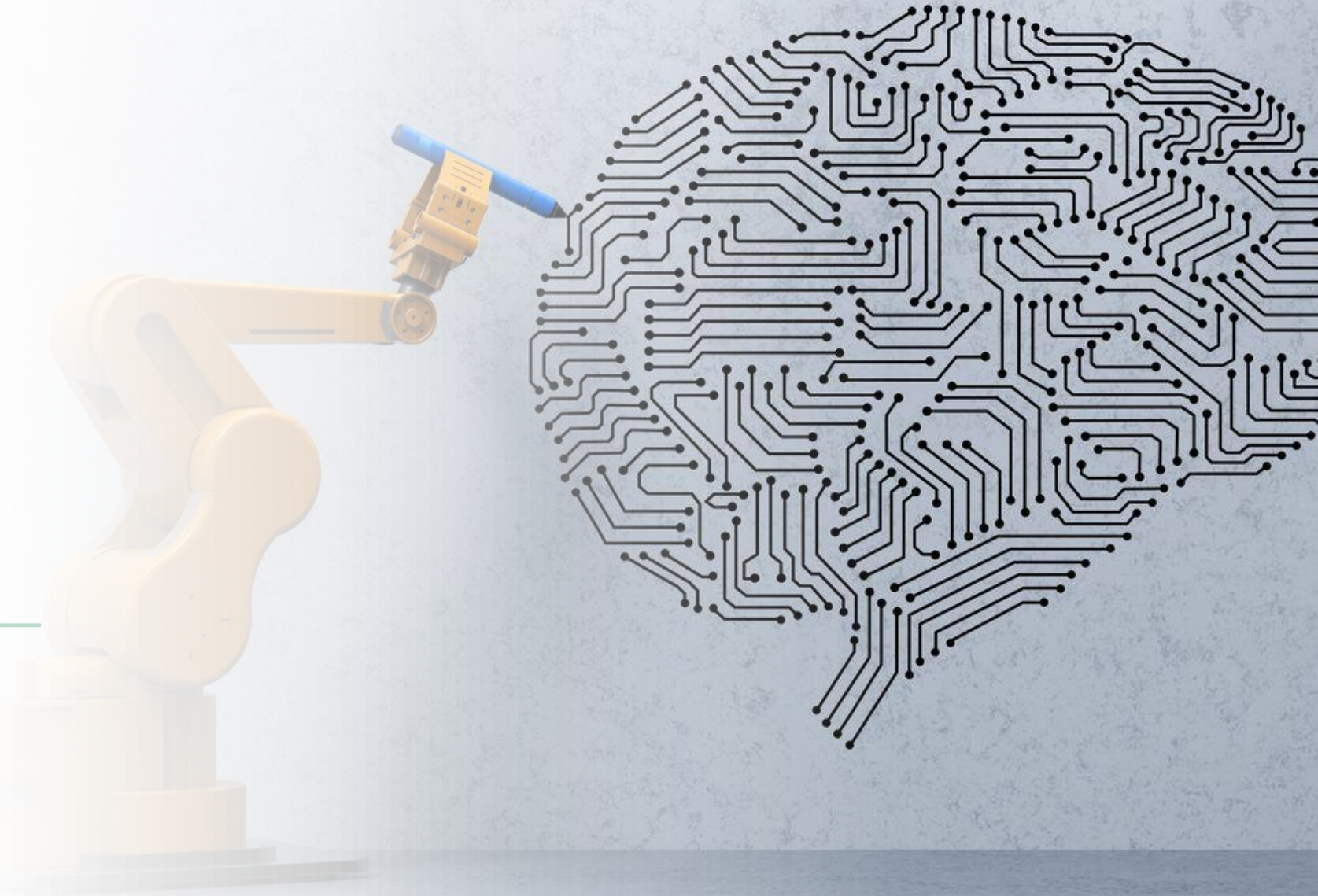
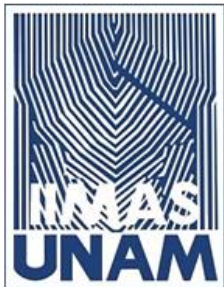


Aprendizaje por refuerzo

Clase 1: Introducción



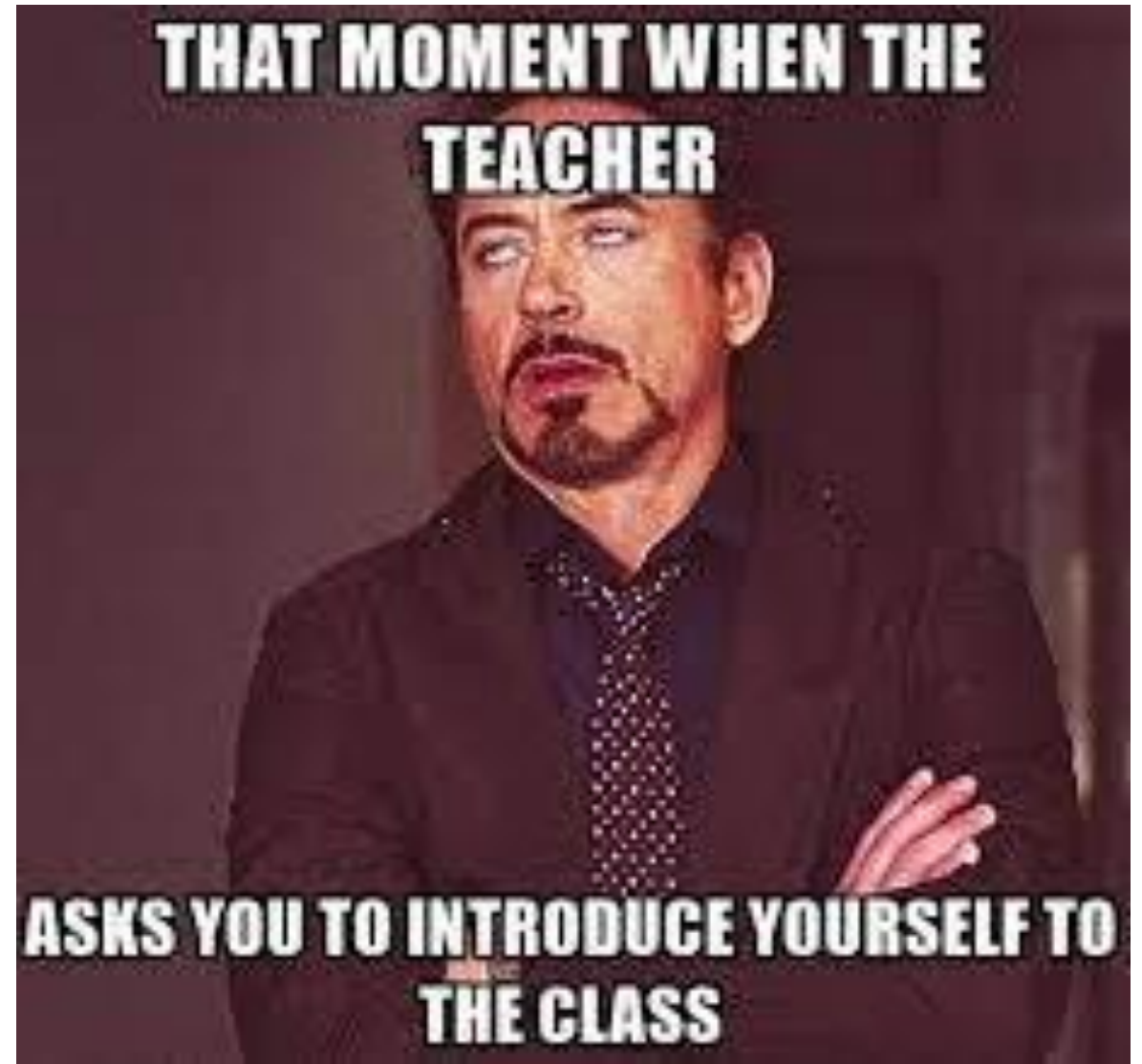
Para el día de hoy...

- Información general
- Evaluación
- Introducción a aprendizaje por refuerzo



Presentaciones

- ¿Quiénes son?
- ¿En qué están trabajando y con quien?
- ¿Qué esperan del curso?



Información del curso

Horarios

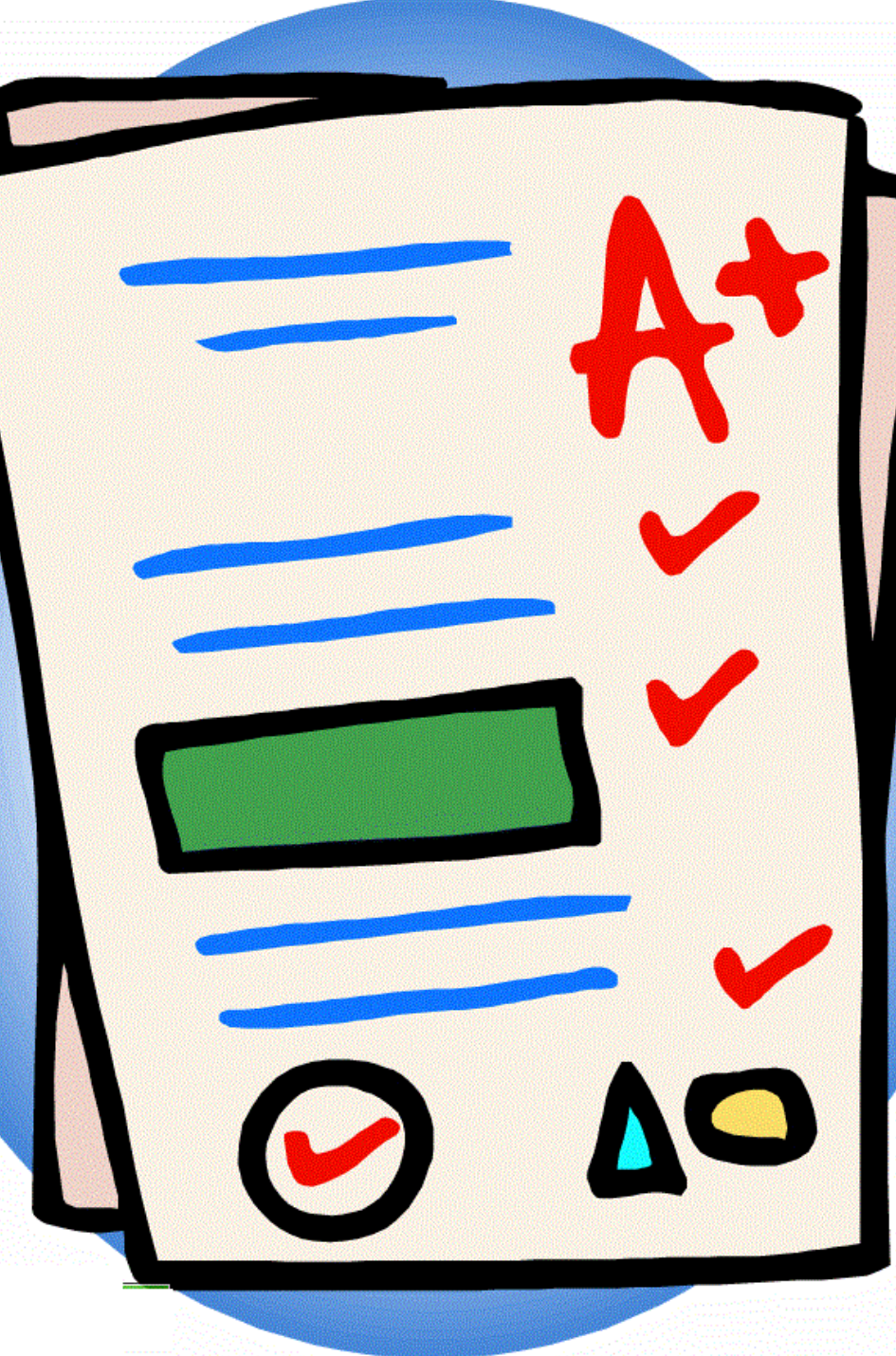
- Clase: lunes y miércoles 13:00 a 14:30

Datos útiles

- 17 semanas
- 26+4 sesiones
- Días inhábiles: febrero 6, marzo 20, abril 3 y 5

Contacto

- carlos.hernandez@iimas.unam.mx
- daniel_gn@comunidad.unam.mx
- vicool2@comunidad.unam.mx



Evaluación

- Tareas 30%
- Exámenes 30%
- Proyecto 30%
- Participación 10%



Tareas

- Los trabajos son individuales
- Se permite la colaboración
- En caso de comprobar plagio, la calificación del estudiante NO será aprobatoria
- Es requisito entregar todas las tareas para aprobar el curso (fecha límite 26/05/2022)
- Es posible entregar tareas posterior a la fecha de entrega con una penalización de 10% por día de atraso (hasta llegar a cero). 20% por cambio de fecha de entrega
- Aplican reglas especiales por Covid. Si tienen algún problema, háganlo saber



Exámenes

- Parcial 1: marzo 28
- Parcial 2: mayo 23
- Final 1: mayo 30
- Final 2: junio 1



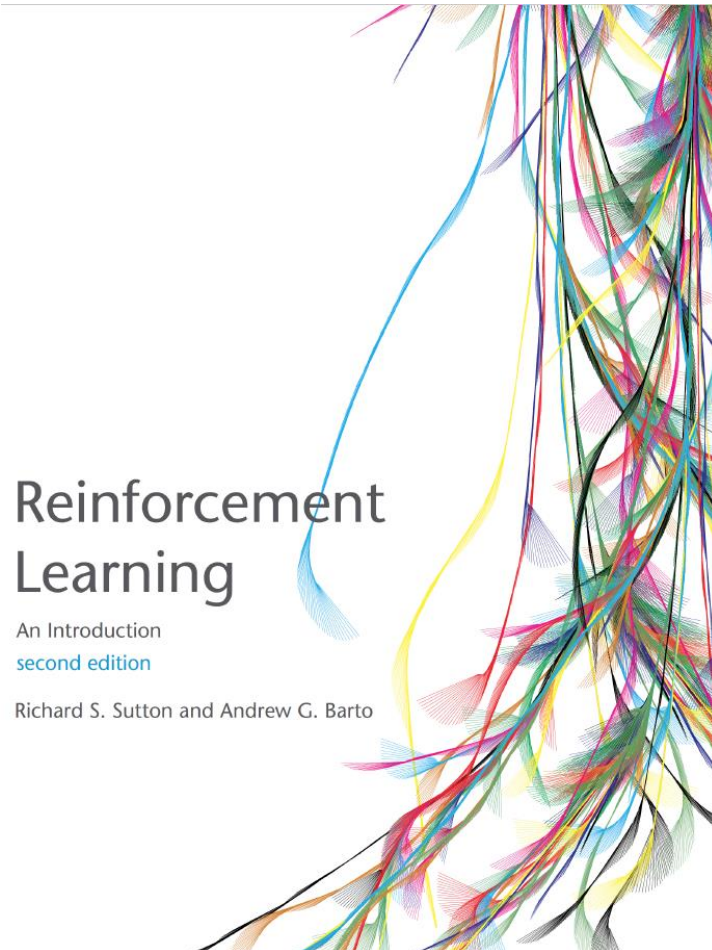
Proyecto

- Aplicación de aprendizaje por refuerzo
 - Contribución al “mundo real”
 - Aplicación a su tesis
 - Replicar estado de arte e investigación de trabajo futuro (hasta 2 personas)
- Se evaluarán 3 componentes
 - 40% Implementación: código reproducible y pruebas (indispensable)
 - 30% Científico: resumen en extenso en inglés de 4 páginas
 - 30% Presentación en inglés
- Fecha de entrega: 25 de mayo
- Se detallarán los elementos posteriormente

Otros comentarios

- Se utilizará Google classroom para dar seguimiento al curso: [n4zyeed](#)
- Todas las sesiones serán grabadas y podrán ser tomadas como sustituto a la clase en caso de no contar con conexión
- Durante las clases habrá ejercicios que contarán para participación. En caso de no asistir se pueden reponer las actividades sin bonificación
- Los oyentes son bienvenidos y serán tratados como estudiantes regulares
- Y más importante... la idea es que ¡aprendan los contenidos!

Recursos



- Libro: Reinforcement Learning An introduction
- Cursos de otras instituciones
 - UCL
<https://www.davidsilver.uk/teaching/>
 - UC Berkeley
<https://rail.eecs.berkeley.edu/deep-rlcourse/>
 - Stanford
<https://web.stanford.edu/class/cs234/>
 - University of Waterloo
<https://cs.uwaterloo.ca/~ppoupart/teaching/cs885-fall21/>

Objetivos del curso

General

- Diseñar agentes capaces de mejorar con la experiencia en un problema dado

Específicos

- Explicar los fundamentos de métodos de aprendizaje por refuerzo, sus alcances y limitaciones para solucionar problemas prácticos.
- Desarrollar y comparar métodos clásicos de planeación y control
- Resolver problemas donde no existe un modelo dado

Contenido

1. Conceptos básicos
2. Métodos tabulares
3. Métodos aproximados
4. Aprendizaje por refuerzo profundo
5. Algoritmos evolutivos
6. Temas avanzados

Calendario

Semana	Clase	Tema	Fecha clase	Tarea	Fecha entrega
1	1	Introducción	30/01/2022		
1	2	El problema	01/02/2022		
2		No hay clase	06/02/2022		
2	3	Control óptimo	08/02/2022		
3	4	Procesos de decisión de Markov	13/02/2022		
3	5	Programación dinámica	15/02/2022	T1	22/02/2022
4	6	Predicción libre de modelo	20/02/2022		
4	7	Control libre de modelo	22/02/2022		
5	8	Aproximación de función de valor	27/02/2022		
5	9	Métodos de política de gradiente	01/03/2022		
6	10	Integrando aprendizaje y planeación	06/03/2022	T2	13/03/2022
6	11	Exploración y explotación	08/03/2022		
7	12	CNNs y Deep Q learning	13/03/2022		
7	13	Bandido multi-brazo	15/03/2022		
8		No hay clase	20/03/2022		
8	14	AR basado en modelo	22/03/2022		
9	15	AR Bayesiano	27/03/2022		
9		Examen 1	29/03/2022	T3/Propuesta	05/04/2022
10		No hay clase	03/04/2022		
10		No hay clase	05/04/2022		
11	16	AR parcialmente observable	10/04/2022		
11	17	Búsqueda con adversarios	12/04/2022		
12	18	AR para optimización de caja negra	17/04/2022		
12	19	AR para optimización discreta	19/04/2022		
13	20	AR para multi-agentes	24/04/2022	T4	01/05/2022
13	21	AR con multiple recompensa	26/04/2022		
14	22	AR inverso	01/05/2022		
14	23	Aprendizaje multi-tarea y transf	03/05/2022		
15	24	AR distribuido	08/05/2022		
15	25	Meta aprendizaje	10/05/2022		
16	26	Teoría de la información	15/05/2022		
16		Cierre	17/05/2022		
17		Examen 2	22/05/2022		
17		Entrega de proyecto	24/05/2022		
18		Final 1	29/05/2022		
18		Final 2	31/05/2022		

Prerequisitos

- Nociones básicas de Python, pip y uso de bibliotecas de terceros
- Análisis y diseño de algoritmos
- Álgebra lineal
- Probabilidad





¡Empezamos!

- ¿Qué es una máquina?
- ¿Qué significa aprender?
- Y.. ¿Por qué queremos que aprenda?

Queremos
máquinas
que...





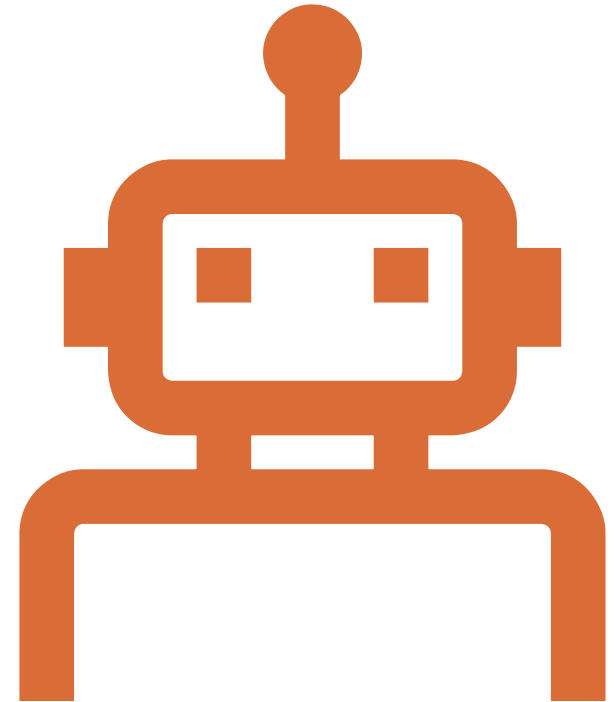
Nuestro enfoque: máquinas que actúen racionalmente

- Alcanzar metas preestablecidas
- Solo nos enfocamos en cuales decisiones se toman
- Las metas se expresan en terminos de la utilidad de los resultados
- ¡Maximizar la utilidad esperada!

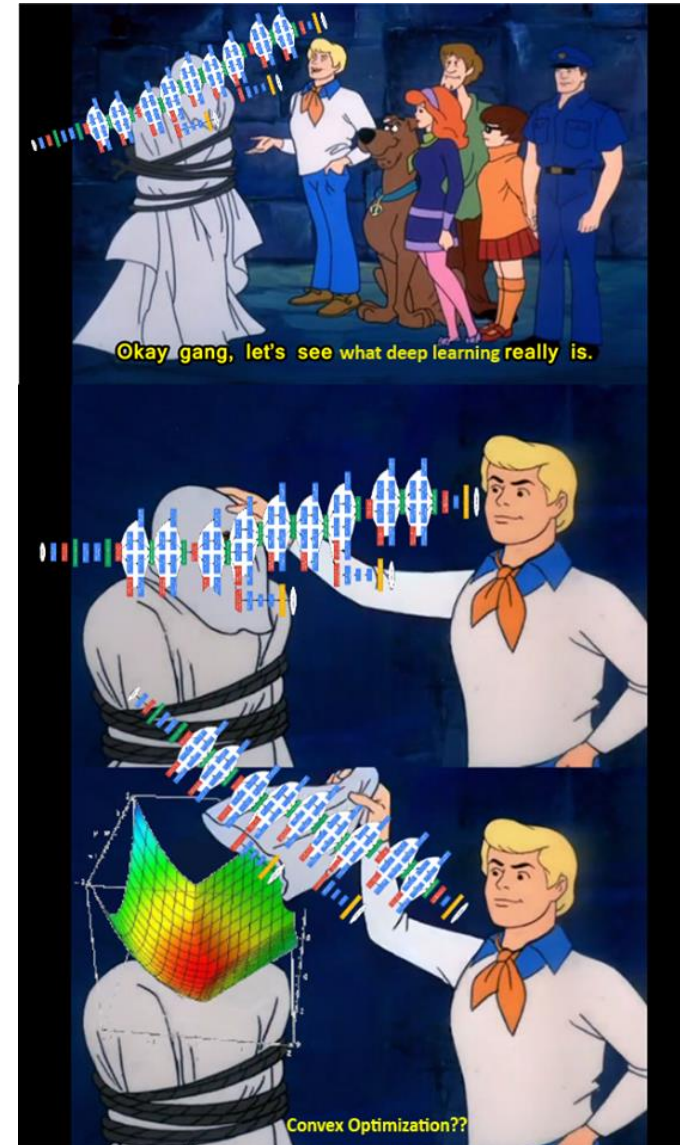
¿Qué es el aprendizaje máquina?


- “Learning is any process by which a system improves performance from experience.” - Herbert Simon
- Definition by Tom Mitchell (1998): Machine Learning is the study of algorithms that
 - improve their performance P
 - at some task T
 - with experience E

A well-defined learning task is given by $\langle P, T, E \rangle$



Una definición alternativa de aprendizaje máquina...





¿Cuándo
utilizar
aprendizaje
máquina?

No existe experiencia humana

Los humanos no pueden explicar su conocimiento

Los modelos deben ser personalizados

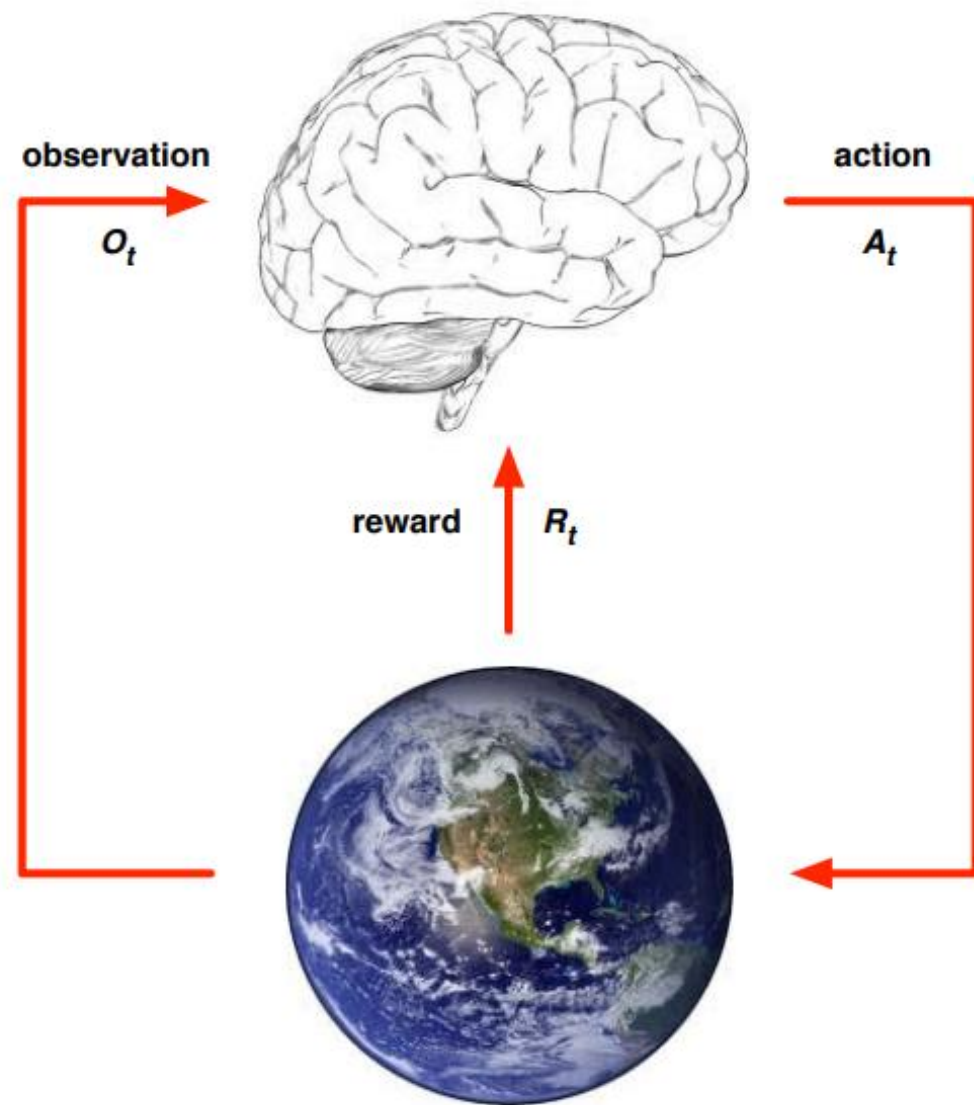
Los modelos están basados en una gran cantidad de datos

Tipos de aprendizaje

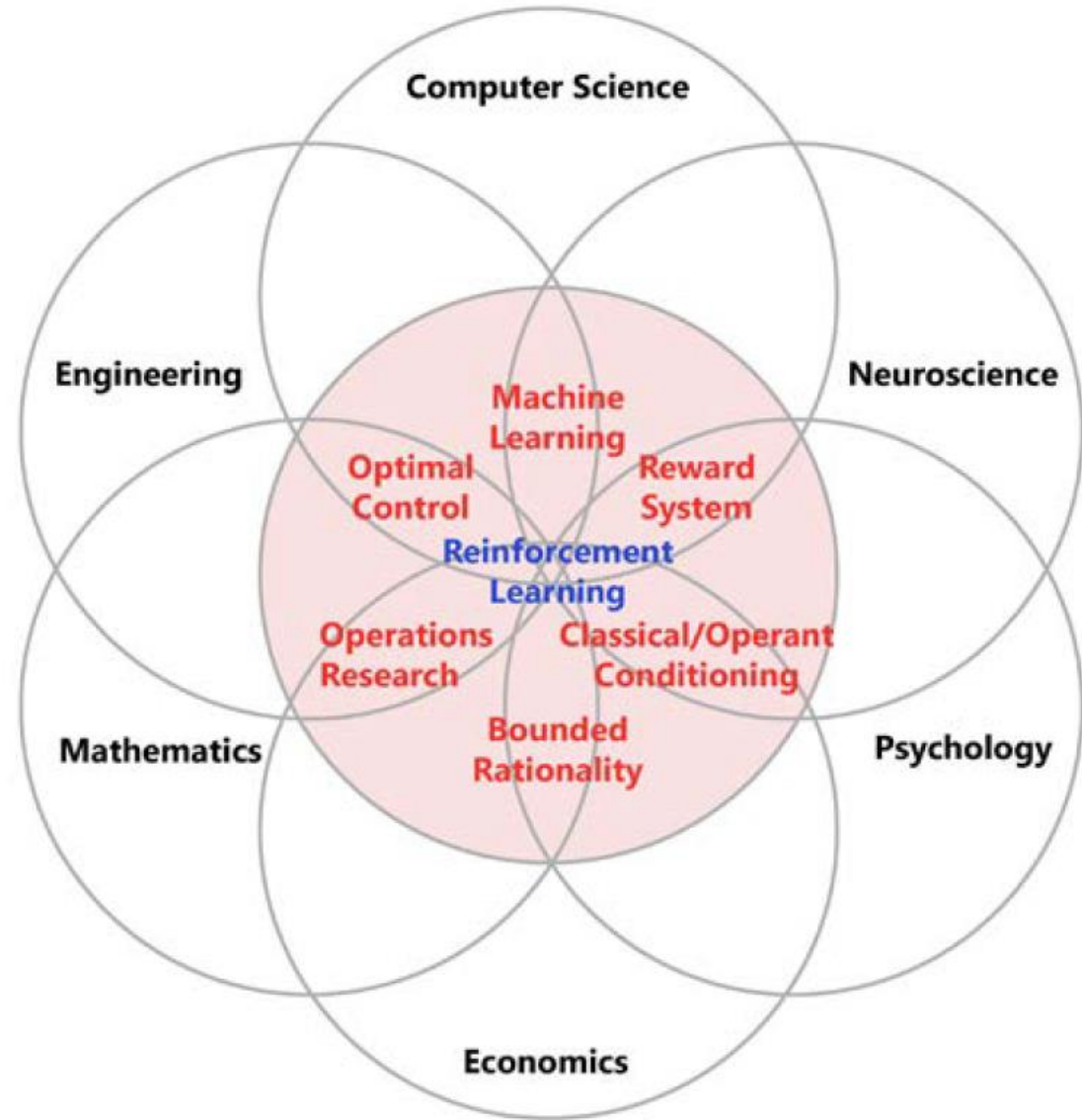
- Supervisado (regresión y clasificación)
 - Datos de entrenamiento
 - Salidas deseadas (etiquetas)
- No supervisado (agrupamiento)
 - Datos de entrenamiento
- Semi supervisado
 - Datos de entrenamiento
 - Algunas salidas deseadas
- Por refuerzo
 - Recompensa de la secuencia de acciones



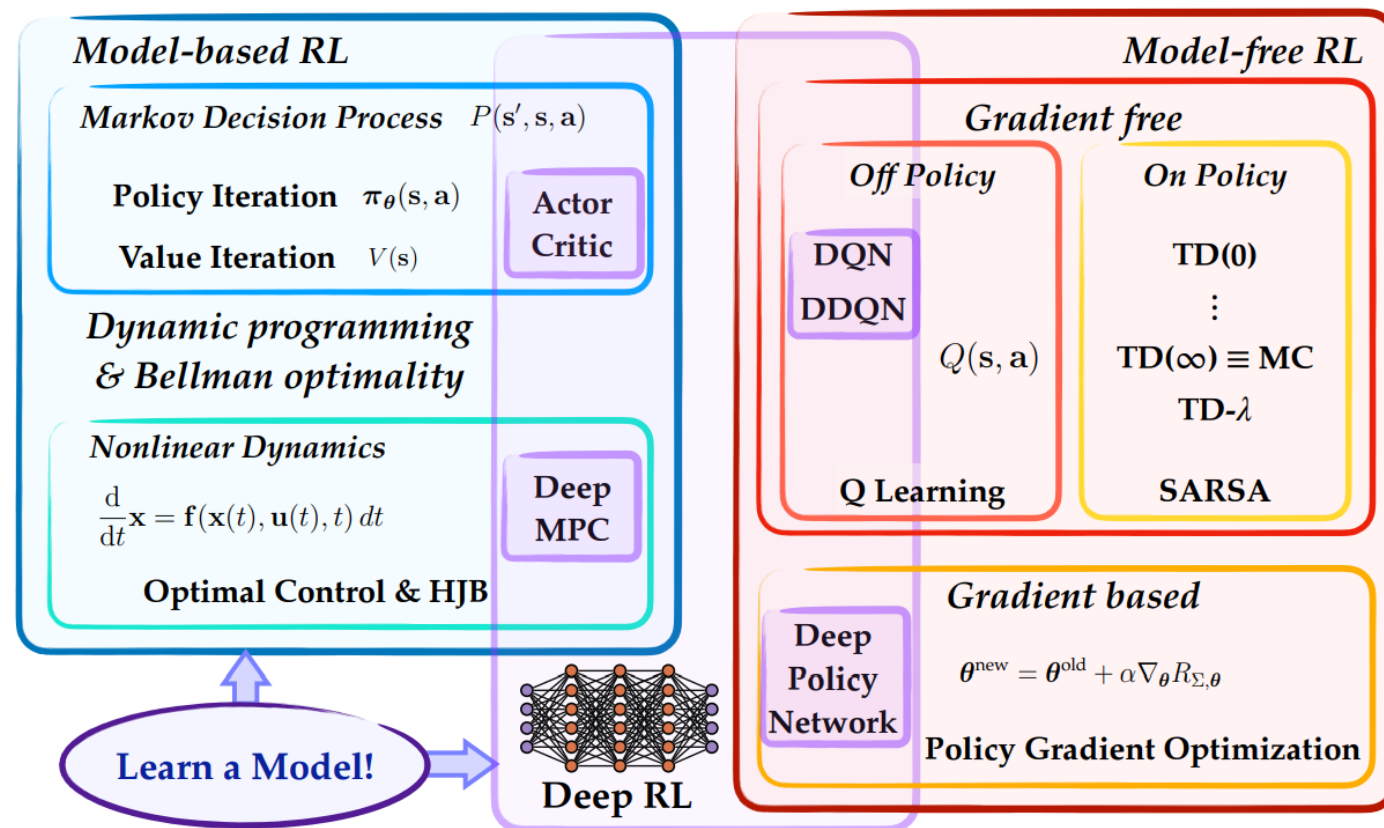
Aprendizaje por refuerzo



¿En dónde
está
ubicado?



Taxonomía de los métodos



El trabajo...

nature

[Explore content](#) ▾ [About the journal](#) ▾ [Publish with us](#) ▾ [Subscribe](#)

[nature](#) > [letters](#) > article

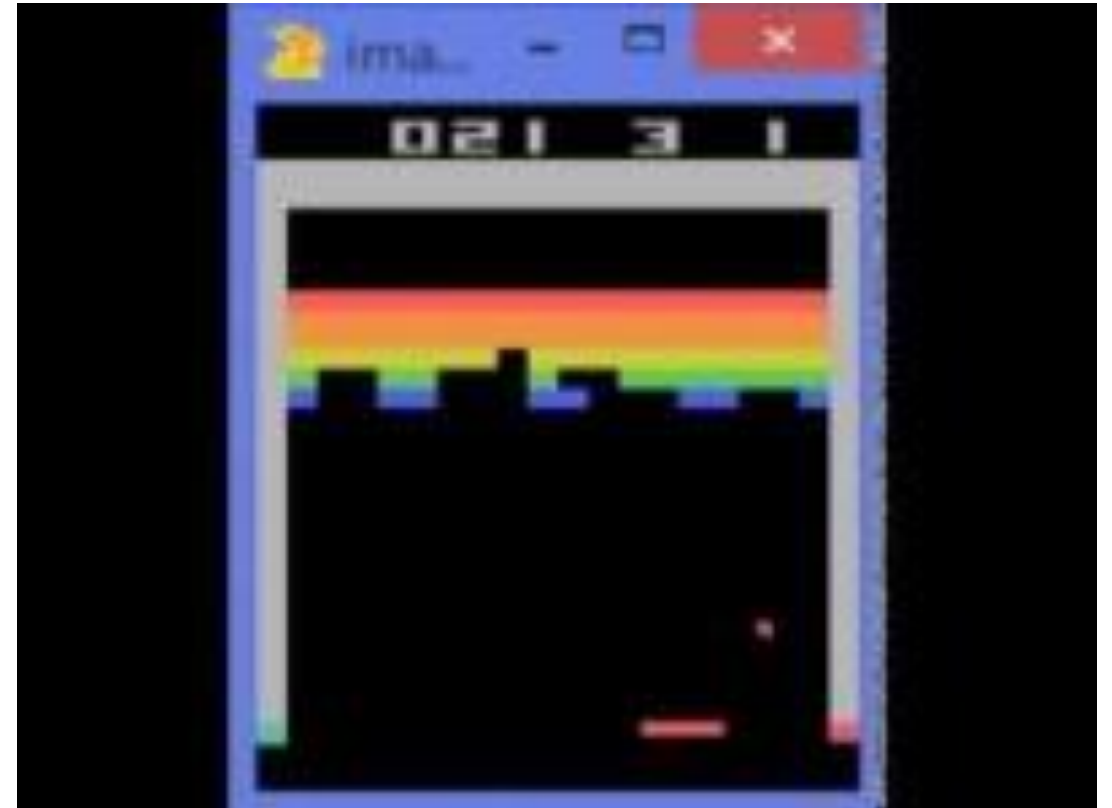
Published: 25 February 2015

Human-level control through deep reinforcement learning

[Volodymyr Mnih](#), [Koray Kavukcuoglu](#) , [David Silver](#), [Andrei A. Rusu](#), [Joel Veness](#), [Marc G. Bellemare](#), [Alex Graves](#), [Martin Riedmiller](#), [Andreas K. Fidjeland](#), [Georg Ostrovski](#), [Stig Petersen](#), [Charles Beattie](#), [Amir Sadik](#), [Ioannis Antonoglou](#), [Helen King](#), [Dharshan Kumaran](#), [Daan Wierstra](#), [Shane Legg](#) & [Demis Hassabis](#) 

[Nature](#) **518**, 529–533 (2015) | [Cite this article](#)

397k Accesses | **7885** Citations | **1550** Altmetric | [Metrics](#)



Y desde
entonces...

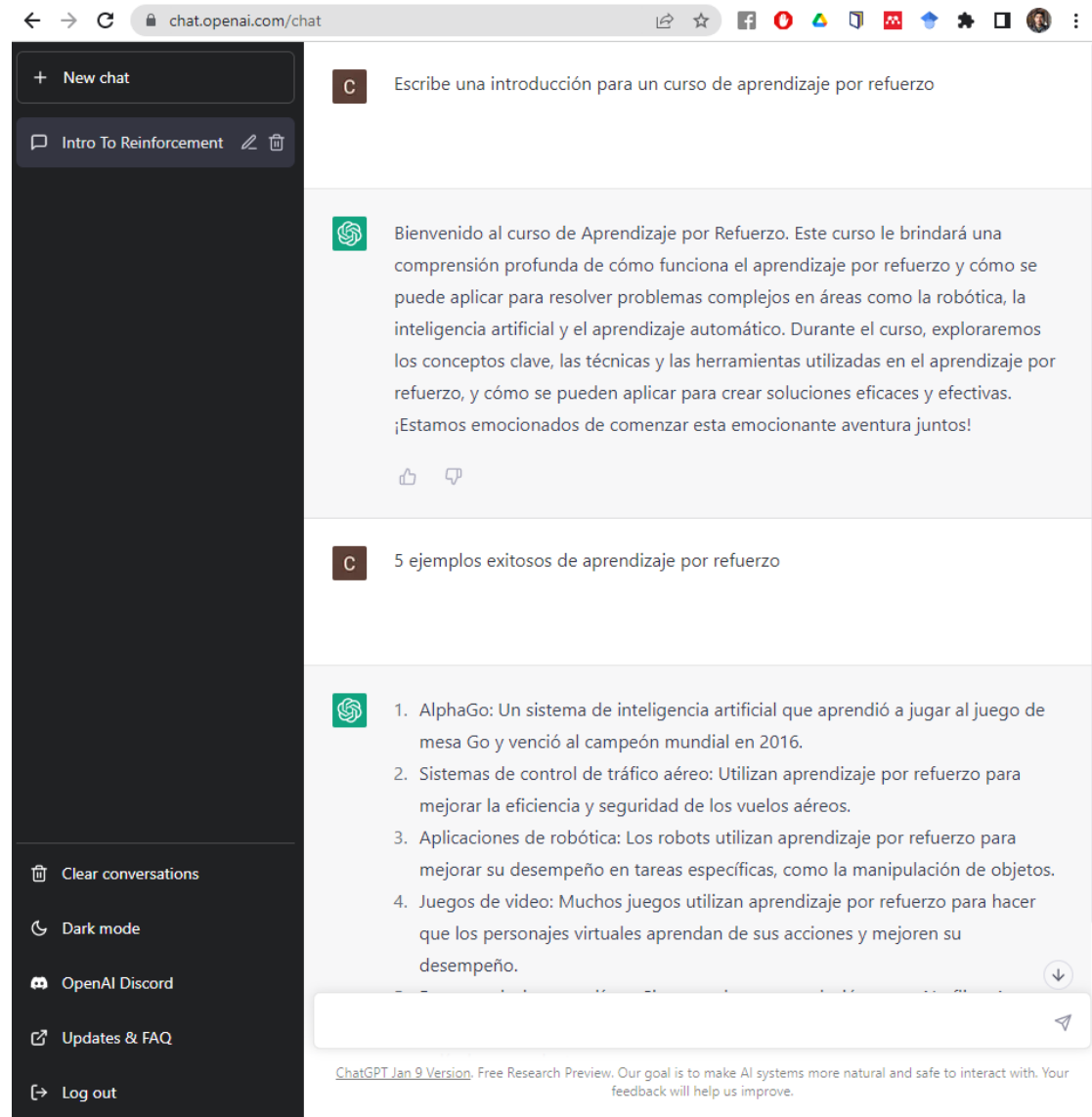


Áreas de aplicación

- Búsqueda web
- Biología computacional
- Finanzas
- E-commerce
- Exploración del espacio
- Robótica
- Redes sociales
- Y un gran etcétera



Otro ejemplo ¿exitoso? de RL



chat.openai.com/chat

+ New chat

Intro To Reinforcement

C Escribe una introducción para un curso de aprendizaje por refuerzo

Bienvenido al curso de Aprendizaje por Refuerzo. Este curso le brindará una comprensión profunda de cómo funciona el aprendizaje por refuerzo y cómo se puede aplicar para resolver problemas complejos en áreas como la robótica, la inteligencia artificial y el aprendizaje automático. Durante el curso, exploraremos los conceptos clave, las técnicas y las herramientas utilizadas en el aprendizaje por refuerzo, y cómo se pueden aplicar para crear soluciones eficaces y efectivas. ¡Estamos emocionados de comenzar esta emocionante aventura juntos!

5 ejemplos exitosos de aprendizaje por refuerzo

1. AlphaGo: Un sistema de inteligencia artificial que aprendió a jugar al juego de mesa Go y venció al campeón mundial en 2016.
2. Sistemas de control de tráfico aéreo: Utilizan aprendizaje por refuerzo para mejorar la eficiencia y seguridad de los vuelos aéreos.
3. Aplicaciones de robótica: Los robots utilizan aprendizaje por refuerzo para mejorar su desempeño en tareas específicas, como la manipulación de objetos.
4. Juegos de video: Muchos juegos utilizan aprendizaje por refuerzo para hacer que los personajes virtuales aprendan de sus acciones y mejoren su desempeño.

ChatGPT Jan 9 Version. Free Research Preview. Our goal is to make AI systems more natural and safe to interact with. Your feedback will help us improve.

Tarea 0

- Desarrollar en Python un juego de gato (Tic-tac-toe)
 - El ambiente debe contar una función `step(action)`
 - `action` son las coordenadas donde se colocará un "O" o "X"
 - La función debe regresar: `observation` (el tablero); el premio (1 si ganó, -1 si perdió o 0 en empate o si el juego no ha terminado); `done` (0 si el juego sigue, 1 si el juego terminó)
 - Un agente que juegue de forma aleatoria
 - Debe recibir una observación, elegir una posición vacía aleatoria y enviar una acción
 - Un despachador
 - Inicializa el juego y pide al jugador correspondiente que tome acción enviando la observación correspondiente
- Realizar un jugador aleatorio para el juego de FrozenLake (<https://gym.openai.com/envs/FrozenLake-v0/>)
- Para ambos problemas repetir 10 veces el juego y reportar la recompensa obtenida por los agentes.
- Entrega: 6 de febrero 2022 23:59
 - Archivos `tictactoe.py` y `frozenlake.py` que ejecuten los 10 experimentos y muestren en pantalla la recompensa de cada agente

Para la otra vez...

- Conceptos básicos

The End.