



## Grado en Inteligencia Artificial Modelos Avanzados de Aprendizaje Automático II

### Práctica 3: Aprendizaje por Refuerzo

Curso 2024/2025

En esta práctica desarrollarás un agente que, mediante un entrenamiento basado en aprendizaje por refuerzo, maneje con destreza el entorno Pendulum-v1 de OpenGym. Este entorno simula un péndulo simple al que se le puede aplicar una fuerza rotacional; el objetivo del agente es mantener el péndulo en posición vertical por encima de su eje.

---

#### • Trabajo a Realizar

- Implementa los algoritmos MonteCarlo control, SARSA y Q-Learning en el entorno Pendulum-v1.
- Realiza los experimentos necesarios y contesta **de forma justificada** a las siguientes preguntas.
  1. ¿Qué hiperparámetros has utilizado? ¿Cómo has seleccionado estos hiperparámetros? ¿Por qué son más adecuados que otros valores?
  2. ¿Qué algoritmo obtiene mejor política determinista?
  3. ¿Qué algoritmo obtiene mejor rendimiento como algoritmo de control (es decir, mayores retornos mientras sigue aprendiendo)?
  4. Si introducimos perturbaciones\*, ¿qué algoritmo funciona mejor? ¿Cómo podríamos adaptar el aprendizaje para que funcione mejor en estos casos?

\* Nota: las perturbaciones consistirán en probabilidades aleatorias con  $p=0.05$  de que la acción ejecutada sea  $[-2.0]$  o  $[2.0]$  en lugar de la acción determinada por el agente.

#### • Consejos

- Utiliza Jupyter si quieres renderizar el entorno (usando `"render_mode='human'"`).
  - Limita la duración de los episodios a 10s.
  - Deberás representar el estado del entorno (a partir de las observaciones) y las acciones posibles, que toman valores continuos; puedes utilizar la representación que veas conveniente, pero se recomienda codificarlos de forma discreta (p.ej. en 20 intervalos).
- 

#### • Entrega

- La entrega se realizará antes del viernes, 9 de mayo, a las 23:59.
- La práctica se realizará por parejas, de forma obligatoria. Sólo uno de los miembros de cada pareja debe entregar la práctica.
- En la entrega deben incluirse:
  - i. Uno o varios notebooks o scripts de Python que contengan los algoritmos implementados y los experimentos realizados
  - ii. Un informe en formato PDF en el que justifiques las decisiones tomadas y contestes de forma exhaustiva a las preguntas del enunciado.

