

# EL7021: Tarea 5

Profesor: Javier Ruiz del Solar  
Auxiliar: Francisco Leiva  
Ayudante: Rodrigo Salas

Junio, 2023

## Requisitos

- Python  $\geq 3.x$
- NumPy
- Matplotlib
- PyTorch
- OpenAI Gym (`pip install gym==0.23.1`)

## Fechas de entrega

Partes I y II: 12 de junio, hasta las 23:59

## Parte I

### Descripción

En esta parte de la tarea implementará una política experta para un agente en el ambiente `CartPole`, y algunas funciones auxiliares para la recolección de datos empleando dicha política.

### Instrucciones

1. Complete la clase `PIDController` del archivo `pid_controller.py`.
2. Ajuste los parámetros del controlador para maximizar la recompensa en el ambiente `CartPole`.
3. Reporte los parámetros empleados, y los resultados obtenidos al evaluar el desempeño del controlador (recompensa promedio, índice de éxito promedio).

## Parte II

### Descripción

En esta parte de la tarea implementará Conservative Q-Learning en conjunción a DQN y probará dicho algoritmo en problemas de control clásicos.

## Instrucciones

1. Complete la función `update` del archivo `conservative_q_learning.py` siguiendo las instrucciones del código base.
2. Experimentos
  - 2.1 Entrene un agente en el ambiente `CartPole` usando la configuración del código base, y las siguientes combinaciones de parámetros:

ID Experimento	nb_rollouts	alpha
exp_11	10	0.0
exp_21	100	0.0
exp_31	1000	0.0
exp_12	10	5.0
exp_22	100	5.0
exp_32	1000	5.0
exp_13	10	10.0
exp_23	100	10.0
exp_33	1000	10.0
exp_14	10	20.0
exp_24	100	20.0
exp_34	1000	20.0

Reporte sus resultados en gráficos de índice de éxito y recompensas por episodio promedio (cada gráfico debe estar asociado a un ID de experimento dado, y debe ser construido a partir de tres entrenamientos independientes). Adjunte los archivos .csv con los resultados obtenidos, y también el código empleado para cargar, procesar y graficar estos datos.

- 2.2 Entrene un agente en el ambiente `MountainCar` usando la configuración del código base, y la política experta proporcionada, las mismas combinaciones de parámetros que para `CartPole`, y la misma forma de reportar sus resultados.
- 2.3 Analice los resultados obtenidos para el ambiente `CartPole` y `MountainCar`. En particular, comente las variaciones en los resultados obtenidos al modificar los parámetros `nb_rollouts` y `alpha`.

## Reglas de formato

Las entregas deben cumplir con los siguientes requerimientos:

- Reporte en formato PDF.
- Incluya un archivo `README.txt` junto al código, donde indique las versiones de las dependencias que utilizó, y las instrucciones de ejecución de su código.
- Entregas parciales y finales en formato zip (reporte y código en un único archivo).
- Figuras legibles, de preferencia vectorizadas.
- Enumerar las respuestas de la misma forma en que se enumeran las preguntas.
- Respuestas concisas. No es necesario describir gráficos, es suficiente con mostrarlos en el reporte, y responder las preguntas que se realizan textualmente.

- No es necesario crear una portada, no obstante, la primera página debe contar con:
  - Título con formato “Avance Tarea X o Entrega Tarea X”, según corresponda.
  - Código del curso.
  - Nombre del estudiante.