



Machine Learning.

Aprendizaje por refuerzo.

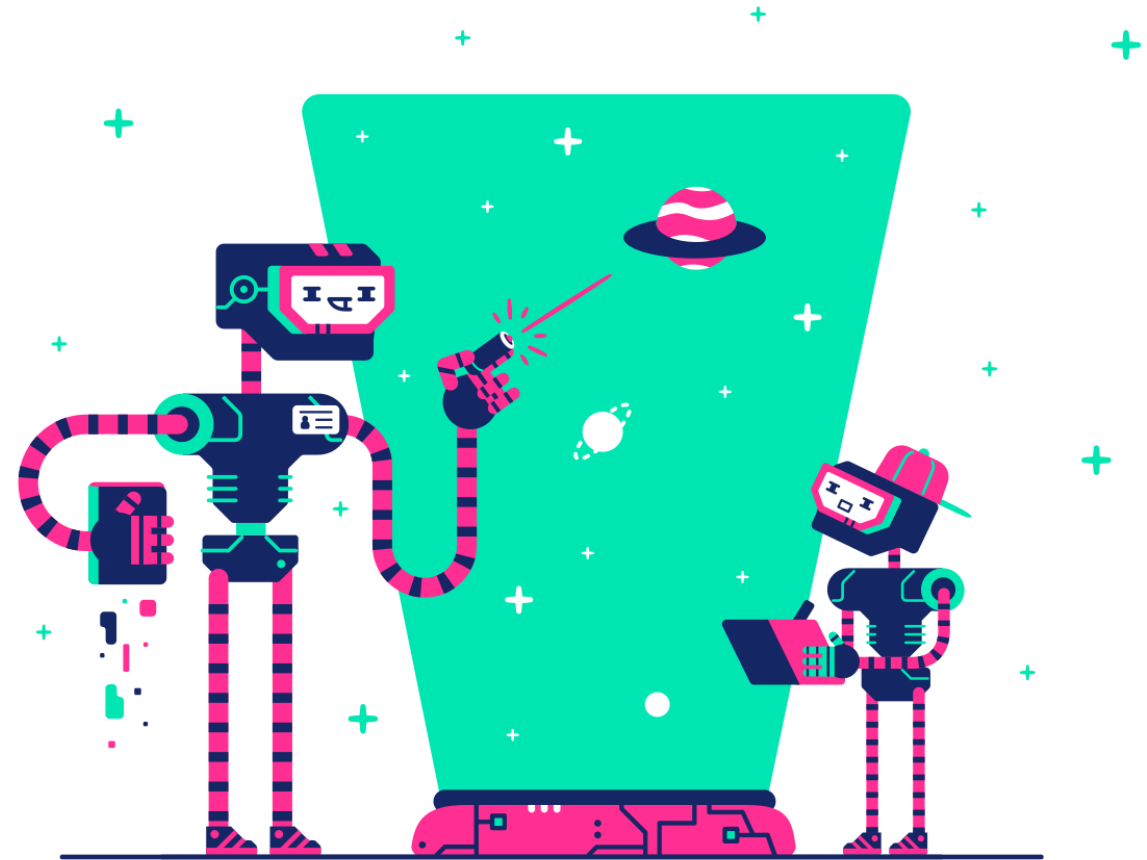
Programa de Tecnología en Cómputo

Instructor:

Juan Angeles Hernández (Adjunto)

Abraham Galindo Ruíz (Auxiliar)

Samuel Arturo Garrido Sánchez (Titular)



Aprendizaje por Refuerzo.

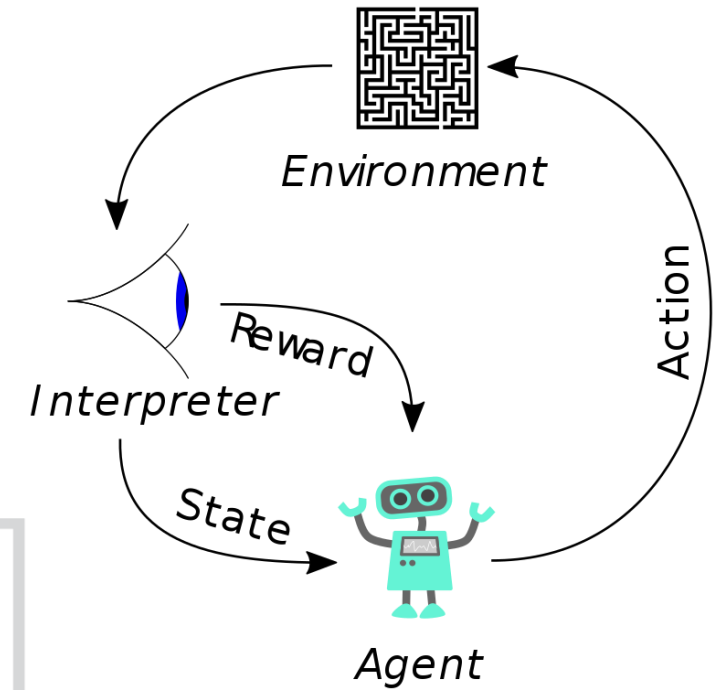
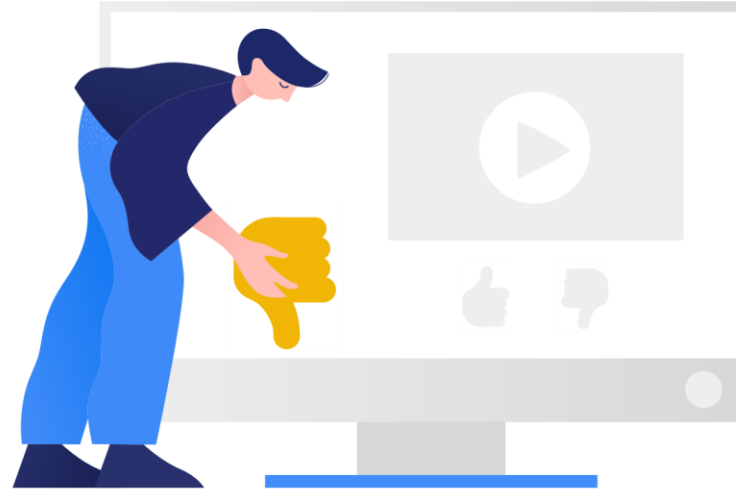
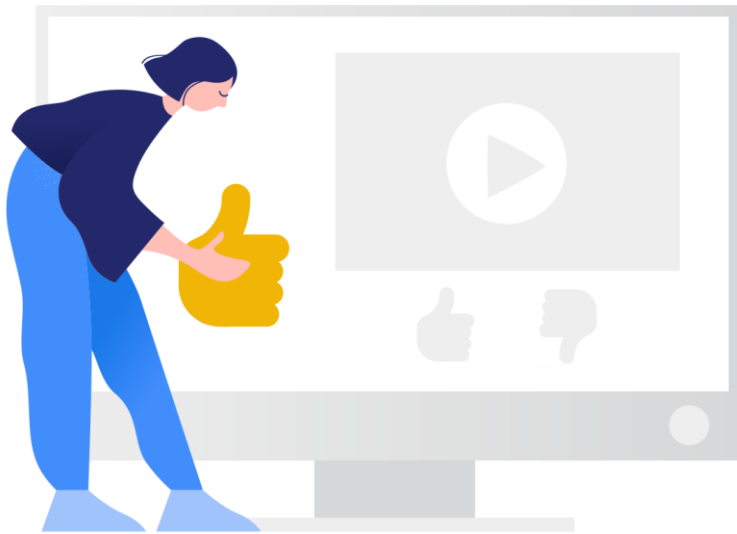
- ¿Qué es un aprendizaje por refuerzo?

Situación: Imagina que apenas estás aprendiendo a cocinar o hacer una actividad nueva.

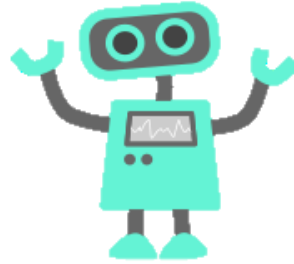


¿Cómo aprenden las máquinas?



- Usando recompensas y no recompensas.



El camino del laberinto.

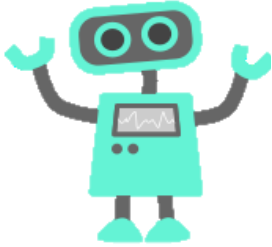


s estados
 a acciones
 R recompensas
 γ Descuento

$V=1$	$V=1$	$V=1$
		$V=1$
		$V=1$
		
	$R=-1$	$R=1$

Ecuación de Bellman.

$$V(s) = \max_a (R(s, a) + \gamma V(s'))$$



Conjunto de
Acciones

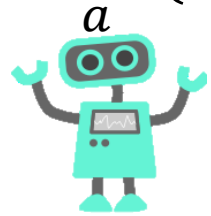
$A = \{\text{arriba, abajo, izquierda, derecha}\}$

$$\gamma = 0.9$$



$V=0.65$	$V=0.73$	$V=0.81$
$V=0.73$		$V=0.9$
$V=0.81$	$V=0.9$	$V=1$
$V=0.73$		
	$R=-1$	$R=1$

Ecuación de Bellman.

$$V(s) = \max_a (R(s, a) + \gamma V(s'))$$

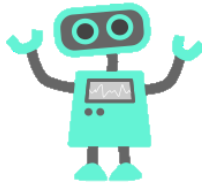


$$\gamma = 0.7$$

$V = 0.16$	$V = 0.240$	$V = 0.168$	$V = 0.1176$
	$V = 0.343$		
$V = 0.7$	$V = 0.49$		$V = 0.7$
$V = 1$	$V = 0.7$	$V = 0.7$	$V = 1$
	$V = 1$	$V = 1$	

Ecuación de Bellman.

$$V(s) = \max_a (R(s, a) + \gamma V(s'))$$



$$\gamma = 0.7$$

