

# Reinforcement Learning: Cómo los Agentes Autónomos Aprenden a Decidir

## Introducción

### ¿Qué es?

Es una técnica de Machine Learning donde un agente aprende a actuar en un entorno interactuando con él.

Aprende por prueba y error, recibiendo recompensas o penalizaciones según sus acciones, con el objetivo de maximizar su recompensa acumulada.



Learning	Training	Tags	Goal
Supervisado	Ejemplos y etiquetas	✓	Clasificación o Regresión
No supervisado	Sin etiquetas	✗	Clustering o Reducir dimensiones
Por refuerzo	Interacción con el entorno	✗	Maximizar recompensa

### Elementos Básicos

- Agente:** Toma decisiones.
- Entorno:** Donde actúa el agente.
- Estado:** Situación actual del agente.
- Acción:** Opción que elige el agente.
- Recompensa:** Valor numérico por una acción.
- Política:** Estrategia de acción del agente.
- Función de valor:** Evalúa la calidad de una acción o estado.

## Tipos de Algoritmos

- **Q-learning:** Aprende valores de acción sin modelo del entorno.
- **SARSA:** Similar al Q-learning pero usa la acción realmente tomada.
- **Deep Q-Learning (DQN):** Usa redes neuronales para aprender valores.
- **Policy Gradient:** Aprende políticas directamente, útil en acciones continuas.

### Actualización de Q-learning

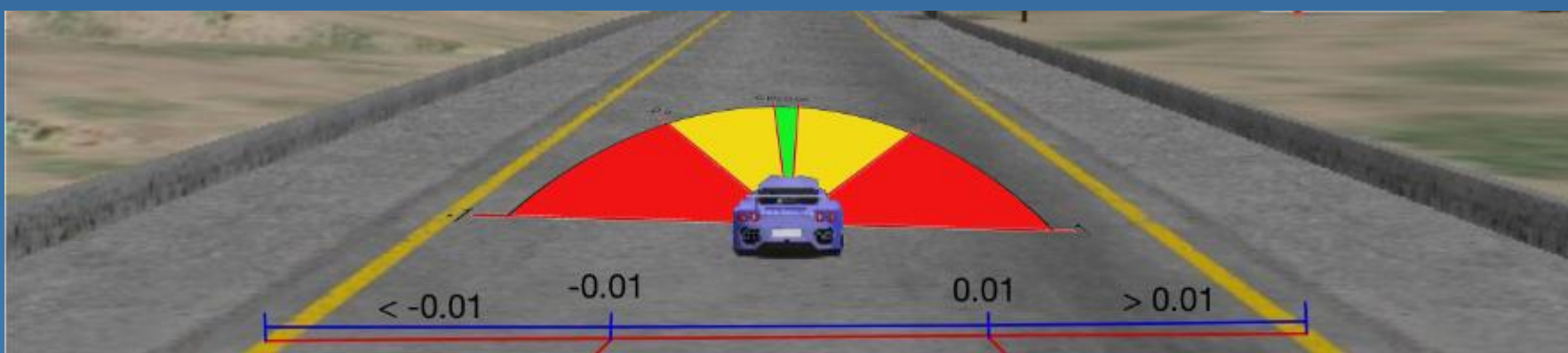
$$Q(s, a) \leftarrow Q(s, a) + \alpha \left[ R + \gamma \max_{a'} Q(s', a') - Q(s, a) \right]$$

### Donde:

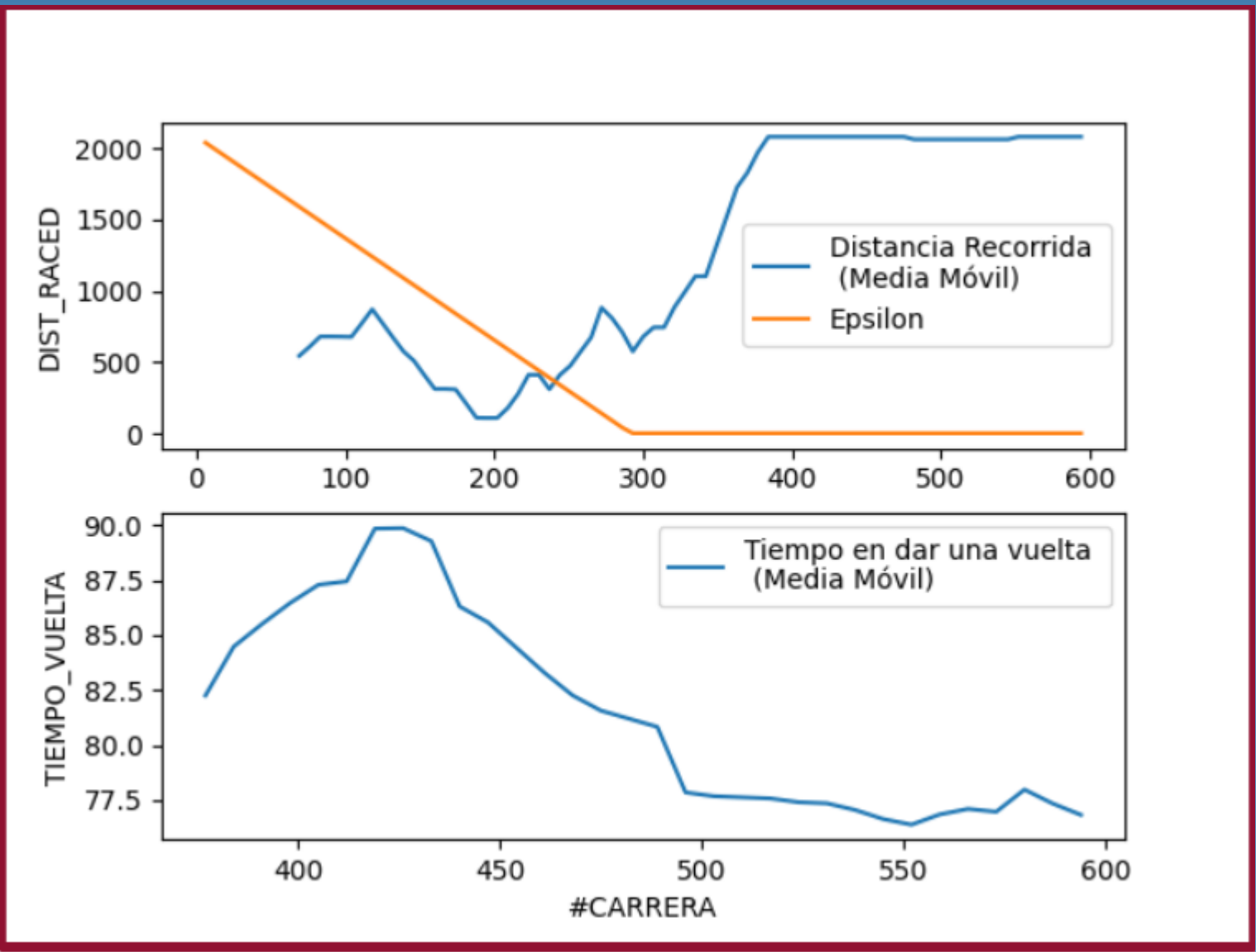
- **Q(s,a):** valor actual estimado para el estado s y acción a
- **R:** recompensa recibida
- **γ:** descuento futuro
- **α:** tasa de aprendizaje

## Aplicaciones

- Robótica (Boston Dynamics)
- Juegos (AlphaGo, ATARI)
- Finanzas y trading algorítmico
- Logística y planificación de rutas
- **Conducción autónoma**
- Sistemas de recomendación (Netflix, YouTube)
- Domótica e IoT inteligente



Caso Práctico. Conducción Autónoma con RL en TORCS.



Gráfica de aprendizaje del agente de conducción autónoma en TORCS.

## Ventajas y Retos

- ✓ Aprende por experiencia y se adapta a entornos complejos.
- ✓ Aplicable donde no se dispone de datos etiquetados.
- ✗ Costoso computacionalmente.
- ✗ Difícil de entrenar en entornos con recompensas escasas.



Un agente de Aprendizaje por Refuerzo entrenado para conducir en el simulador TORCS. Aprende mediante prueba y error a mantenerse en pista, girar y acelerar de forma óptima. Utiliza el algoritmo QLearning.

Trabajo propio: <https://github.com/albertofermer/QLearningTorcs>



UNIVERSIDAD  
DE  
CÓRDOBA

Basado en: Sutton & Barto (2018). Reinforcement Learning: An Introduction.