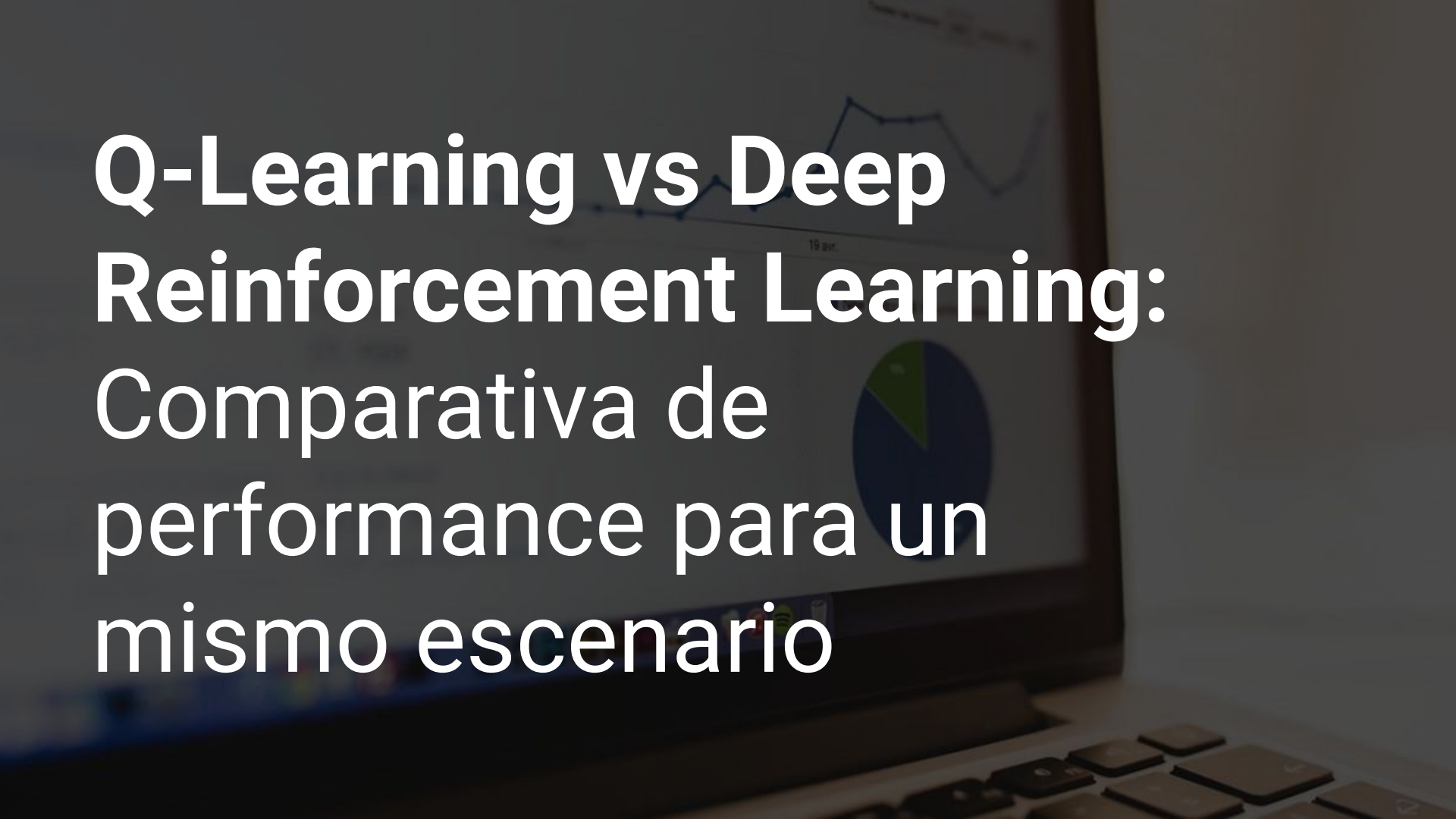


Inteligencia Artificial I

Cavallo Gaston
Fernandez Juan Manuel

Q-Learning vs Deep Reinforcement Learning: Comparativa de performance para un mismo escenario

The background of the image is a dark, slightly blurred photograph of a laptop. The laptop screen is visible, showing a line graph with a blue line and a pie chart with a green slice. The text is overlaid on the left side of the screen in a large, white, sans-serif font.

Problema

Probamos los algoritmos de Deep Reinforcement Learning y QLearning y los comparamos según distintas métricas para verificar la eficiencia que presentaban en un mismo entorno

En este caso, ambos agentes jugaron al snake 600 episodios



A close-up photograph of a person's hand playing a piano. The hand is positioned over the keys, with fingers slightly curved. The background is blurred, showing some indistinct shapes and colors, possibly a room with a lamp. The overall tone is artistic and focused on the action of playing music.

Q-learning

¿Que es?

¿ Qué hace ?

Q-Learning es un algoritmo que utiliza un conjunto de funciones para la exploración del entorno, luego, lo “aprende” o “memoriza” utilizando una matriz denominada Q-Table.

¿Cómo lo hace?

La exploración en Q-Learning apunta a maximizar la recompensa obtenida a cada acción que toma, para lograrlo, se utiliza la Ecuación de Optimalidad de Bellman

$$q_*(s, a) = E \left[R_{t+1} + \gamma \max_{a'} q_*(s', a') \right]$$



Deep Reinforcement Learning

¿ Qué hace ?

Se basa principalmente en las denominadas redes neuronales, las cuales están compuestas por una serie de “capas”, encargadas de procesar la información de manera tal que cada capa realiza una serie de operaciones sobre los datos y luego los propaga a la siguiente capa, para así repetir el proceso y llegar a una salida.

¿Cómo lo hace?

La idea del Deep Learning es la de lograr una representación más “simple” de los datos, realizando operaciones sobre la entrada y ajustando valores internos dentro de cada una de las capas, estos valores internos se denominan pesos o variables ocultas, los cuales resultan del ajuste iterativo haciendo uso de una función de optimización y otra de “pérdida”.

Pasos en la implementación

- Representación del entorno (entrada)
 - Optimización del agente (procesamiento)
 - Obtención de resultados (comparativas)

An aerial photograph of the New York City skyline at dusk. The sky is a mix of dark blue and orange, with scattered clouds. The city is densely packed with skyscrapers, many of which are illuminated with their interior lights. The Empire State Building is prominent in the center, with its top lit in red and green. The Hudson River is visible in the background, and the city lights reflect on the water. The title 'Métricas y comparativas' is overlaid in a large, white, sans-serif font on the left side of the image.

Métricas y comparativas

Tiempo transcurrido en segundos hasta obtener una recompensa media de 7



Cantidad de episodios hasta la recompensa media de 7



Máximas puntuaciones



Tiempo total de ejecución

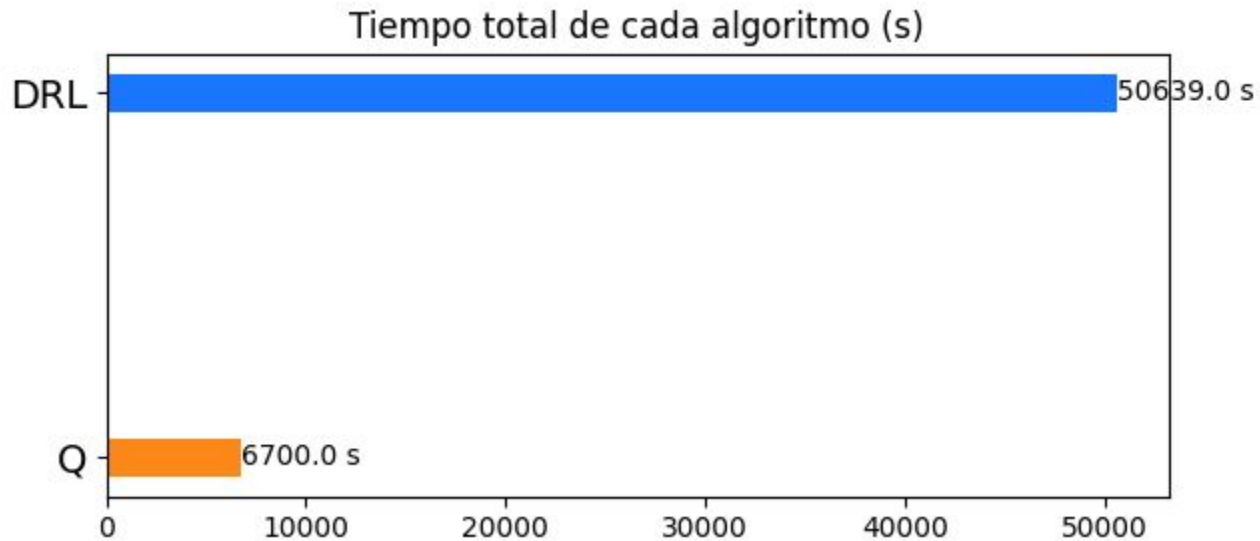
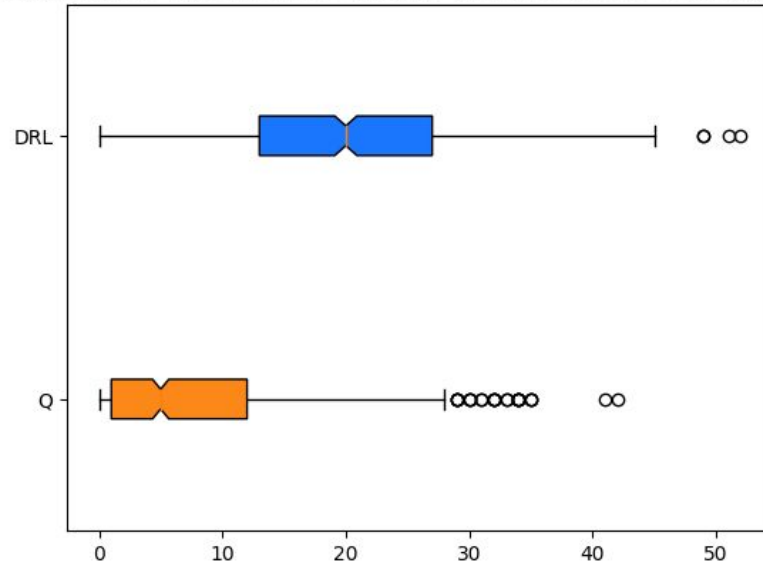


Diagrama de caja y bigotes para ambos algoritmos

Diagrama de cajas y bigotes para las puntuaciones de cada algoritmo



Conclusiones

- Tiempo que le toma a cada agente lograr una política óptima del entorno: DRL fue 17 veces más rápido que Q-Learning.
- Ritmo de aprendizaje del agente: DRL fue 39 veces más rápido
- Ambos agentes logran resolver el problema de manera eficaz: hay una diferencia del 20% de puntuación máxima entre ambos agentes



Opiniones y puntos de vista