10.0	10.0	0.0
2024-06-14_11-09-14_dqn_cartpole	255.0	255.0	0.0
2024-06-13_19-37-45_dqn_cartpole	9.0	9.0	0.0
2024-06-13_21-45-52_dqn_cartpole	230.0	230.0	0.0
2024-06-13_22-34-12_dqn_cartpole	96.0	96.0	0.0
2024-06-13_20-39-12_dqn_cartpole	14.0	14.0	0.0
2024-06-14_19-20-40_dqn_cartpole	500.0	500.0	0.0
2024-06-14_18-05-51_dqn_cartpole	97.0	97.0	0.0
2024-06-14_13-24-09_dqn_cartpole	44.0	44.0	0.0
2024-06-13_23-49-30_dqn_cartpole	92.0	92.0	0.0
2024-06-14_17-21-25_dqn_cartpole	86.0	86.0	0.0
2024-06-13_22-09-48_dqn_cartpole	90.0	90.0	0.0
2024-06-13_19-16-11_dqn_cartpole	49.0	49.0	0.0

model	hidden_layers	units_per_layer	activation_per_layer
2024-06-13_17-34-57_dqn_cartpole	4	[512, 512, 512, 512]	['relu', 'relu', 'relu', 'relu']
2024-06-14_00-20-45_dqn_cartpole	2	[256, 256]	['relu', 'relu']
2024-06-13_21-08-24_dqn_cartpole	2	[256, 256]	['relu', 'relu']
2024-06-14_09-34-50_dqn_cartpole	2	[64, 64]	['relu', 'relu']
2024-06-13_18-59-17_dqn_cartpole	2	[256, 256]	['relu', 'relu']
2024-06-14_17-31-42_dqn_cartpole	2	[64, 64]	['relu', 'relu']
2024-06-14_18-54-27_dqn_cartpole	2	[64, 64]	['relu', 'relu']
2024-06-14_18-17-10_dqn_cartpole	2	[64, 64]	['relu', 'relu']
2024-06-13_18-42-33_dqn_cartpole	2	[512, 512]	['relu', 'relu']
2024-06-13_23-08-40_dqn_cartpole	2	[64, 64]	['relu', 'relu']
2024-06-14_16-42-26_dqn_cartpole	2	[64, 64]	['relu', 'relu']
2024-06-15_00-29-02_dqn_cartpole	2	[64, 64]	['relu', 'relu']
2024-06-13_18-50-05_dqn_cartpole	2	[256, 256]	['relu', 'relu']
2024-06-14_18-50-42_dqn_cartpole	2	[64, 64]	['relu', 'relu']
2024-06-13_19-11-27_dqn_cartpole	2	[256, 256]	['relu', 'relu']
2024-06-14_13-15-56_dqn_cartpole	2	[256, 256]	['relu', 'relu']
2024-06-14_15-38-10_dqn_cartpole	2	[64, 64]	['relu', 'relu']
2024-06-13_18-56-41_dqn_cartpole	2	[128, 128]	['relu', 'relu']
2024-06-14_11-09-14_dqn_cartpole	2	[64, 64]	['relu', 'relu']
2024-06-13_19-37-45_dqn_cartpole	2	[256, 256]	['relu', 'relu']
2024-06-13_21-45-52_dqn_cartpole	2	[256, 256]	['relu', 'relu']
2024-06-13_22-34-12_dqn_cartpole	2	[64, 64]	['relu', 'relu']
2024-06-13_20-39-12_dqn_cartpole	2	[256, 256]	['relu', 'relu']
2024-06-14_19-20-40_dqn_cartpole	2	[64, 64]	['relu', 'relu']
2024-06-14_18-05-51_dqn_cartpole	2	[64, 64]	['relu', 'relu']
2024-06-14_13-24-09_dqn_cartpole	2	[64, 64]	['relu', 'relu']
2024-06-13_23-49-30_dqn_cartpole	2	[64, 64]	['relu', 'relu']
2024-06-14_17-21-25_dqn_cartpole	2	[64, 64]	['relu', 'relu']
2024-06-13_22-09-48_dqn_cartpole	2	[256, 256]	['relu', 'relu']
2024-06-13_19-16-11_dqn_cartpole	2	[256, 256]	['relu', 'relu']

model	output_layer_activation	num_episodes	learning_rate
2024-06-13_17-34-57_dqn_cartpole	linear	10000	0.01
2024-06-14_00-20-45_dqn_cartpole	linear	6000	0.001
2024-06-13_21-08-24_dqn_cartpole	linear	5000	0.001
2024-06-14_09-34-50_dqn_cartpole	linear	4000	0.001
2024-06-13_18-59-17_dqn_cartpole	linear	3000	0.01
2024-06-14_17-31-42_dqn_cartpole	linear	6000	0.001
2024-06-14_18-54-27_dqn_cartpole	sigmoid	6000	0.005
2024-06-14_18-17-10_dqn_cartpole	linear	6000	0.005
2024-06-13_18-42-33_dqn_cartpole	linear	10000	0.01
2024-06-13_23-08-40_dqn_cartpole	linear	4000	0.001
2024-06-14_16-42-26_dqn_cartpole	linear	4000	0.001
2024-06-15_00-29-02_dqn_cartpole	linear	7000	0.001
2024-06-13_18-50-05_dqn_cartpole	linear	3000	0.01
2024-06-14_18-50-42_dqn_cartpole	linear	6000	0.005
2024-06-13_19-11-27_dqn_cartpole	linear	3000	0.005
2024-06-14_13-15-56_dqn_cartpole	linear	4000	0.001
2024-06-14_15-38-10_dqn_cartpole	linear	4000	0.001
2024-06-13_18-56-41_dqn_cartpole	linear	3000	0.01
2024-06-14_11-09-14_dqn_cartpole	linear	4000	0.001
2024-06-13_19-37-45_dqn_cartpole	linear	6000	0.005
2024-06-13_21-45-52_dqn_cartpole	linear	5000	0.001
2024-06-13_22-34-12_dqn_cartpole	linear	5000	0.001
2024-06-13_20-39-12_dqn_cartpole	linear	5000	0.005
2024-06-14_19-20-40_dqn_cartpole	linear	6000	0.005
2024-06-14_18-05-51_dqn_cartpole	linear	6000	0.001
2024-06-14_13-24-09_dqn_cartpole	linear	4000	0.001
2024-06-13_23-49-30_dqn_cartpole	linear	4000	0.001
2024-06-14_17-21-25_dqn_cartpole	linear	10000	0.001
2024-06-13_22-09-48_dqn_cartpole	linear	6000	0.001
2024-06-13_19-16-11_dqn_cartpole	linear	3000	0.01

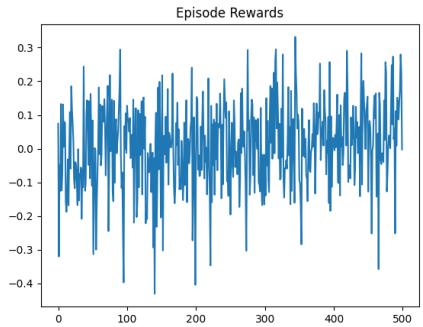
model	epsilon	epsilon_min	epsilon_decay
2024-06-13_17-34-57_dqn_cartpole	0.000993147795920851	0.001	0.99
2024-06-14_00-20-45_dqn_cartpole	0.004998338275642187	0.005	0.999
2024-06-13_21-08-24_dqn_cartpole	0.006721111959865607	0.005	0.999
2024-06-14_09-34-50_dqn_cartpole	0.018279019827489446	0.005	0.999
2024-06-13_18-59-17_dqn_cartpole	0.74080710775091	0.001	0.9999
2024-06-14_17-31-42_dqn_cartpole	0.5487951708942683	0.005	0.9999
2024-06-14_18-54-27_dqn_cartpole	0.004998338275642187	0.005	0.999
2024-06-14_18-17-10_dqn_cartpole	0.1652542546206636	0.005	0.9997
2024-06-13_18-42-33_dqn_cartpole	0.000993147795920851	0.001	0.99
2024-06-13_23-08-40_dqn_cartpole	0.018279019827489446	0.005	0.999
2024-06-14_16-42-26_dqn_cartpole	0.018279019827489446	0.005	0.999
2024-06-15_00-29-02_dqn_cartpole	0.004998338275642187	0.005	0.999
2024-06-13_18-50-05_dqn_cartpole	0.04971239399803625	0.001	0.999
2024-06-14_18-50-42_dqn_cartpole	0.01497353923933631	0.005	0.9993
2024-06-13_19-11-27_dqn_cartpole	0.04971239399803625	0.001	0.999
2024-06-14_13-15-56_dqn_cartpole	0.018279019827489446	0.005	0.999
2024-06-14_15-38-10_dqn_cartpole	0.018279019827489446	0.005	0.999
2024-06-13_18-56-41_dqn_cartpole	0.04971239399803625	0.001	0.999
2024-06-14_11-09-14_dqn_cartpole	0.018279019827489446	0.005	0.999
2024-06-13_19-37-45_dqn_cartpole	0.00247132211701599	0.001	0.999
2024-06-13_21-45-52_dqn_cartpole	0.006721111959865607	0.005	0.999
2024-06-13_22-34-12_dqn_cartpole	0.006721111959865607	0.005	0.999
2024-06-13_20-39-12_dqn_cartpole	0.006721111959865607	0.005	0.999
2024-06-14_19-20-40_dqn_cartpole	0.004998338275642187	0.005	0.999
2024-06-14_18-05-51_dqn_cartpole	0.04974972962361759	0.005	0.9995
2024-06-14_13-24-09_dqn_cartpole	0.018279019827489446	0.005	0.999
2024-06-13_23-49-30_dqn_cartpole	0.018279019827489446	0.005	0.999
2024-06-14_17-21-25_dqn_cartpole	0.004998338275642187	0.005	0.999
2024-06-13_22-09-48_dqn_cartpole	0.00247132211701599	0.001	0.999
2024-06-13_19-16-11_dqn_cartpole	0.22304647413401948	0.001	0.9995

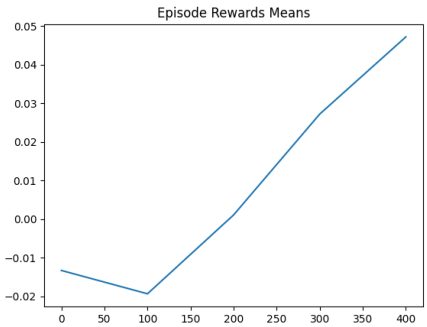
model	batch_size	discount	update_target_every
2024-06-13_17-34-57_dqn_cartpole	32	0.9	20
2024-06-14_00-20-45_dqn_cartpole	128	0.9	40
2024-06-13_21-08-24_dqn_cartpole	32	0.9	20
2024-06-14_09-34-50_dqn_cartpole	128	0.9	30
2024-06-13_18-59-17_dqn_cartpole	32	0.9	20
2024-06-14_17-31-42_dqn_cartpole	128	0.95	40
2024-06-14_18-54-27_dqn_cartpole	128	0.95	40
2024-06-14_18-17-10_dqn_cartpole	128	0.95	40
2024-06-13_18-42-33_dqn_cartpole	32	0.9	20
2024-06-13_23-08-40_dqn_cartpole	128	0.9	40
2024-06-14_16-42-26_dqn_cartpole	128	0.95	40
2024-06-15_00-29-02_dqn_cartpole	128	0.95	40
2024-06-13_18-50-05_dqn_cartpole	32	0.9	20
2024-06-14_18-50-42_dqn_cartpole	128	0.95	40
2024-06-13_19-11-27_dqn_cartpole	32	0.9	20
2024-06-14_13-15-56_dqn_cartpole	128	0.99	30
2024-06-14_15-38-10_dqn_cartpole	128	0.8	30
2024-06-13_18-56-41_dqn_cartpole	32	0.9	20
2024-06-14_11-09-14_dqn_cartpole	128	0.99	10
2024-06-13_19-37-45_dqn_cartpole	32	0.9	20
2024-06-13_21-45-52_dqn_cartpole	128	0.9	20
2024-06-13_22-34-12_dqn_cartpole	128	0.9	20
2024-06-13_20-39-12_dqn_cartpole	32	0.9	20
2024-06-14_19-20-40_dqn_cartpole	128	0.95	40
2024-06-14_18-05-51_dqn_cartpole	128	0.95	40
2024-06-14_13-24-09_dqn_cartpole	128	0.99	40
2024-06-13_23-49-30_dqn_cartpole	128	0.9	10
2024-06-14_17-21-25_dqn_cartpole	128	0.8	30
2024-06-13_22-09-48_dqn_cartpole	32	0.9	20
2024-06-13_19-16-11_dqn_cartpole	32	0.9	20

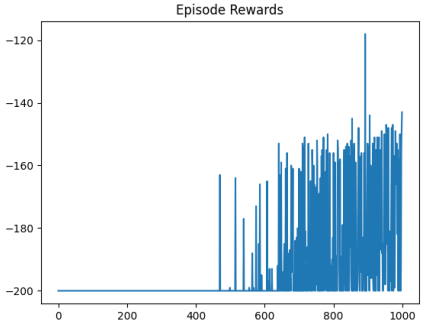
4 Discusión

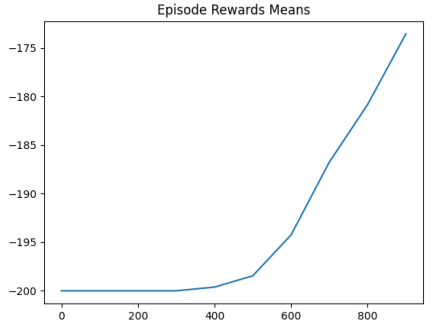
Se entrenaron múltiples modelos variando los hiperparámetros como la cantidad de epochs, la cantidad de capas y la cantidad de neuronas en cada capa. La decisión del mejor modelo se tomó a partir del valor mas alto en la recompensa de evaluación.

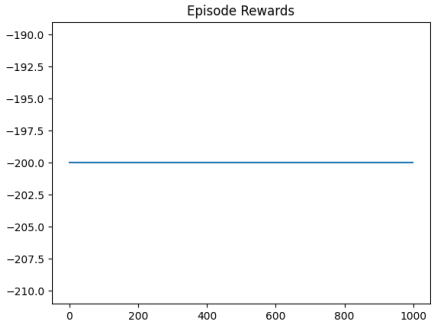
4.1 Mountain

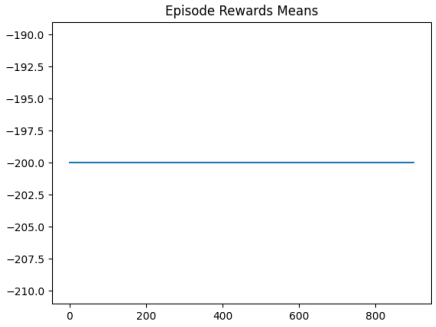

Episode Rewards
0.3
0.2
0.1
0.0
−0.1
−0.2
−0.3
−0.4
0
100
200
300
400
500

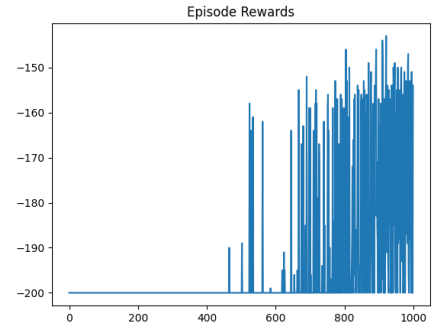

Episode Rewards Means
0.05
0.04
0.03
0.02
0.01
0.00
−0.01
−0.02
0
50
100
150
200
250
300
350
400

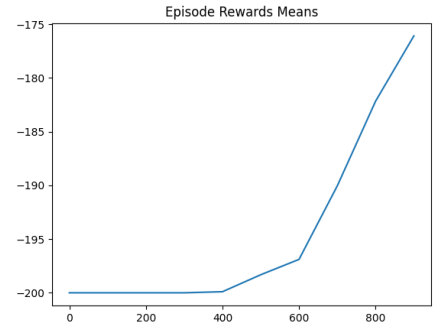

Episode Rewards
−120
−140
−160
−180
−200
0
200
400
600
800
1000

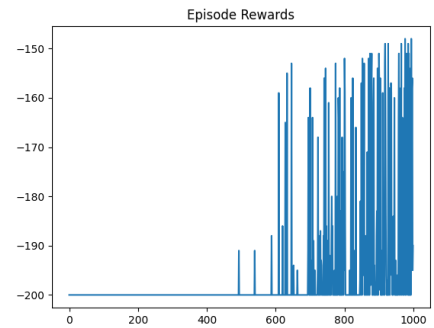

Episode Rewards Means
−175
−180
−185
−190
−195
−200
0
200
400
600
800

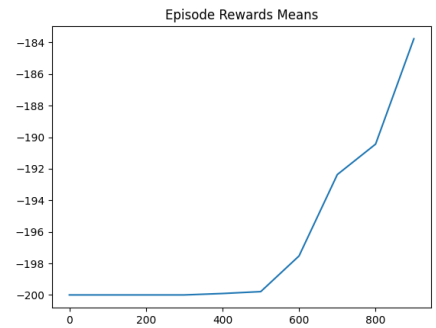

Episode Rewards
−190.0
−192.5
−195.0
−197.5
−200.0
−202.5
−205.0
−207.5
−210.0
0
200
400
600
800
1000

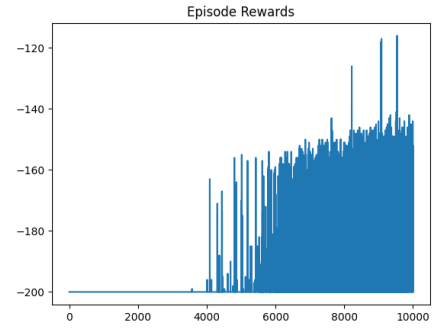

Episode Rewards Means
−190.0
−192.5
−195.0
−197.5
−200.0
−202.5
−205.0
−207.5
−210.0
0
200
400
600
800

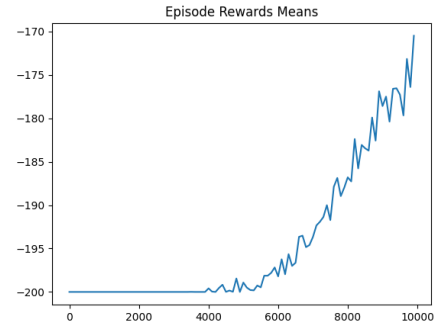

Episode Rewards
−150
−160
−170
−180
−190
−200
0
200
400
600
800
1000

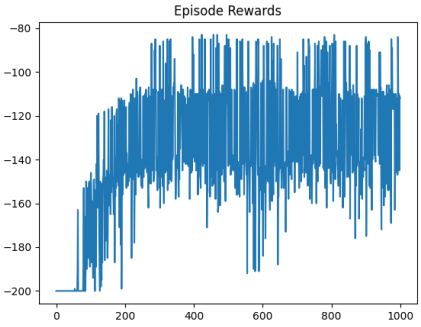

Episode Rewards Means
−175
−180
−185
−190
−195
−200
0
200
400
600
800

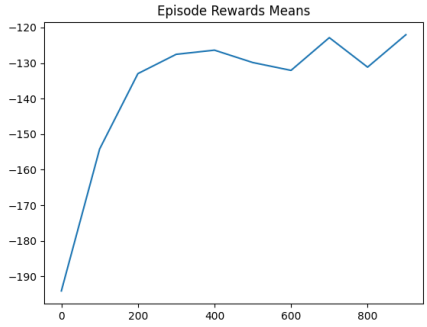

Episode Rewards
−150
−160
−170
−180
−190
−200
0
200
400
600
800
1000

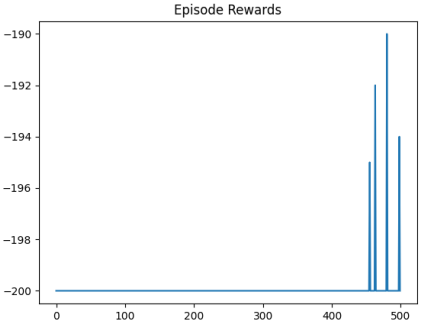

Episode Rewards Means
−184
−186
−188
−190
−192
−194
−196
−198
−200
0
200
400
600
800

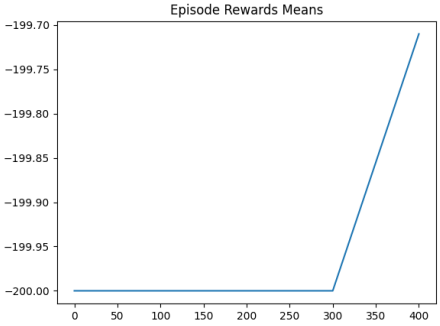

Episode Rewards
−120
−140
−160
−180
−200
0
2000
4000
6000
8000
10000

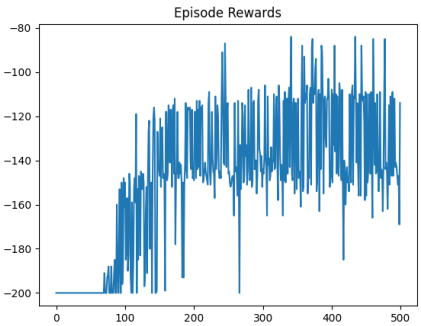

Episode Rewards Means
−170
−175
−180
−185
−190
−195
−200
0
2000
4000
6000
8000
10000

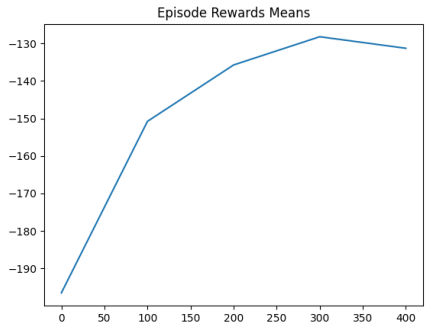

Episode Rewards
−80
−100
−120
−140
−160
−180
−200
0
200
400
600
800
1000

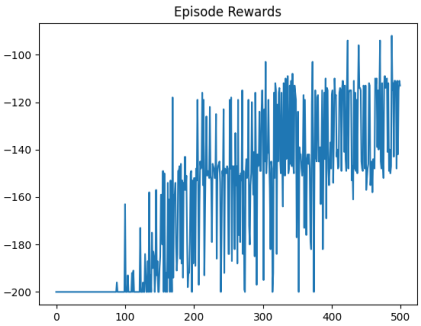

Episode Rewards Means
−120
−130
−140
−150
−160
−170
−180
−190
0
200
400
600
800

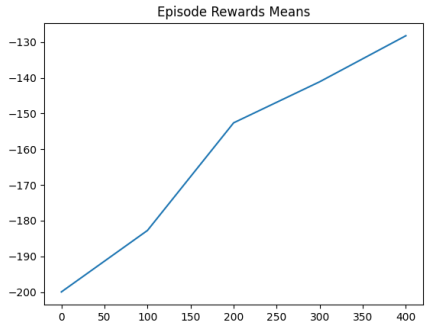

Episode Rewards
−190
−192
−194
−196
−198
−200
0
100
200
300
400
500

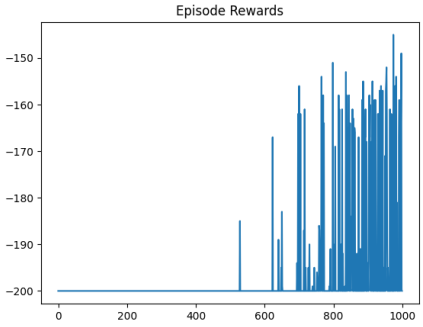

Episode Rewards Means
−199.70
−199.75
−199.80
−199.85
−199.90
−199.95
−200.00
0
50
100
150
200
250
300
350
400

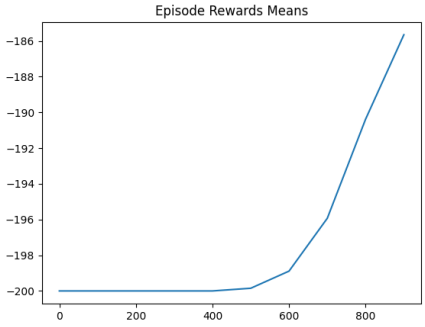

Episode Rewards
−80
−100
−120
−140
−160
−180
−200
0
100
200
300
400
500

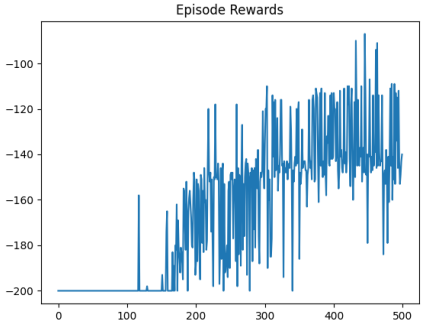

Episode Rewards Means
−130
−140
−150
−160
−170
−180
−190
0
50
100
150
200
250
300
350
400

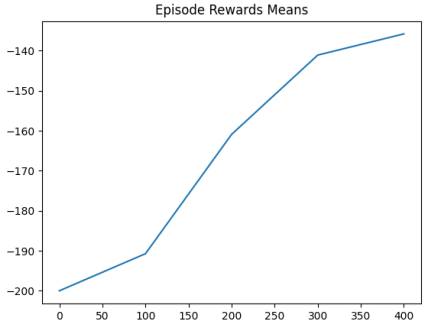

Episode Rewards
−100
−120
−140
−160
−180
−200
0
100
200
300
400
500

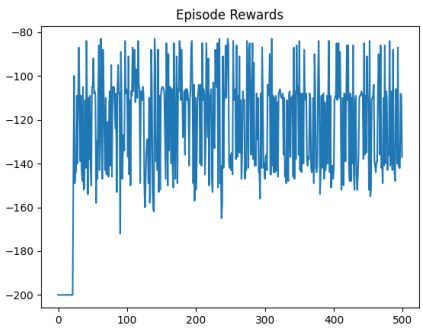

Episode Rewards Means
−130
−140
−150
−160
−170
−180
−190
−200
0
50
100
150
200
250
300
350
400


Episode Rewards
−150
−160
−170
−180
−190
−200
0
200
400
600
800
1000

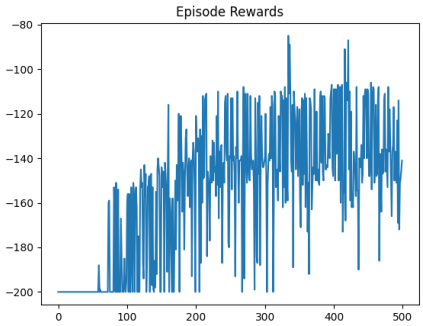

Episode Rewards Means
−186
−188
−190
−192
−194
−196
−198
−200
0
200
400
600
800


Episode Rewards
−100
−120
−140
−160
−180
−200
0
100
200
300
400
500

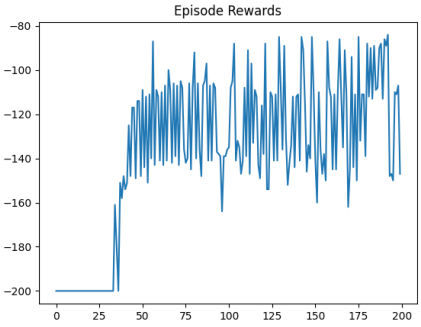

Episode Rewards Means
−140
−150
−160
−170
−180
−190
−200
0
50
100
150
200
250
300
350
400

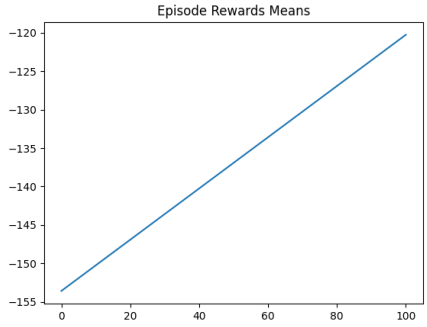

Episode Rewards
−80
−100
−120
−140
−160
−180
−200
0
100
200
300
400
500

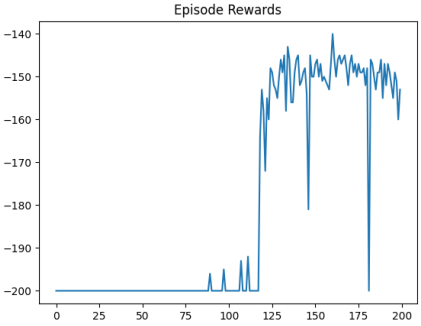

Episode Rewards Means
−120
−125
−130
−135
0
50
100
150
200
250
300
350
400

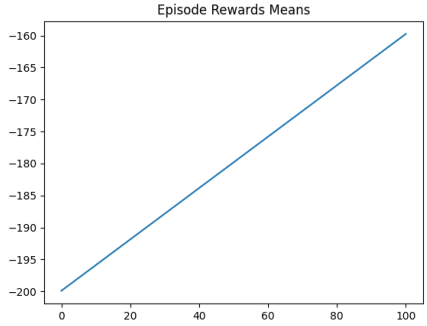

Episode Rewards
−80
−100
−120
−140
−160
−180
−200
0
100
200
300
400
500

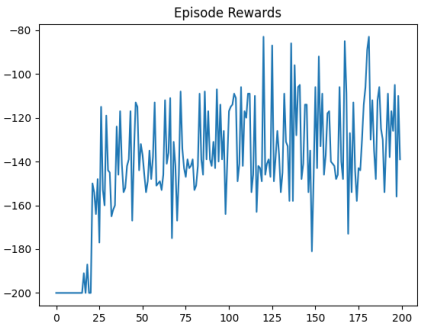

Episode Rewards Means
−140
−150
−160
−170
−180
−190
0
50
100
150
200
250
300
350
400

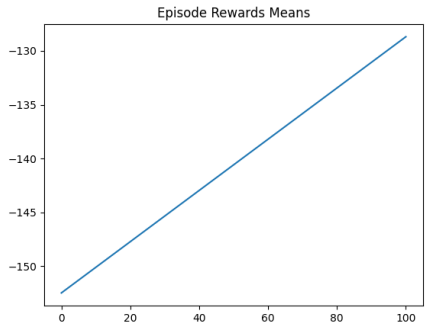

Episode Rewards
−80
−100
−120
−140
−160
−180
−200
0
25
50
75
100
125
150
175
200

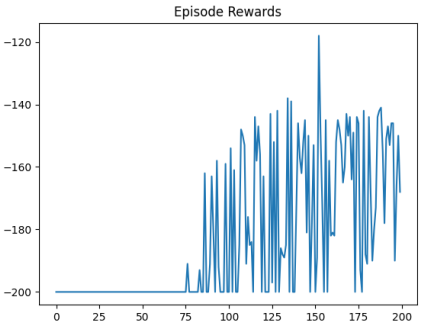

Episode Rewards Means
−120
−125
−130
−135
−140
−145
−150
−155
0
20
40
60
80
100

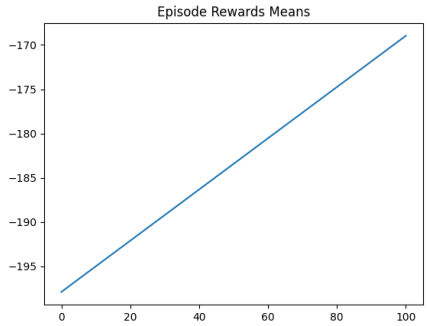

Episode Rewards
−140
−150
−160
−170
−180
−190
−200
0
25
50
75
100
125
150
175
200

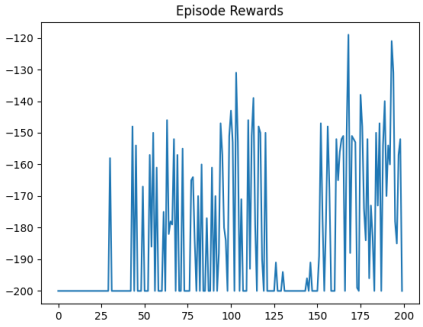

Episode Rewards Means
−160
−165
−170
−175
−180
−185
−190
−195
−200
0
20
40
60
80
100

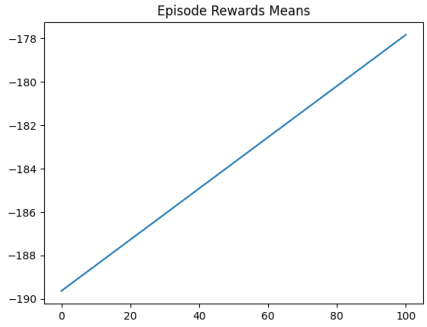

Episode Rewards
−80
−100
−120
−140
−160
−180
−200
0
25
50
75
100
125
150
175
200

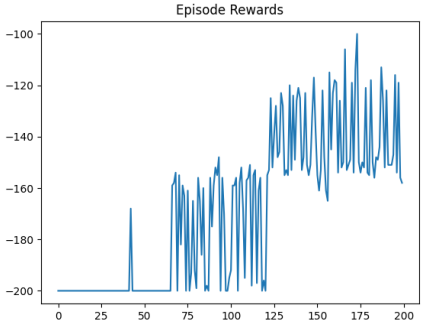

Episode Rewards Means
−130
−135
−140
−145
−150
0
20
40
60
80
100

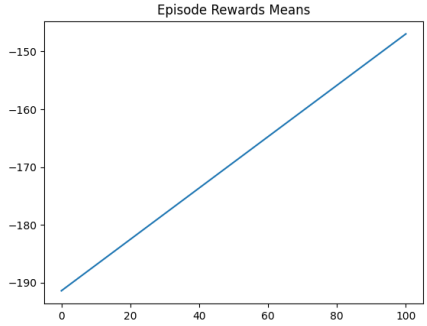

Episode Rewards
−120
−140
−160
−180
−200
0
25
50
75
100
125
150
175
200

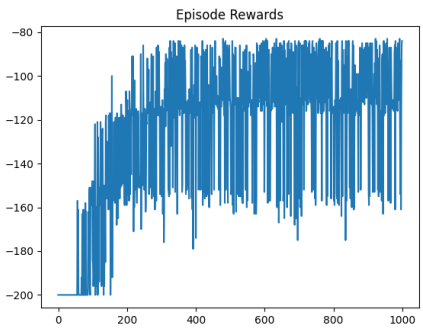

Episode Rewards Means
−170
−175
−180
−185
−190
−195
0
20
40
60
80
100

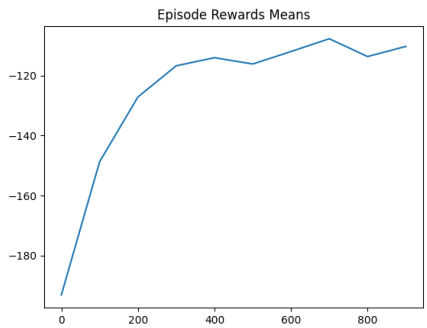

Episode Rewards
−120
−130
−140
−150
−160
−170
−180
−190
−200
0
25
50
75
100
125
150
175
200

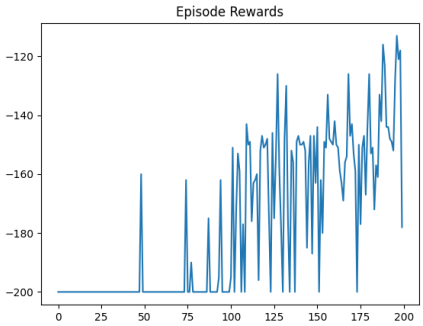

Episode Rewards Means
−178
−180
−182
−184
−186
−188
−190
0
20
40
60
80
100

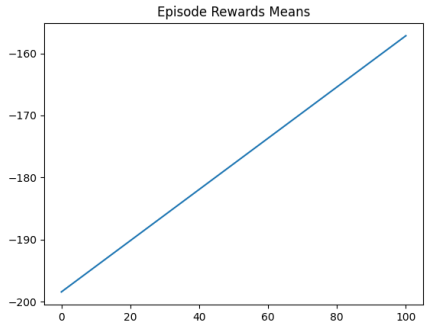

Episode Rewards
−100
−120
−140
−160
−180
−200
0
25
50
75
100
125
150
175
200


Episode Rewards Means
−150
−160
−170
−180
−190
0
20
40
60
80
100

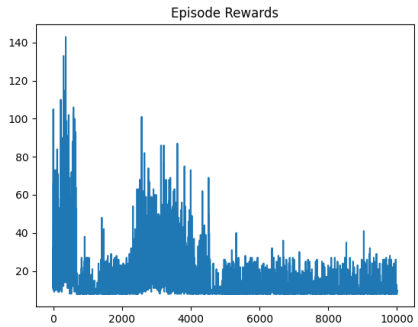

Episode Rewards
−80
−100
−120
−140
−160
−180
−200
0
200
400
600
800
1000

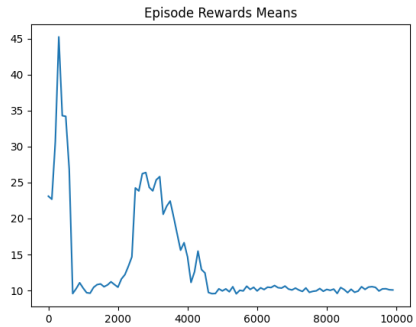

Episode Rewards Means
−120
−140
−160
−180
0
200
400
600
800

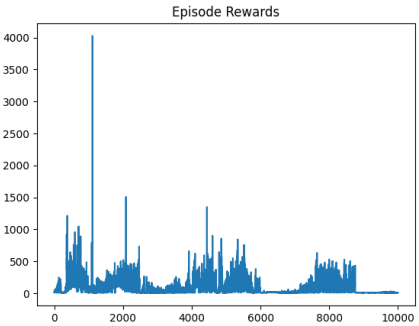

Episode Rewards
−120
−140
−160
−180
−200
0
25
50
75
100
125
150
175
200

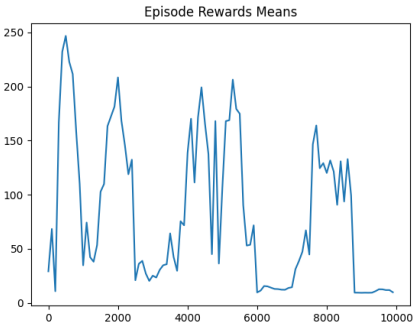

Episode Rewards Means
−160
−170
−180
−190
−200
0
20
40
60
80
100

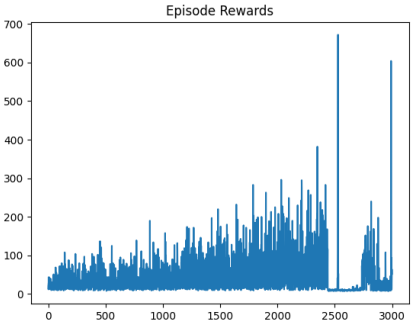
4.2 CartPole


Episode Rewards
140
120
100
80
60
40
20
0
2000
4000
6000
8000
10000

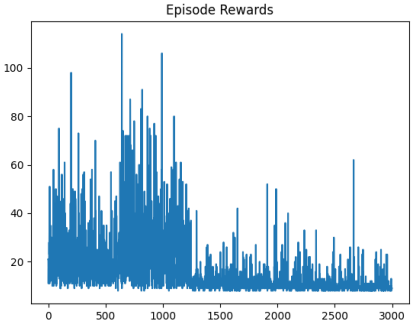

Episode Rewards Means
45
40
35
30
25
20
15
10
0
2000
4000
6000
8000
10000


Episode Rewards
4000
3500
3000
2500
2000
1500
1000
500
0
0
2000
4000
6000
8000
10000

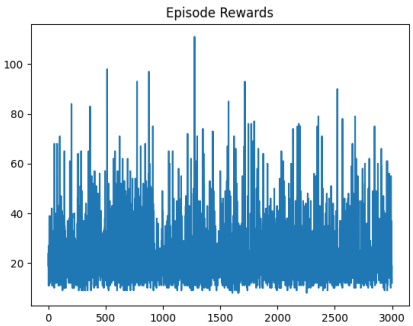

Episode Rewards Means
250
200
150
100
50
0
0
2000
4000
6000
8000
10000


Episode Rewards
0
100
200
300
400
500
600
700
0
500
1000
1500
2000
2500
3000

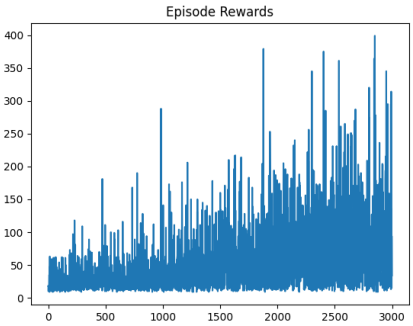

Episode Rewards Means
10
20
30
40
50
60
70
80
0
500
1000
1500
2000
2500
3000

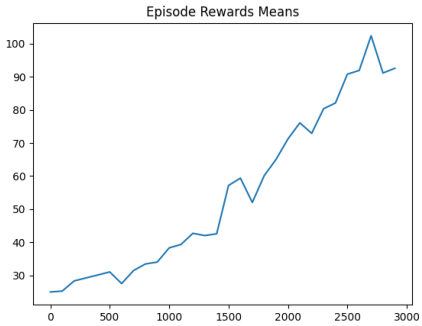

Episode Rewards
20
40
60
80
100
0
500
1000
1500
2000
2500
3000

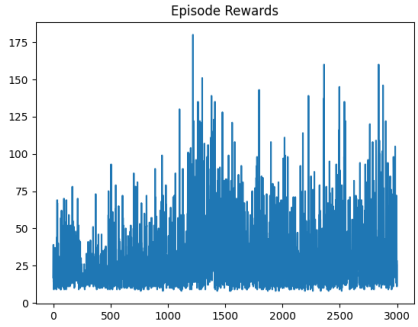

Episode Rewards Means
10.0
12.5
15.0
17.5
20.0
22.5
25.0
27.5
30.0
0
500
1000
1500
2000
2500
3000

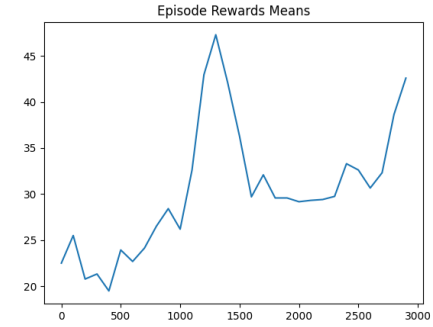

Episode Rewards
20
40
60
80
100
0
500
1000
1500
2000
2500
3000

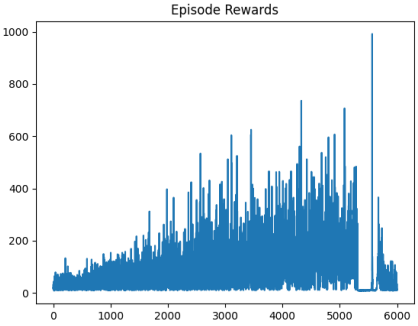

Episode Rewards Means
21
22
23
24
25
26
27
0
500
1000
1500
2000
2500
3000

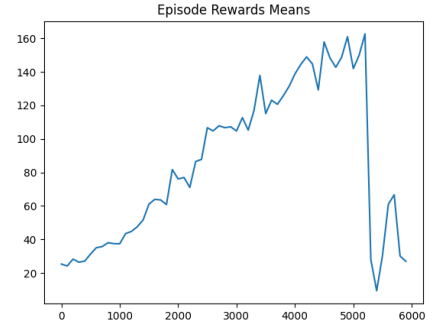

Episode Rewards
0
50
100
150
200
250
300
350
400
0
500
1000
1500
2000
2500
3000

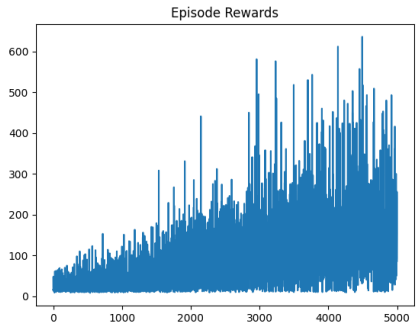

Episode Rewards Means
30
40
50
60
70
80
90
100
0
500
1000
1500
2000
2500
3000

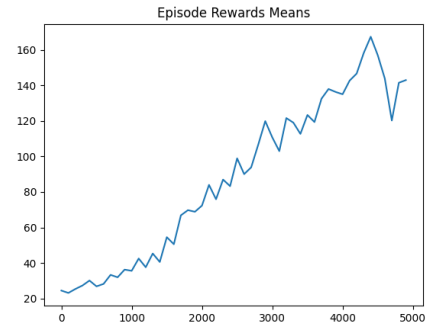

Episode Rewards
0
25
50
75
100
125
150
175
0
500
1000
1500
2000
2500
3000

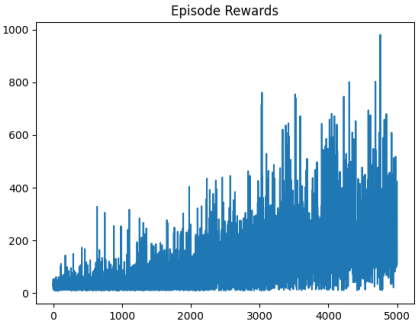

Episode Rewards Means
20
25
30
35
40
45
0
500
1000
1500
2000
2500
3000

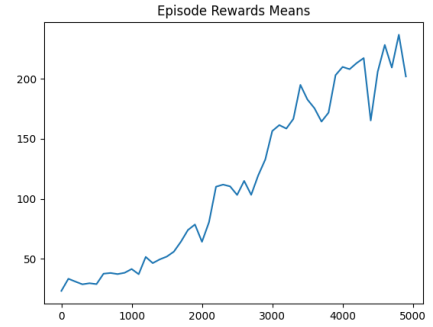

Episode Rewards
0
200
400
600
800
1000
0
1000
2000
3000
4000
5000
6000

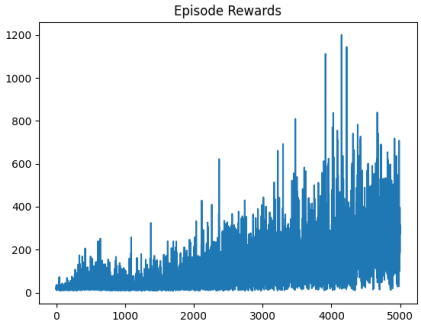

Episode Rewards Means
20
40
60
80
100
120
140
160
0
1000
2000
3000
4000
5000
6000

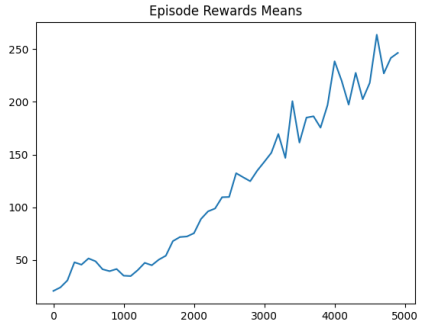

Episode Rewards
0
100
200
300
400
500
600
0
1000
2000
3000
4000
5000

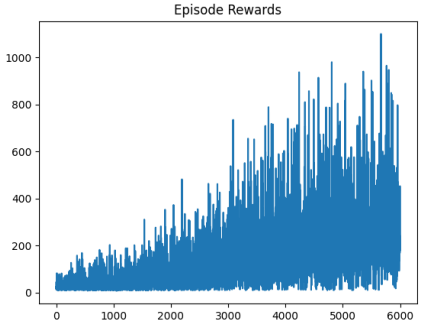

Episode Rewards Means
20
40
60
80
100
120
140
160
0
1000
2000
3000
4000
5000

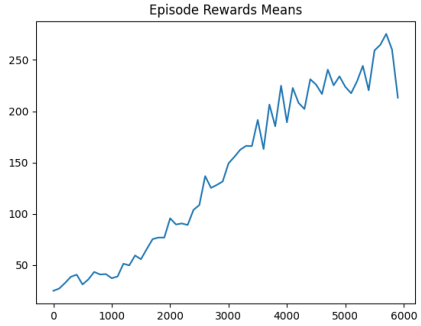

Episode Rewards
0
200
400
600
800
1000
0
1000
2000
3000
4000
5000

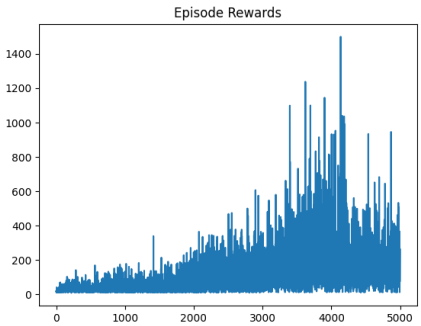

Episode Rewards Means
50
100
150
200
0
1000
2000
3000
4000
5000

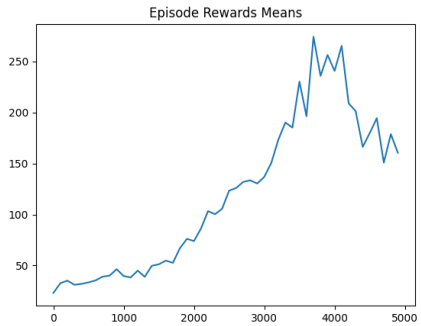

Episode Rewards
0
200
400
600
800
1000
1200
0
1000
2000
3000
4000
5000

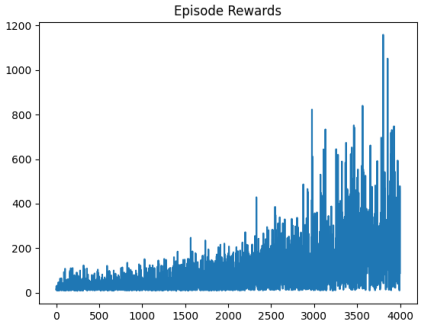

Episode Rewards Means
50
100
150
200
250
0
1000
2000
3000
4000
5000

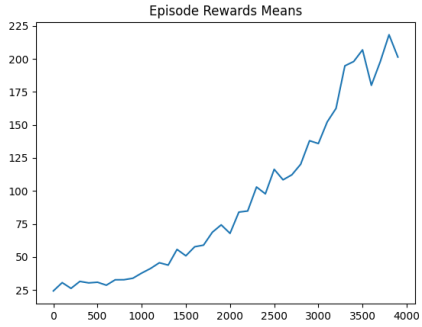

Episode Rewards
0
200
400
600
800
1000
0
1000
2000
3000
4000
5000
6000

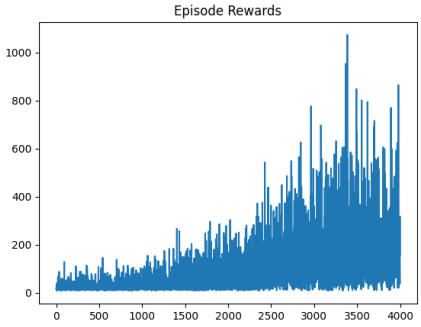

Episode Rewards Means
50
100
150
200
250
0
1000
2000
3000
4000
5000
6000

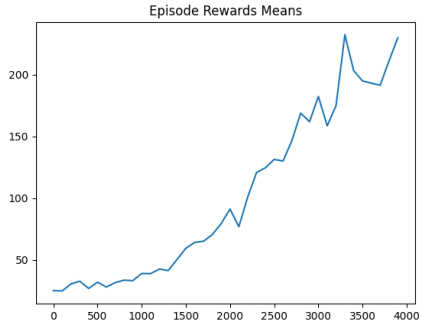

Episode Rewards
0
200
400
600
800
1000
1200
1400
0
1000
2000
3000
4000
5000

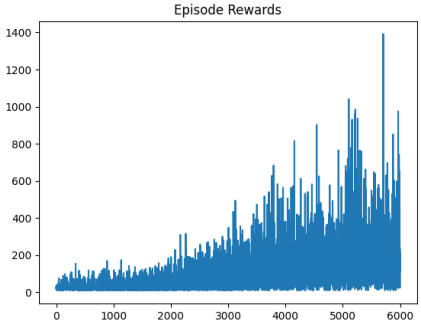

Episode Rewards Means
50
100
150
200
250
0
1000
2000
3000
4000
5000

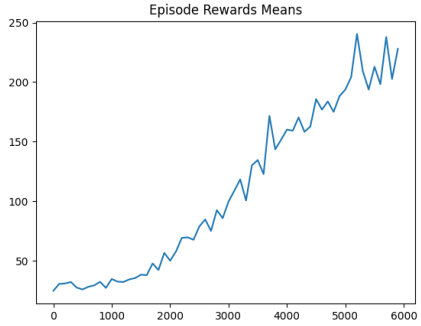

Episode Rewards
0
200
400
600
800
1000
1200
0
500
1000
1500
2000
2500
3000
3500
4000

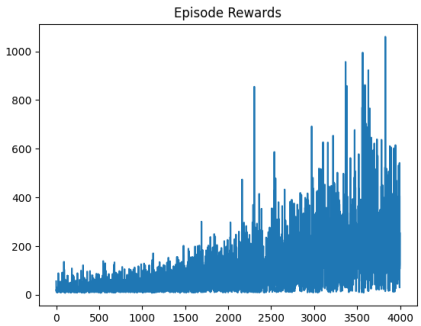

Episode Rewards Means
25
50
75
100
125
150
175
200
225
0
500
1000
1500
2000
2500
3000
3500
4000

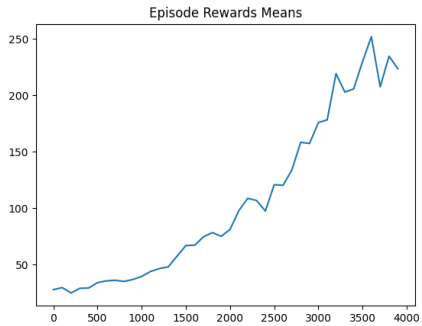

Episode Rewards
0
200
400
600
800
1000
0
500
1000
1500
2000
2500
3000
3500
4000

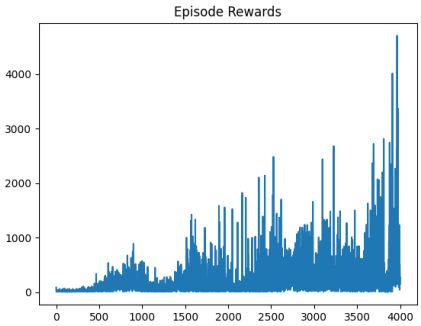

Episode Rewards Means
50
100
150
200
0
500
1000
1500
2000
2500
3000
3500
4000

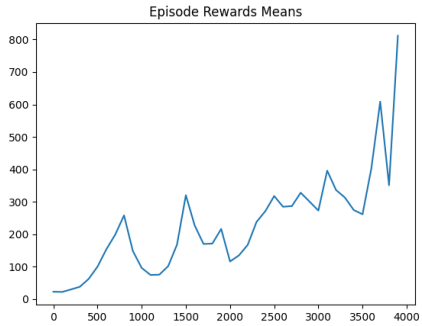

Episode Rewards
0
200
400
600
800
1000
1200
1400
0
1000
2000
3000
4000
5000
6000

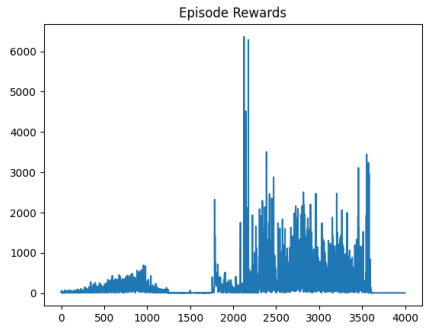

Episode Rewards Means
50
100
150
200
250
0
1000
2000
3000
4000
5000
6000

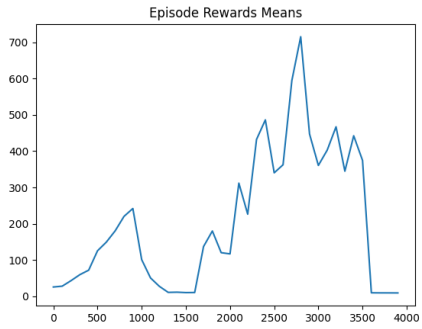

Episode Rewards
0
200
400
600
800
1000
0
500
1000
1500
2000
2500
3000
3500
4000

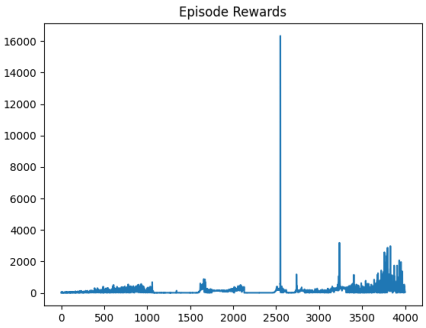

Episode Rewards Means
50
100
150
200
250
0
500
1000
1500
2000
2500
3000
3500
4000

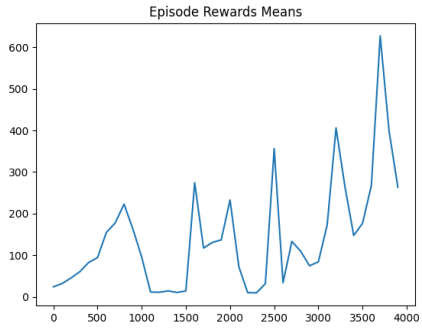

Episode Rewards
0
1000
2000
3000
4000
0
500
1000
1500
2000
2500
3000
3500
4000

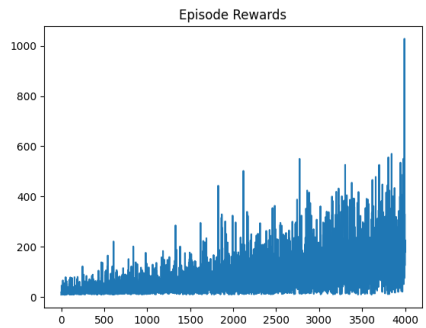

Episode Rewards Means
0
100
200
300
400
500
600
700
800
0
500
1000
1500
2000
2500
3000
3500
4000

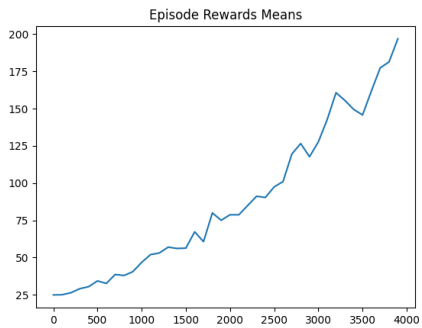

Episode Rewards
0
1000
2000
3000
4000
5000
6000
0
500
1000
1500
2000
2500
3000
3500
4000

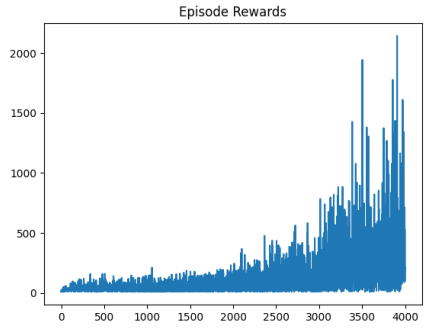

Episode Rewards Means
0
100
200
300
400
500
600
700
0
500
1000
1500
2000
2500
3000
3500
4000

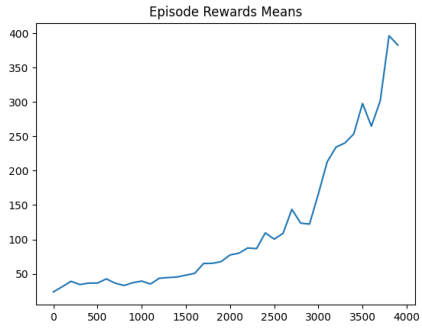

Episode Rewards
0
2000
4000
6000
8000
10000
12000
14000
16000
0
500
1000
1500
2000
2500
3000
3500
4000

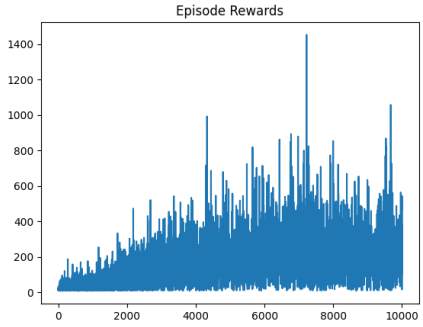

Episode Rewards Means
0
100
200
300
400
500
600
0
500
1000
1500
2000
2500
3000
3500
4000

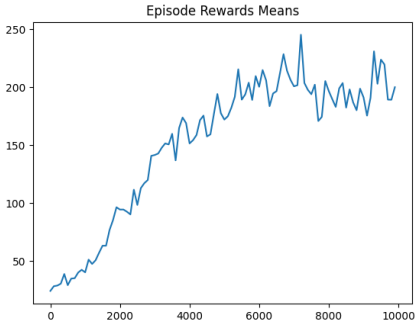

Episode Rewards
0
200
400
600
800
1000
0
500
1000
1500
2000
2500
3000
3500
4000

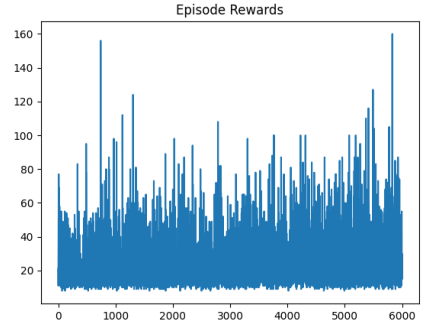

Episode Rewards Means
25
50
75
100
125
150
175
200
0
500
1000
1500
2000
2500
3000
3500
4000

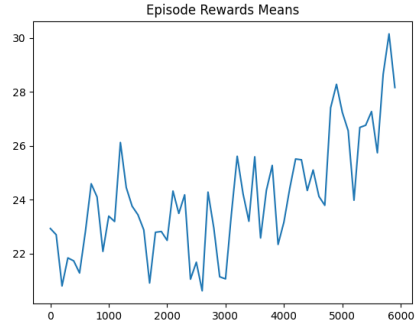

Episode Rewards
0
500
1000
1500
2000
0
500
1000
1500
2000
2500
3000
3500
4000

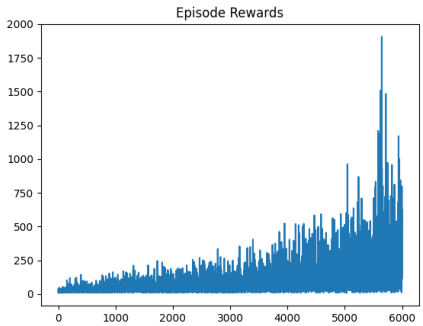

Episode Rewards Means
50
100
150
200
250
300
350
400
0
500
1000
1500
2000
2500
3000
3500
4000

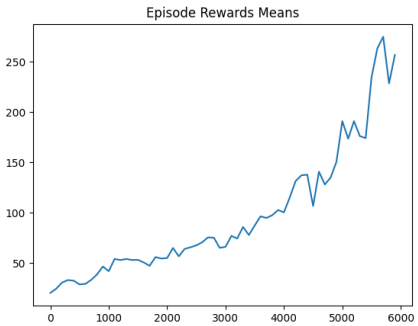

Episode Rewards
0
200
400
600
800
1000
1200
1400
0
2000
4000
6000
8000
10000

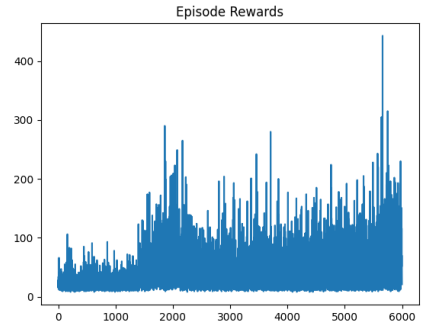

Episode Rewards Means
50
100
150
200
250
0
2000
4000
6000
8000
10000

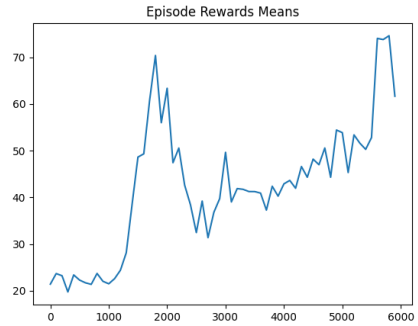

Episode Rewards
20
40
60
80
100
120
140
160
0
1000
2000
3000
4000
5000
6000

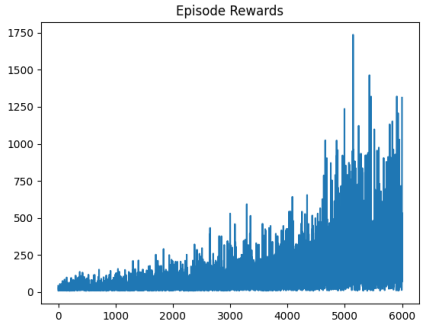

Episode Rewards Means
22
24
26
28
30
0
1000
2000
3000
4000
5000
6000

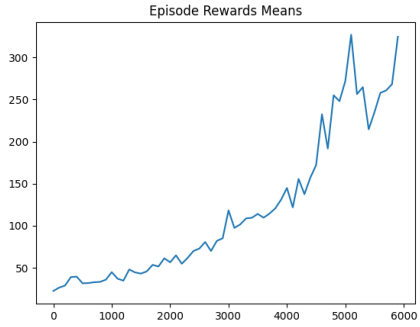

Episode Rewards
0
250
500
750
1000
1250
1500
1750
2000
0
1000
2000
3000
4000
5000
6000

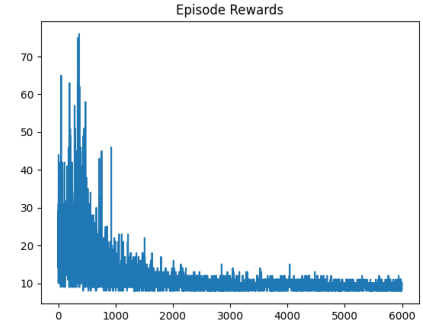

Episode Rewards Means
50
100
150
200
250
0
1000
2000
3000
4000
5000
6000

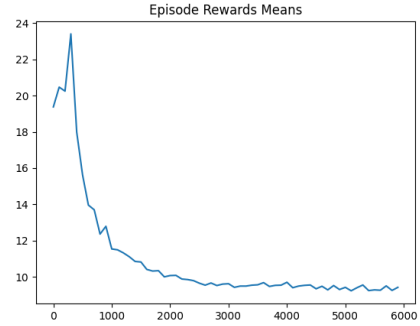

Episode Rewards
0
100
200
300
400
0
1000
2000
3000
4000
5000
6000

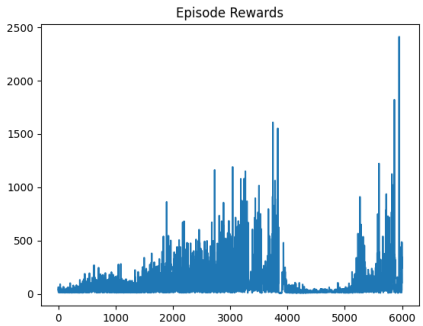

Episode Rewards Means
20
30
40
50
60
70
0
1000
2000
3000
4000
5000
6000

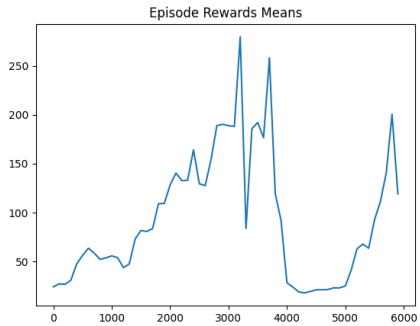

Episode Rewards
0
250
500
750
1000
1250
1500
1750
0
1000
2000
3000
4000
5000
6000

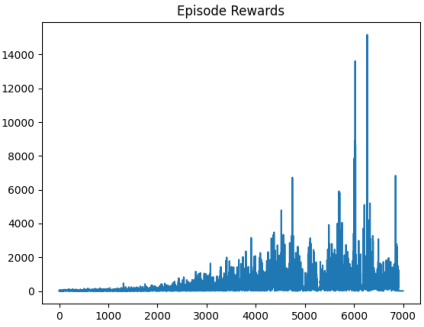

Episode Rewards Means
50
100
150
200
250
300
0
1000
2000
3000
4000
5000
6000

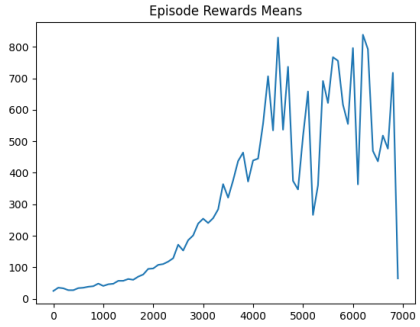

Episode Rewards
10
20
30
40
50
60
70
0
1000
2000
3000
4000
5000
6000


Episode Rewards Means
10
12
14
16
18
20
22
24
0
1000
2000
3000
4000
5000
6000


Episode Rewards
0
500
1000
1500
2000
2500
0
1000
2000
3000
4000
5000
6000


Episode Rewards Means
50
100
150
200
250
0
1000
2000
3000
4000
5000
6000


Episode Rewards
0
2000
4000
6000
8000
10000
12000
14000
0
1000
2000
3000
4000
5000
6000
7000


Episode Rewards Means
0
100
200
300
400
500
600
700
800
0
1000
2000
3000
4000
5000
6000
7000

5 Conclusión

Concluimos que las redes neuronales pueden ser una herramienta eficaz para aproximar la solución de la ecuación de movimiento de un péndulo simple. Sin embargo, es importante considerar la complejidad del modelo y la calidad de los datos de entrada para obtener resultados precisos.