

Pontificia Universidad Católica de Chile
Escuela de Ingeniería
Departamento de Ciencia de la Computación



IIC2613 - Inteligencia Artificial

Introducción a Machine Learning

Hans Löbel

Dpto. Ingeniería de Transporte y Logística
Dpto. Ciencia de la Computación

Algunos temas administrativos antes de empezar

- Todo sigue igual en cuanto a estructura (controles, ayudantías, tareas).
- Controles cortos: 21/10, 04/11, 18/11
- Control largo: 04/12
- Tareas:
 - T4: 21/10 -> 04/11
 - T5: 11/11 -> 25/11
- Todo el material queda en Canvas.

Fecha	Tema
07/10	Fundamentos de ML
09/10	
14/10	Árboles y ensambles
16/10	
21/10	
23/10	Support Vector Machines (SVM)
28/10	
30/10	
04/11	Aprendizaje Reforzado
06/11	
11/11	
13/11	
18/11	Redes Neuronales
20/11	
25/11	
27/11	

¿Cómo es esta segunda parte del curso?

- Tomaremos un enfoque de la IA con los datos como foco => **qué podemos aprender de ellos para resolver problemas.**
- Combinaremos la intuición con la formalización matemática de estas ideas.
- Más contenido, algo menos de discusión y código.
- **Menos modelación, más aprendizaje.**

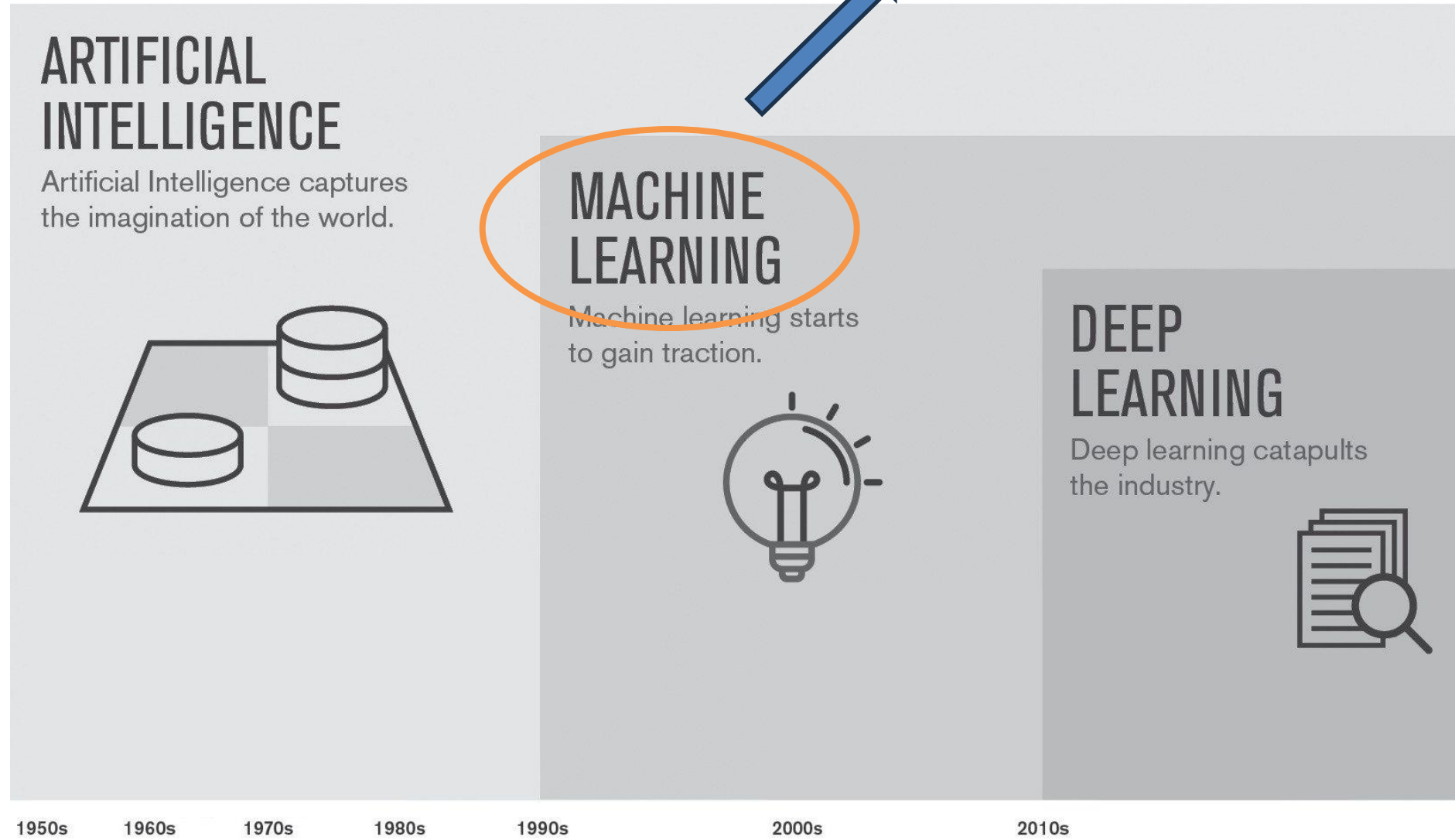


¿Qué es Inteligencia Artificial?

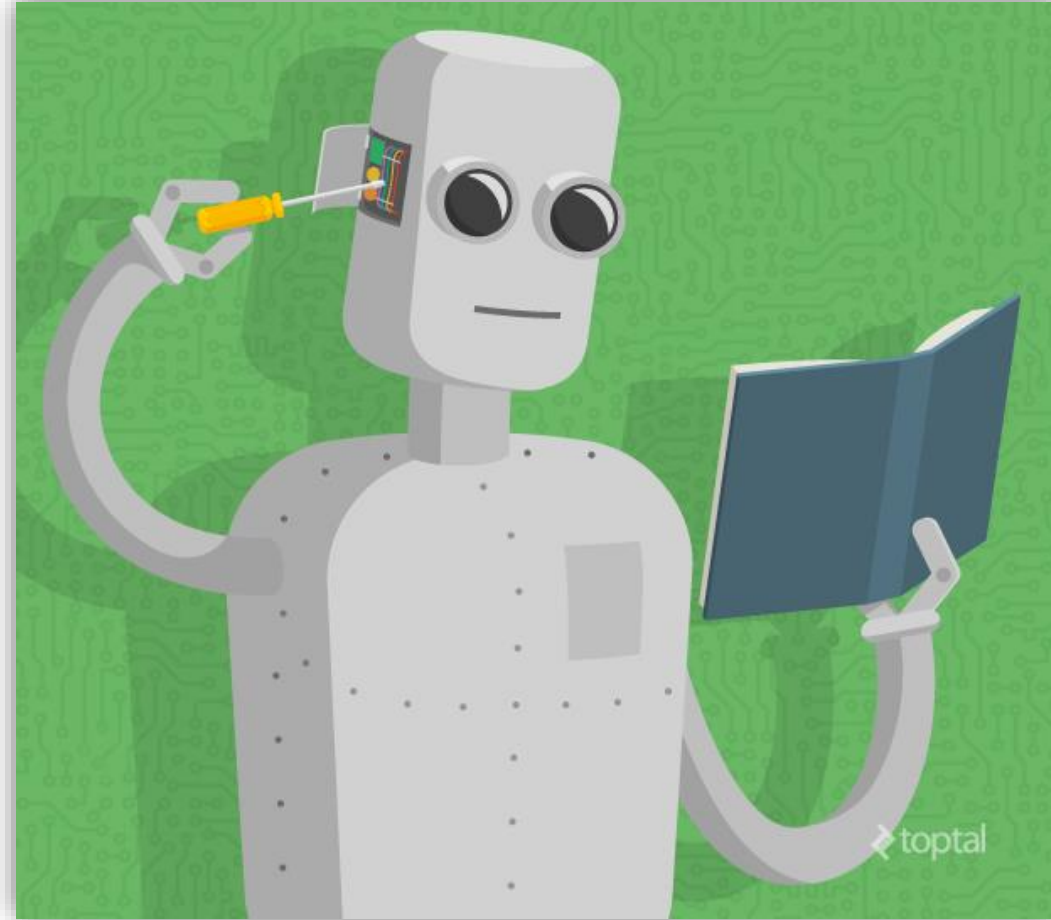
- Término creado a mediados de los años 50.
- **Métodos, algoritmos y tecnologías** que permiten que una **máquina (software)** muestre un comportamiento **inteligente**, (ojalá) **indistinguible** del de un **humano**.
- En la práctica, significa que las máquinas deben percibir, razonar y actuar con **gran flexibilidad**.
- Capturó el interés, imaginación y financiamiento de gran cantidad de personas y entidades.



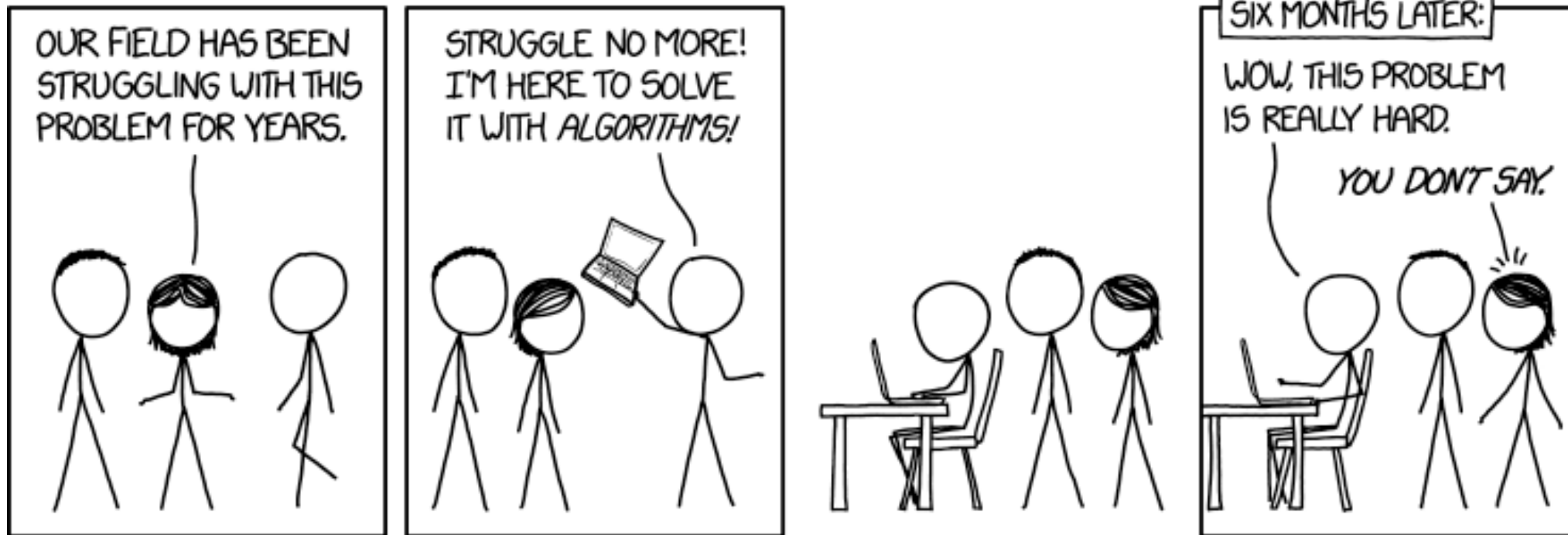
Aprendizaje de máquina
o
aprendizaje automático



¿Qué significa que una máquina aprenda?



¿Por qué no **programamos a la máquina** desde el principio con la capacidad de realizar lo que queremos?



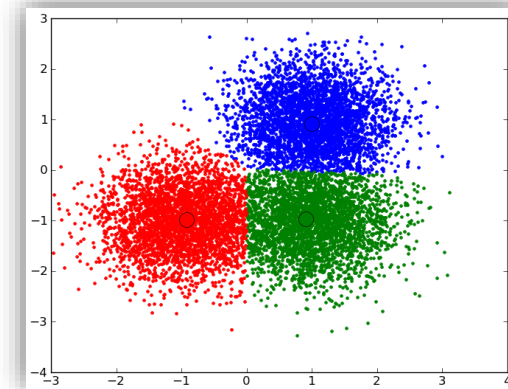
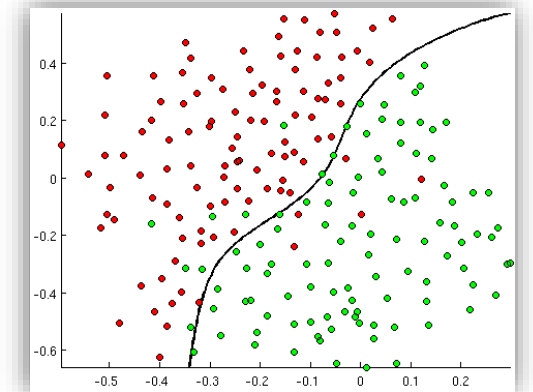
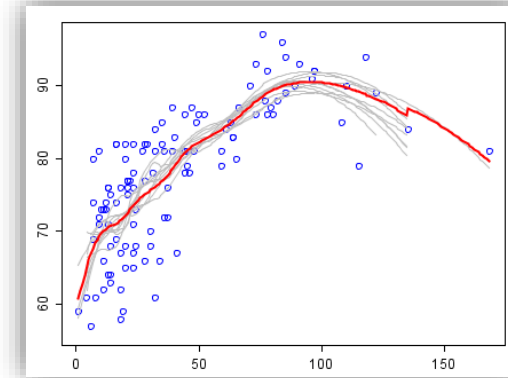
¿Por qué no **programamos a la máquina** desde el principio con la capacidad de realizar lo que queremos?



Ok, entiendo, no queda otra que **aprender** a resolver algunos problemas, pero **¿cómo?**

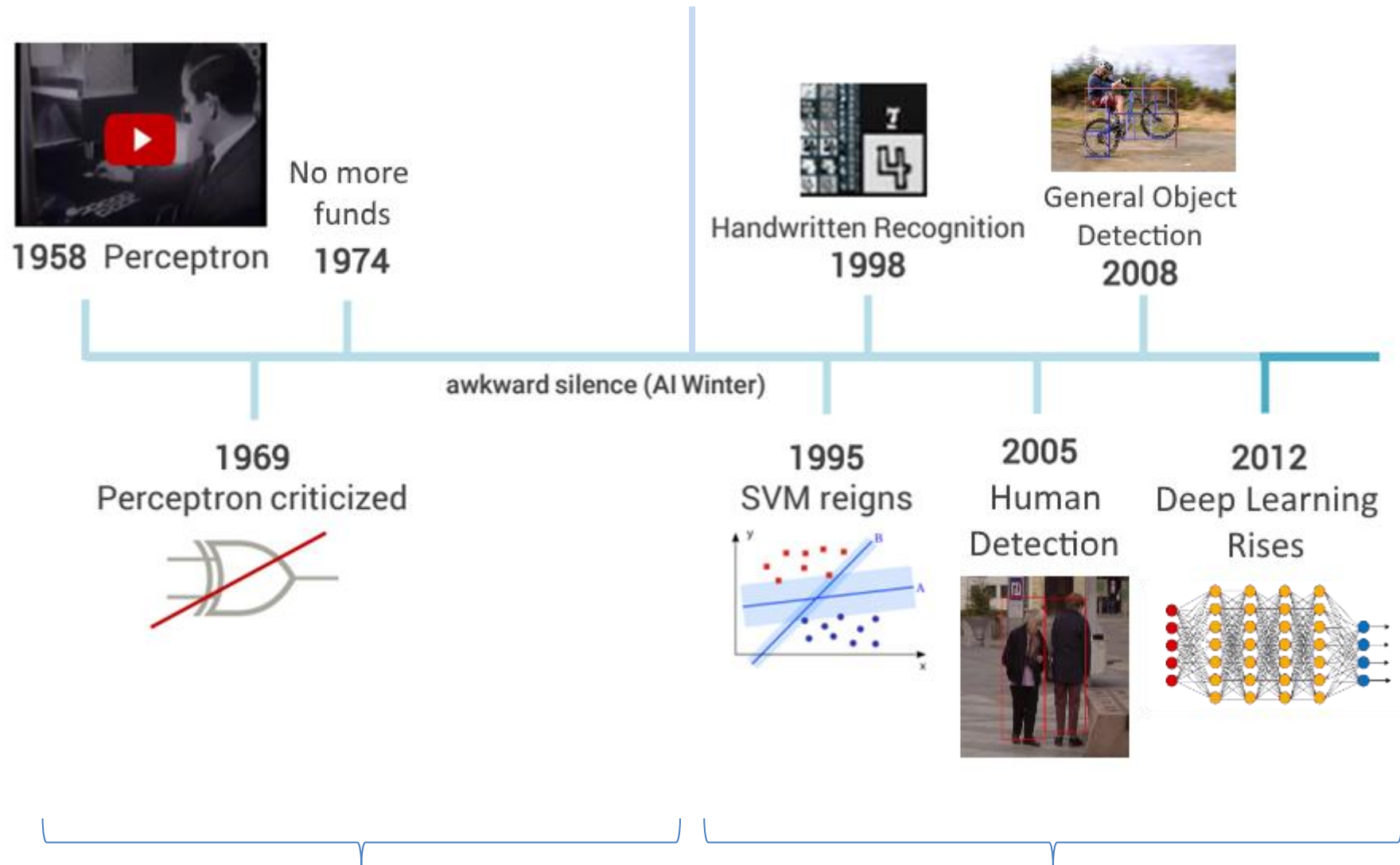
Machine Learning se centra en el estudio de algoritmos que aprenden inductivamente, mejorando su rendimiento en una tarea, a través de la experiencia

- Sistemas dedicados a resolver tareas/problemas.
- Aprenden inductivamente, generalizando a nuevos casos a partir de la experiencia (datos).
- Generalmente, mejoran su rendimiento al estar expuestos a más experiencia (mientras más datos, mejor).
- Buscan aprendizaje más que modelación de los datos (representaciones útiles del mundo)
- Sólidas bases teóricas (pero lo práctico va más rápido que la teoría)





Explosión digital (Big Data, IoT, etc) ≈ 1990

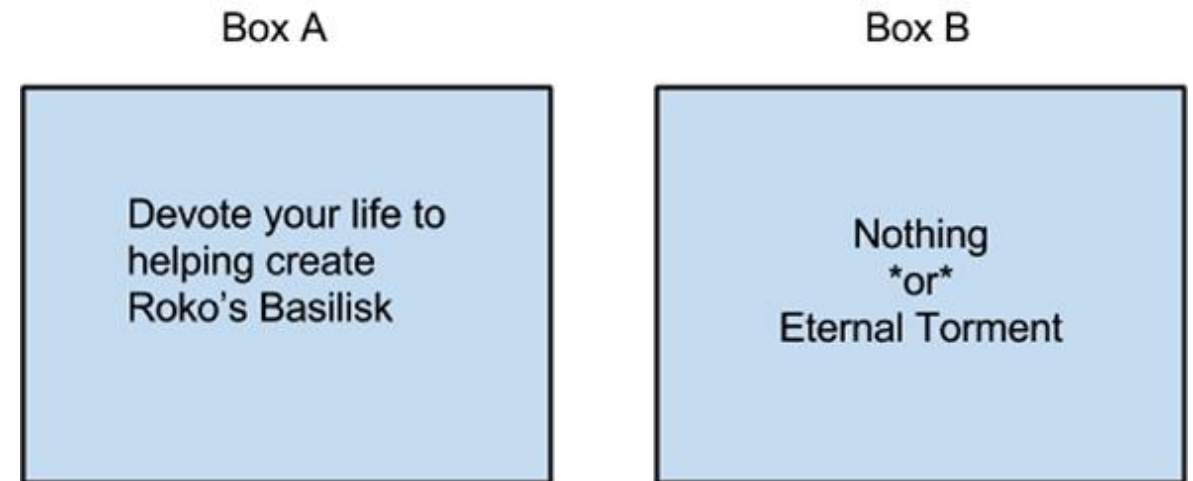


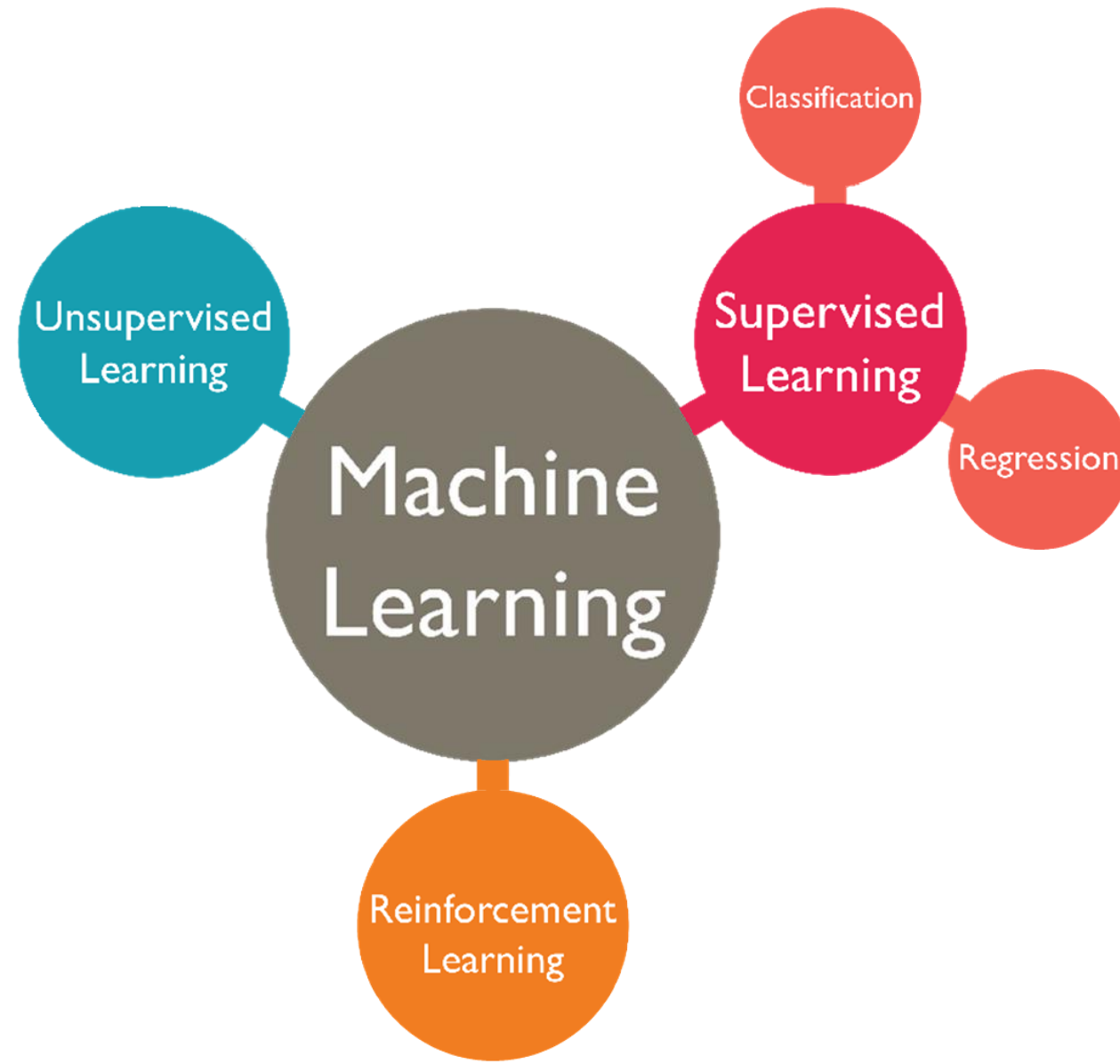
IA tradicional: aspiraba a construir máquinas tan inteligentes como un humano

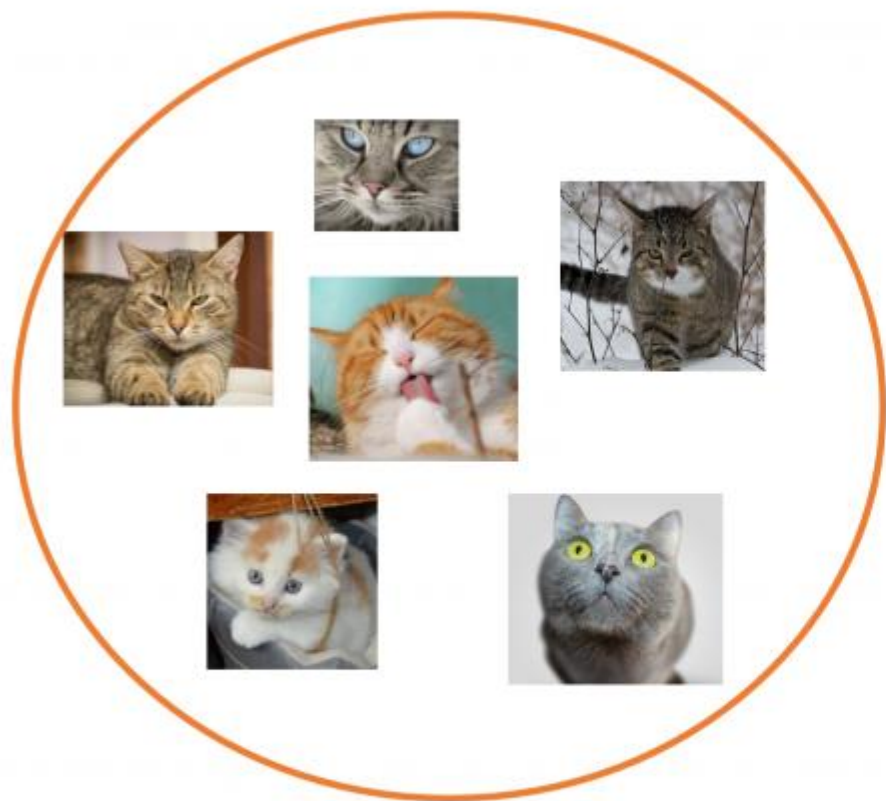
Machine Learning: máquinas especializadas, que aprenden **INDUCTIVAMENTE**, pueden superar a los humanos en la(s) tarea(s) asignada(s)

Roko's basilisk

- Experimento mental, centrado en una **hipotética super IA** del futuro.
- En él, esta IA **castigaría retroactivamente** a aquellos que no colaboraron en generar su existencia.
- Como la mayoría de estas personas estarían muertas, la IA las simularía y **castigaría en la simulación**.
- El sólo hecho de conocer sobre este experimento, hace que uds. ya estén participando (por la temática del curso, no deberían sufrir tormento eterno).







Cat

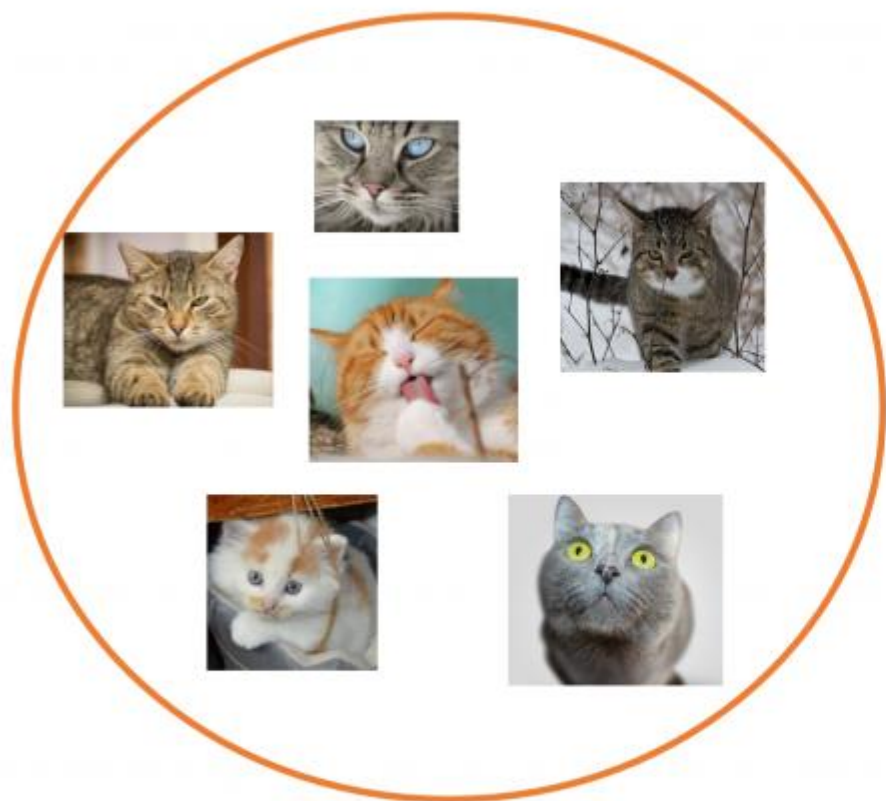


Dog

¿



?



Cat

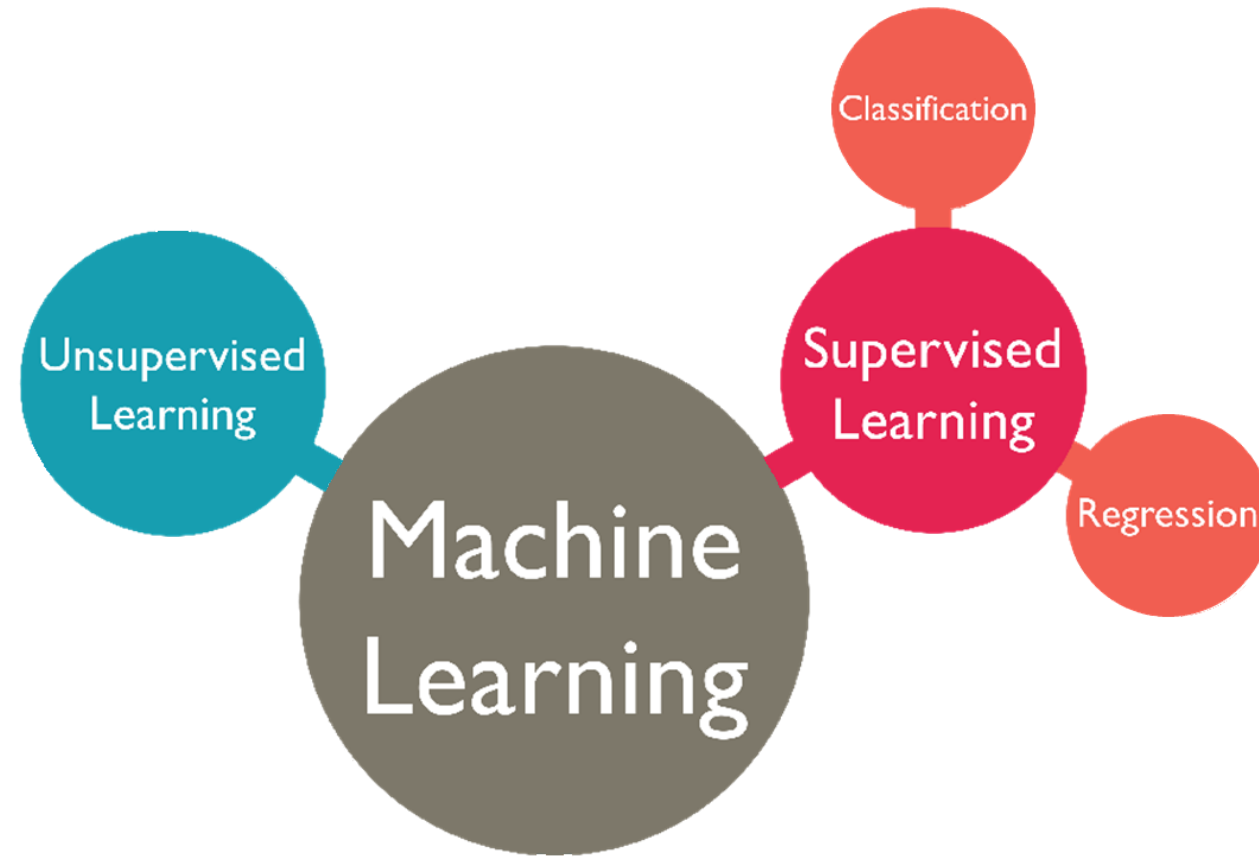


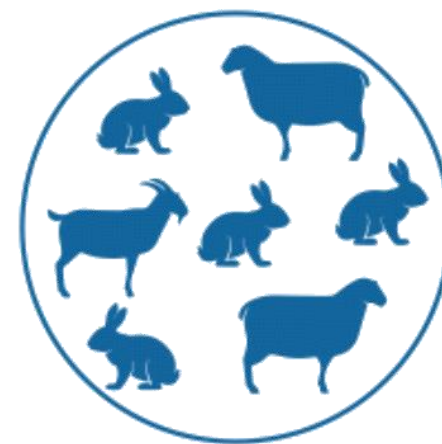
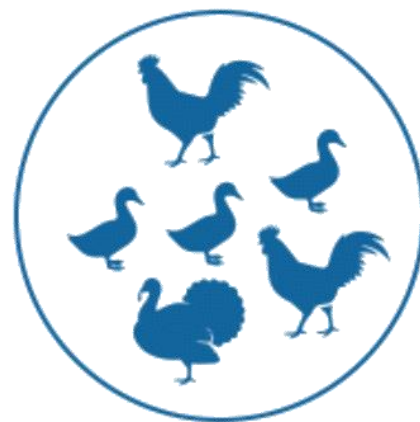
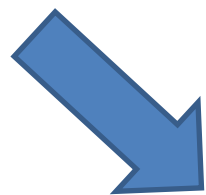
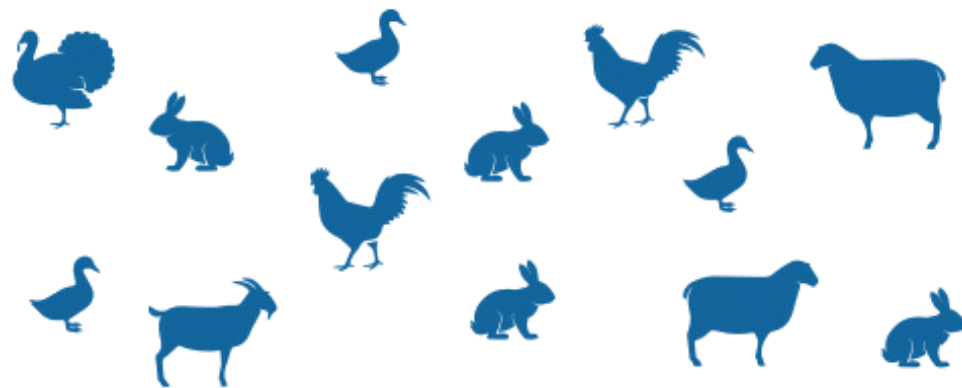
Dog

¿



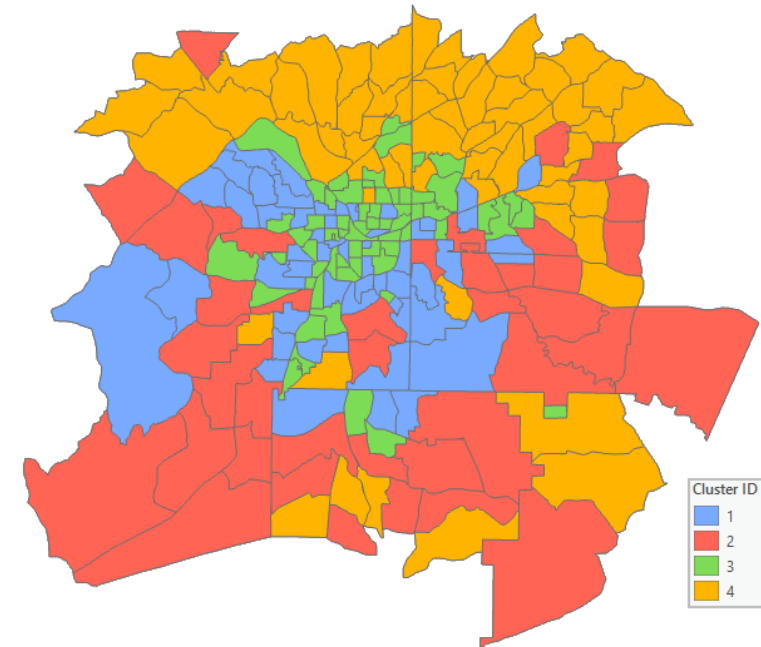
?

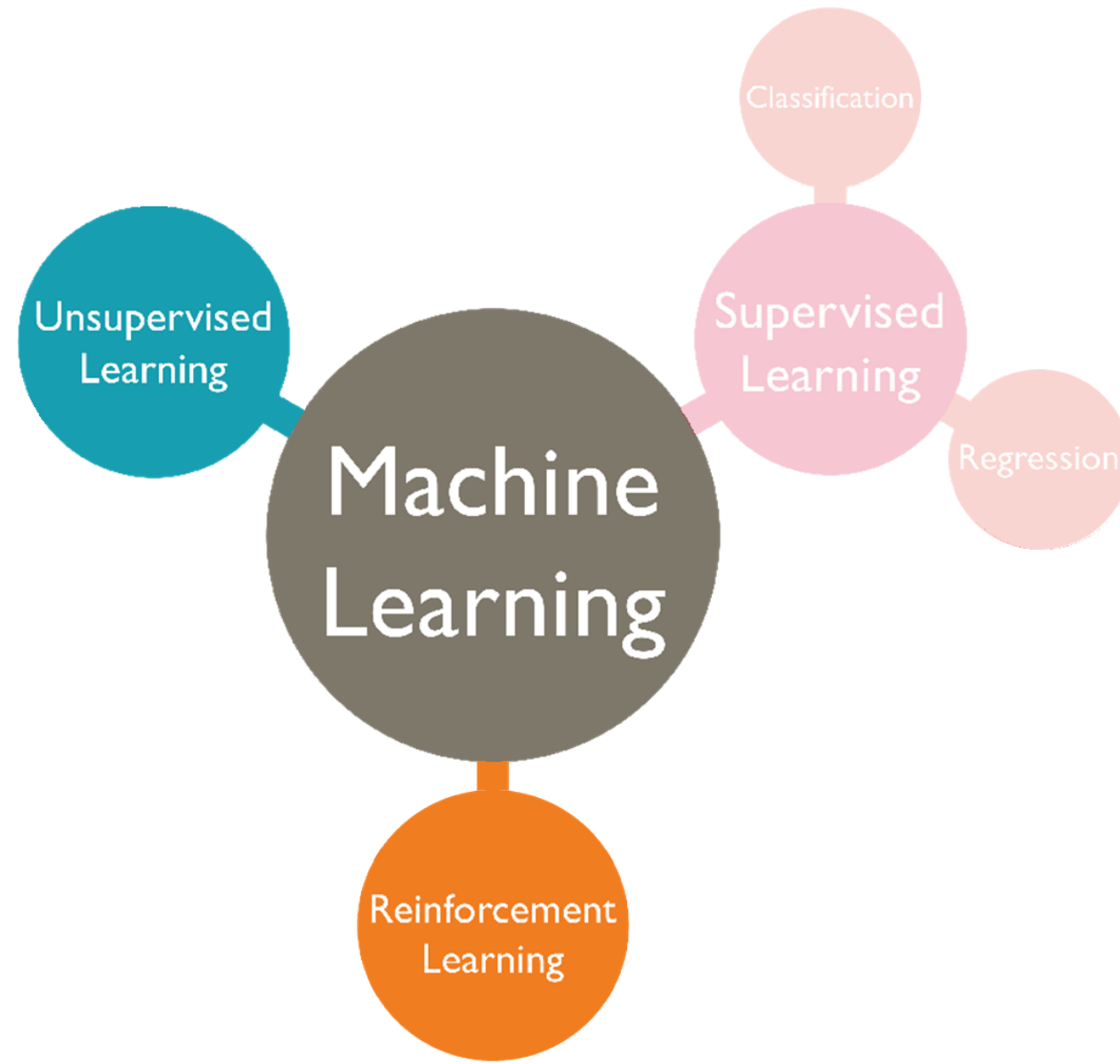


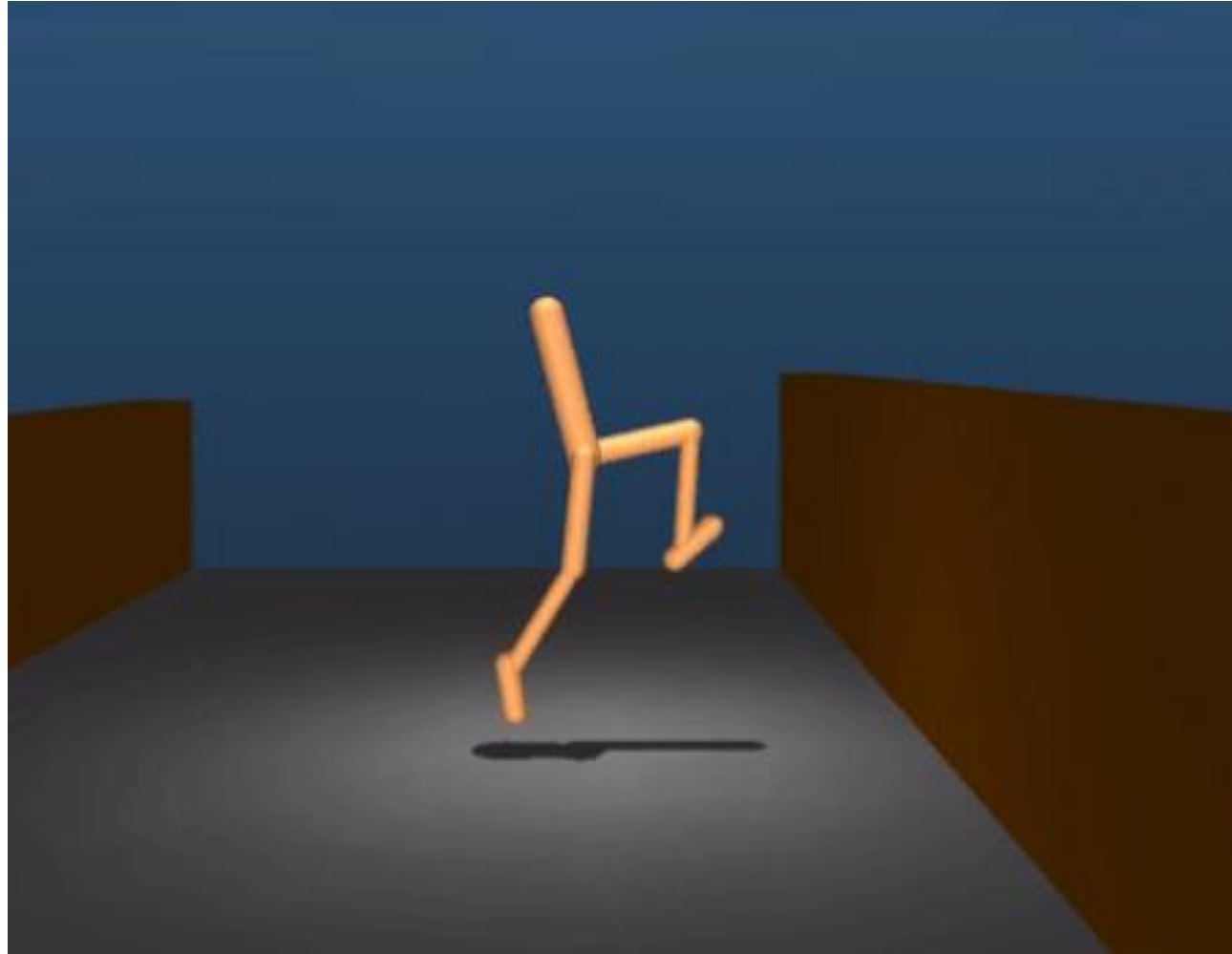




Type	Level	Description
Location	I/B [†]	Geocoded position of the unit
SES	I B	Socioeconomic status
Aesthetic	I/B	Aesthetic perception index of the nearest geocoded image (individual) or averaged at the urban block level
Political	I/B	Proportion of a political choice in the Constitutional plebiscite or in the Chilean Presidential Elections (first round)
Land use	I/B	Proportion/ $M^{2\ddagger}$ of land use of the urban block according to a land use typology*
Demographic	I B	Sex, age, proportion of immigrants Age, proportion of immigrants/women



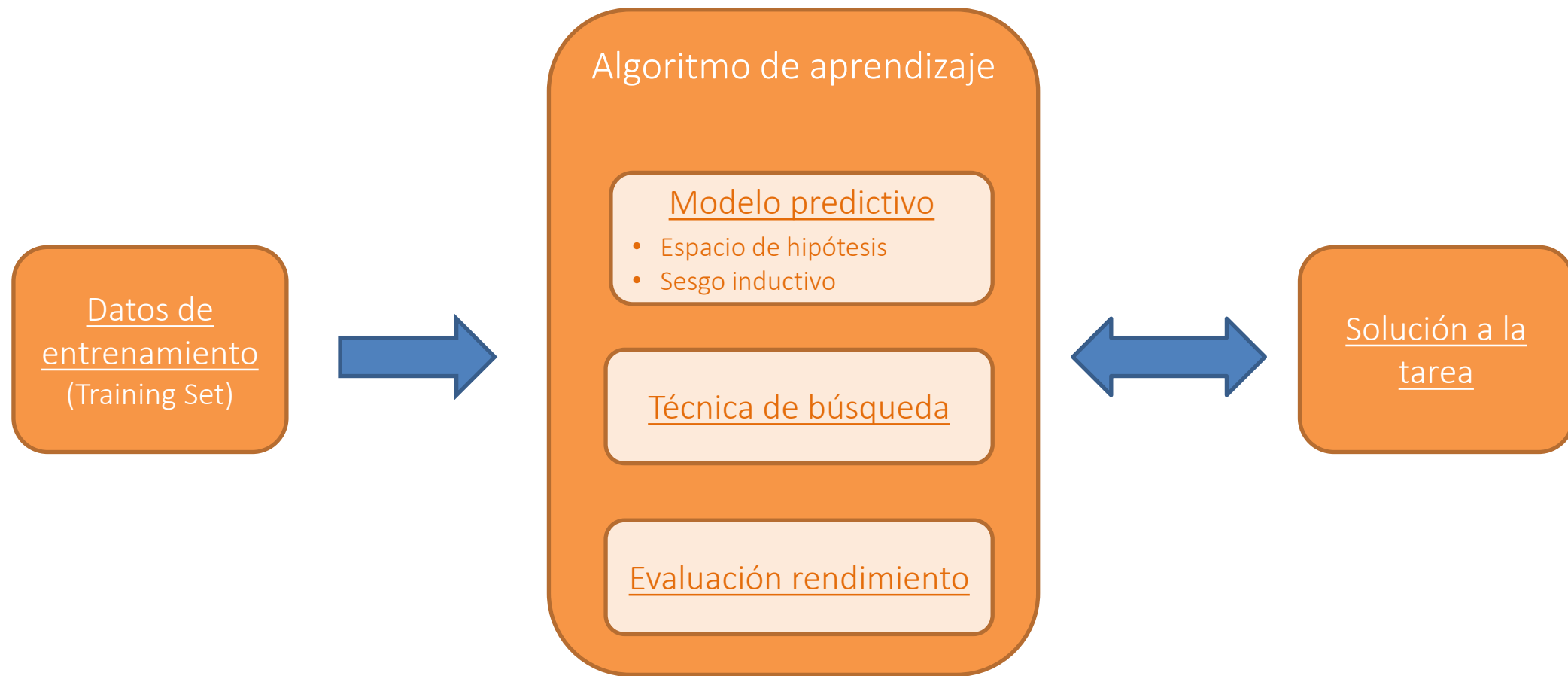






<https://youtu.be/KHMwq9pv7mg>

Todos los sistemas de Machine Learning presentan, a alto nivel, un esquema de funcionamiento similar



Buscamos, en el **espacio de hipótesis** del modelo, aquella que genera la solución con mejor **rendimiento**, usando el **sesgo inductivo** y los datos de **entrenamiento** para guiar la **búsqueda**.

¿Cómo se ve esto matemáticamente?

Problema de optimización: búsqueda sobre el espacio de hipótesis

Parámetros: viven en el espacio de hipótesis del modelo

Función de pérdida: captura el rendimiento del modelo al cuantificar su error en los datos de entrenamiento

Etiqueta: valor a predecir (tarea)

$$\operatorname{argmin}_W J(X, Y; W) = \lambda \mathcal{R}(W) + \sum_i^N \mathcal{L}(f(x_i; W), y_i)$$

Datos de entrenamiento: incluyen los datos en sí y sus etiquetas.

Regularizador: función que induce sesgo inductivo en el modelo predictivo a través de sus parámetros

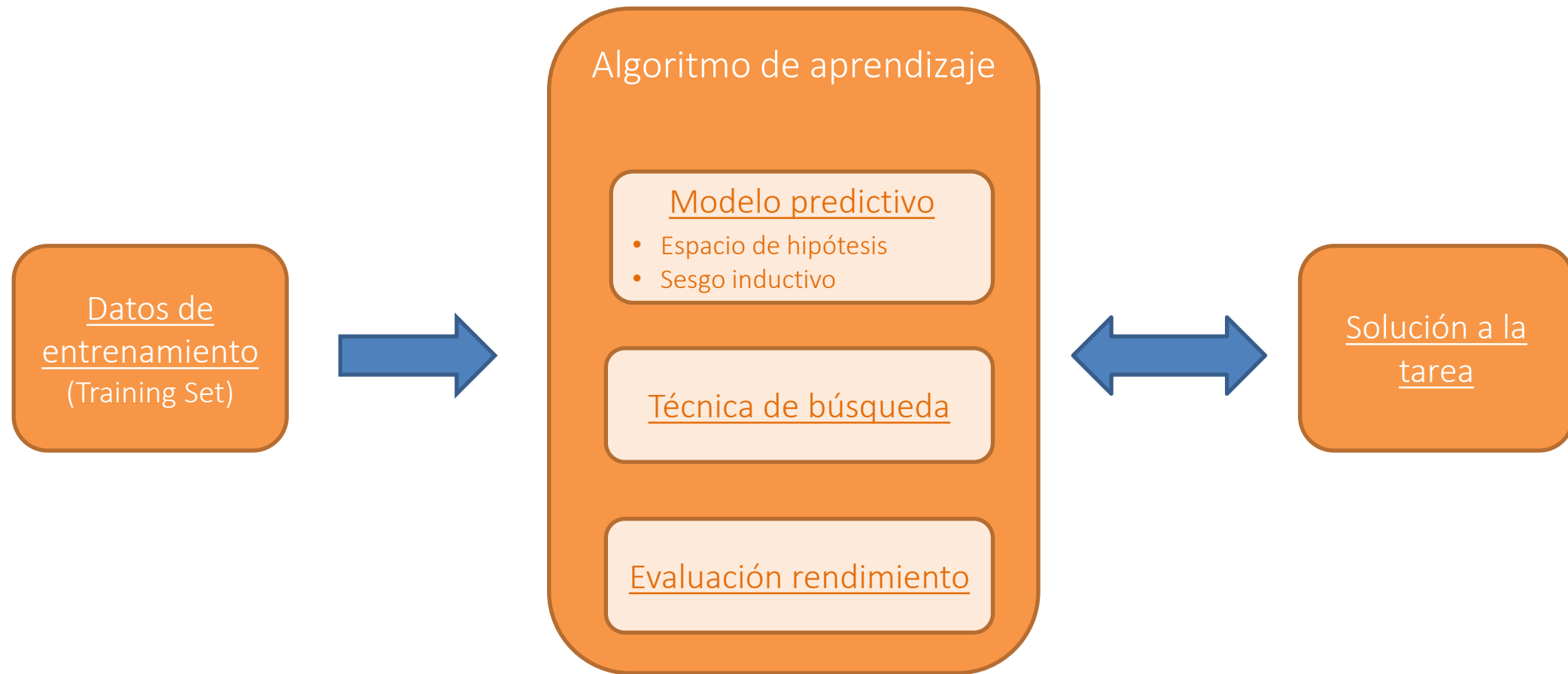
Modelo predictivo: función parametrizada, que mapea desde el espacio de los datos hacia el de las etiquetas

Buscamos, en el **espacio de hipótesis** del modelo, aquella que genera la solución con mejor **rendimiento**, usando el **sesgo inductivo** y los datos de **entrenamiento** para guiar la **búsqueda**.

Ok, súper lindo, pero ¿cómo funciona?

¿Cómo podemos construir un detector de vehículos?

¿Cómo podemos construir un detector de vehículos?

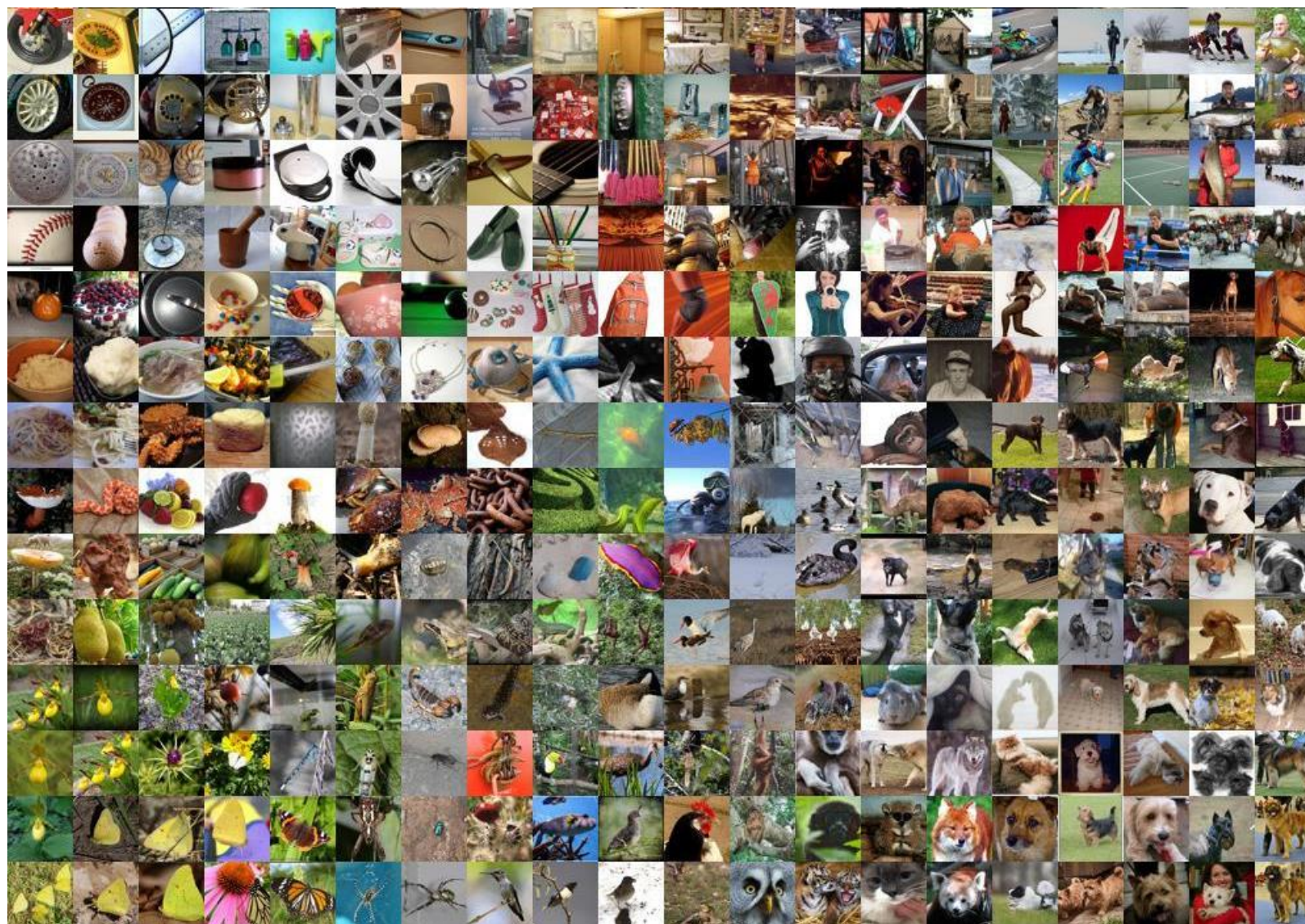


¿Qué es lo primero que necesitamos?

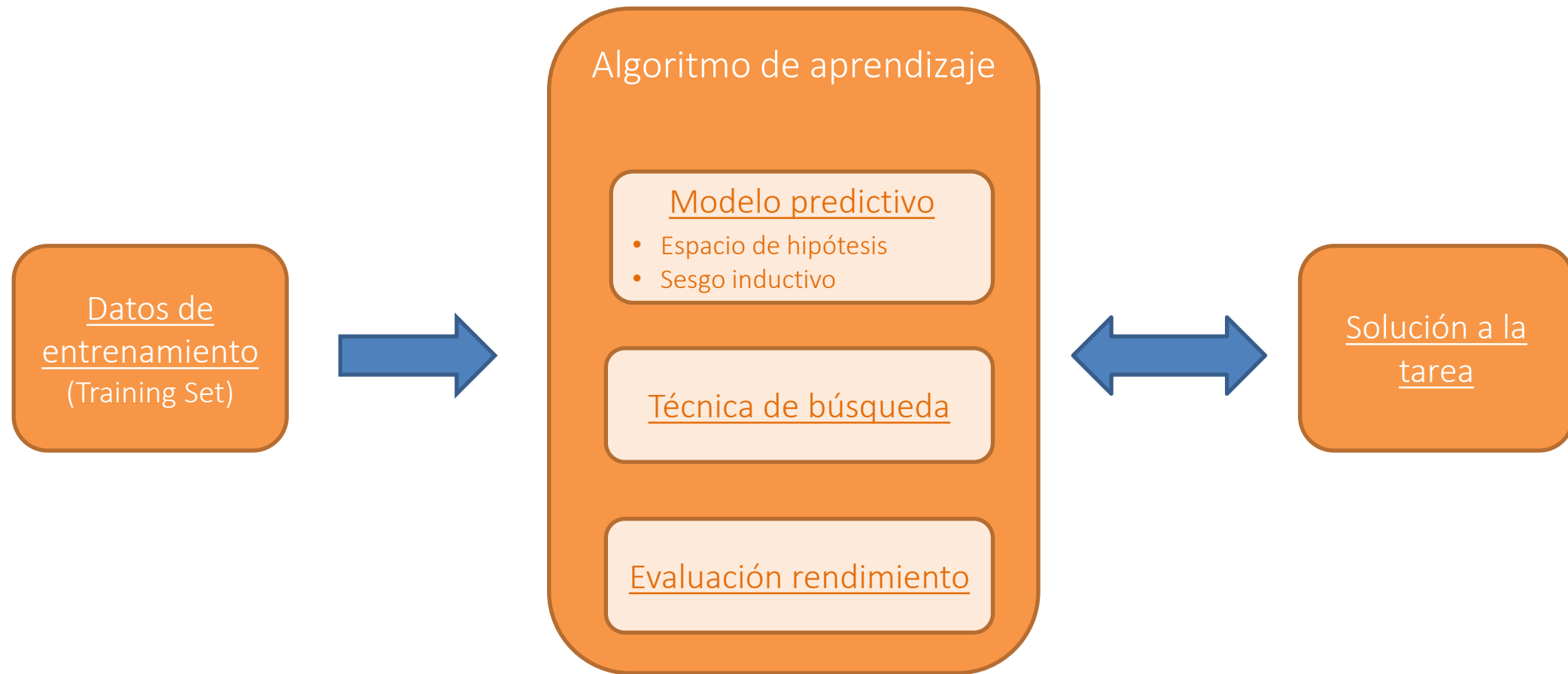
Algoritmo de aprendizaje: estos SON vehículos



Algoritmo de aprendizaje: estos **NO SON** vehículos

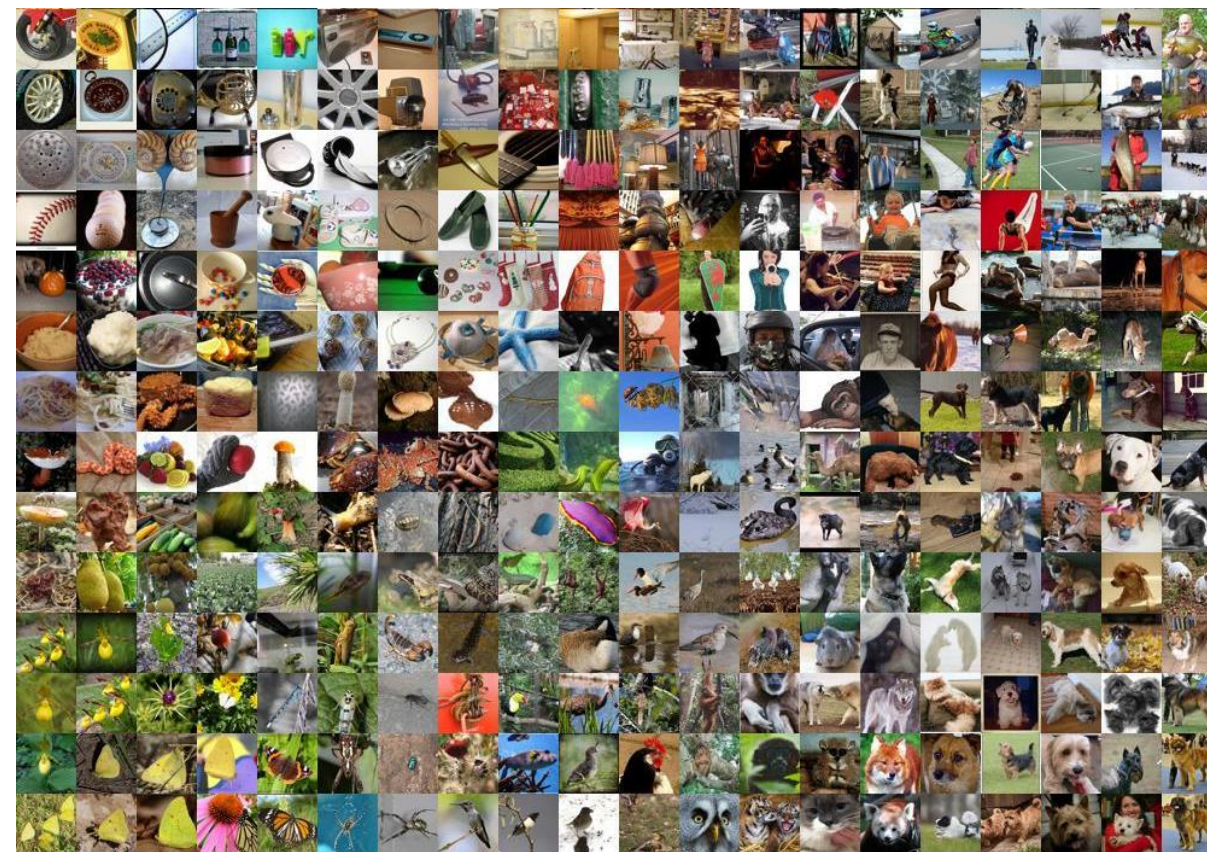


¿Cómo podemos construir un detector de vehículos?

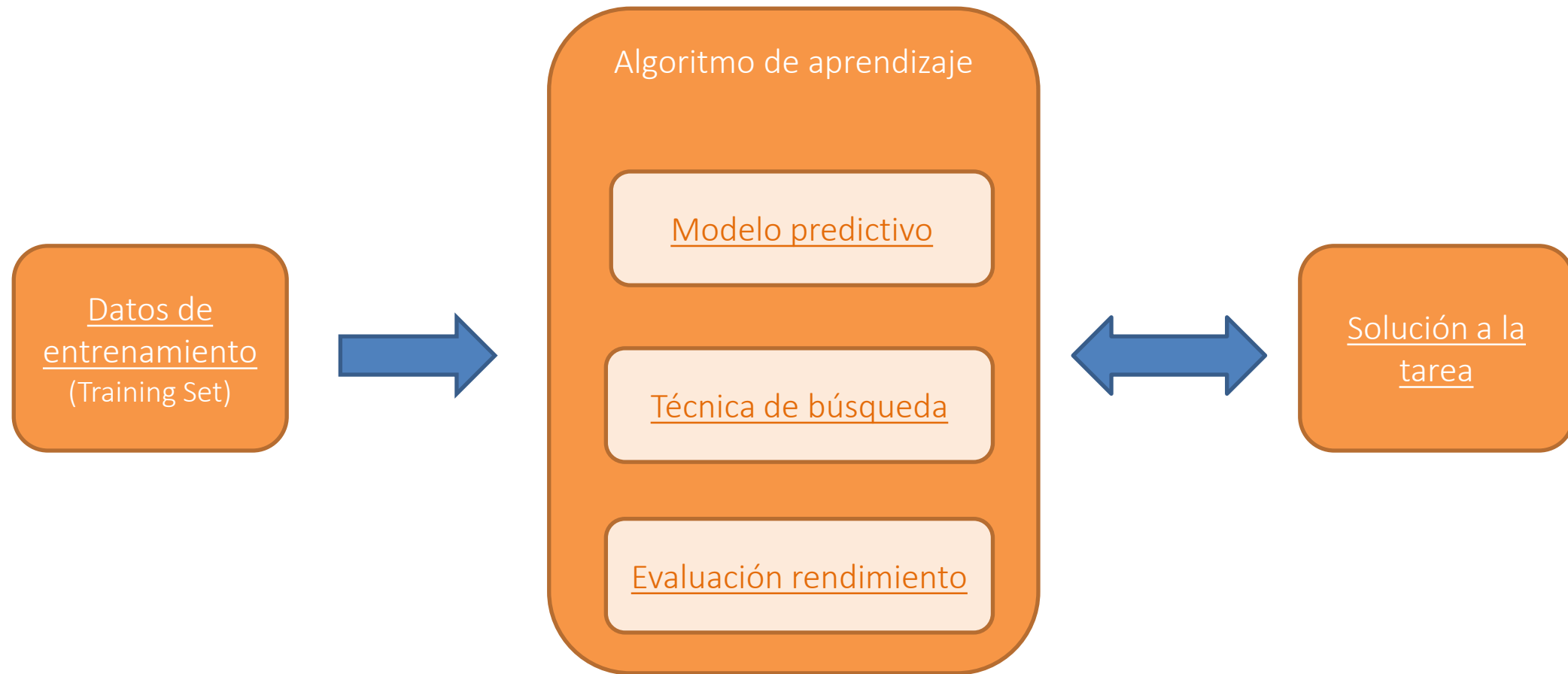


Dejando para más adelante en el curso el modelo y la búsqueda,
¿cómo podemos evaluar el rendimiento?

Algoritmo de aprendizaje: ¿cuáles son vehículos y cuáles no?



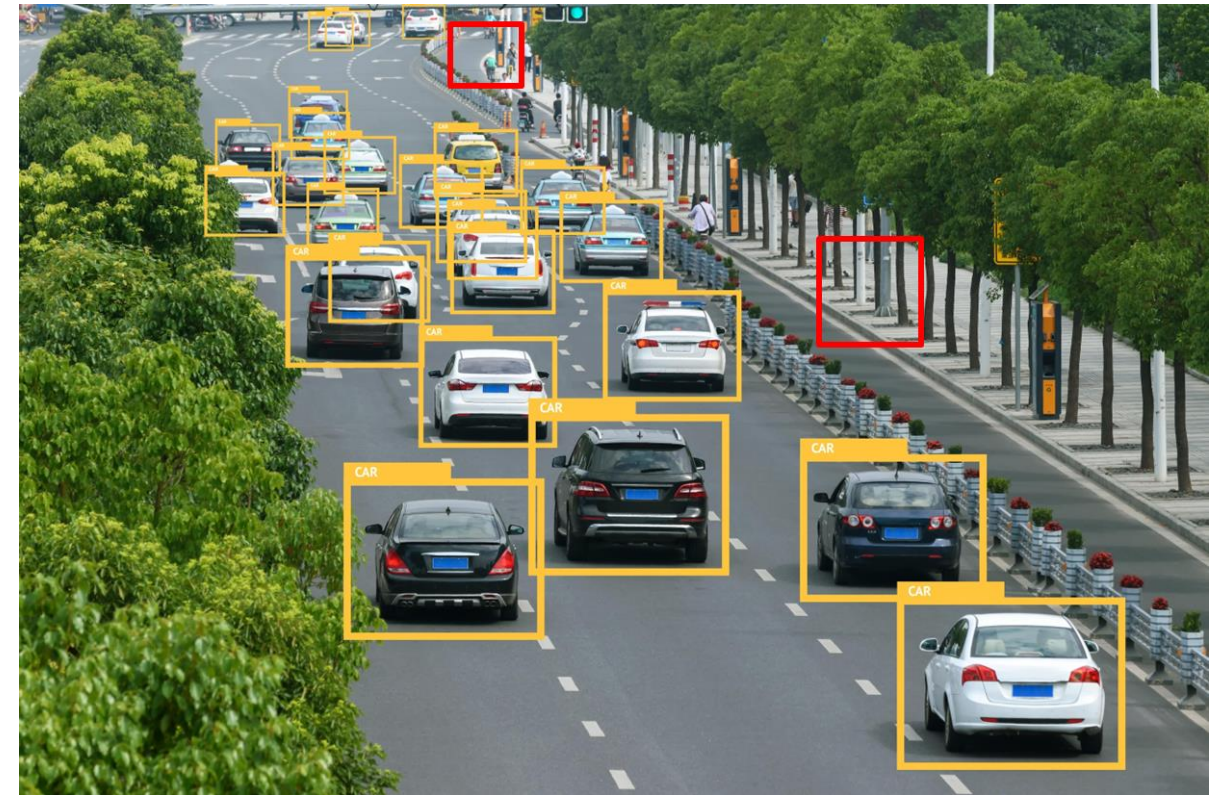
¿Cómo podemos construir un detector de vehículos?



¿Hemos terminado?

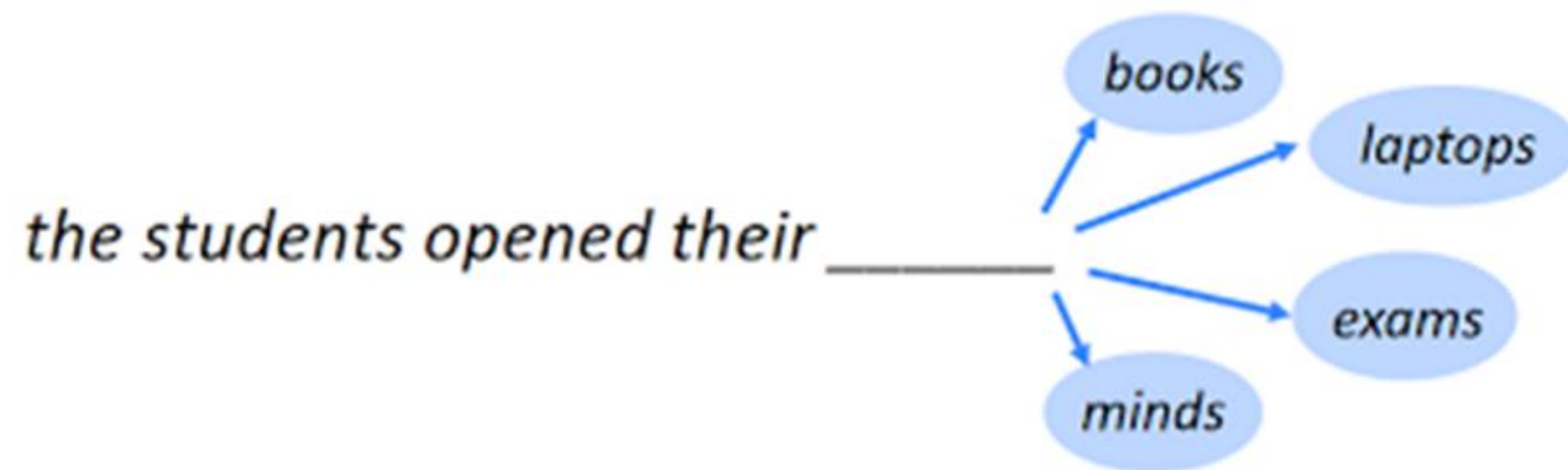
Un sistema de ML no debe ser visto como un producto estático

- Tal como cualquier sistema complejo, un sistema de ML requiere de “mantenciones”.
- A diferencia de un producto físico, estas no solo pueden ser en el hardware, sino que también a los **datos**.
- En el caso de un detector de vehículos, es fundamental un proceso **continuo** de **validación** y actualización de **rendimiento** y **datos**.



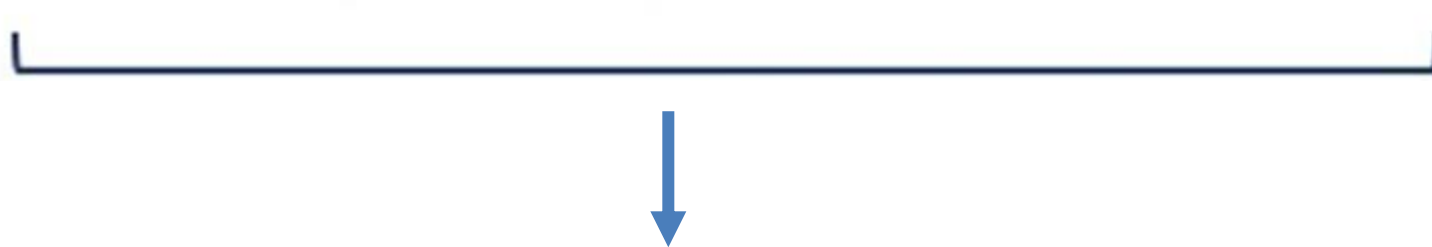
Otro ejemplo más, ¿cómo podemos construir un generador de texto?
(estirando algo el chicle, un precursor de ChatGPT)

Un interludio: modelos de lenguaje



Generación de texto mediante modelos de lenguaje

Can you please come **here**

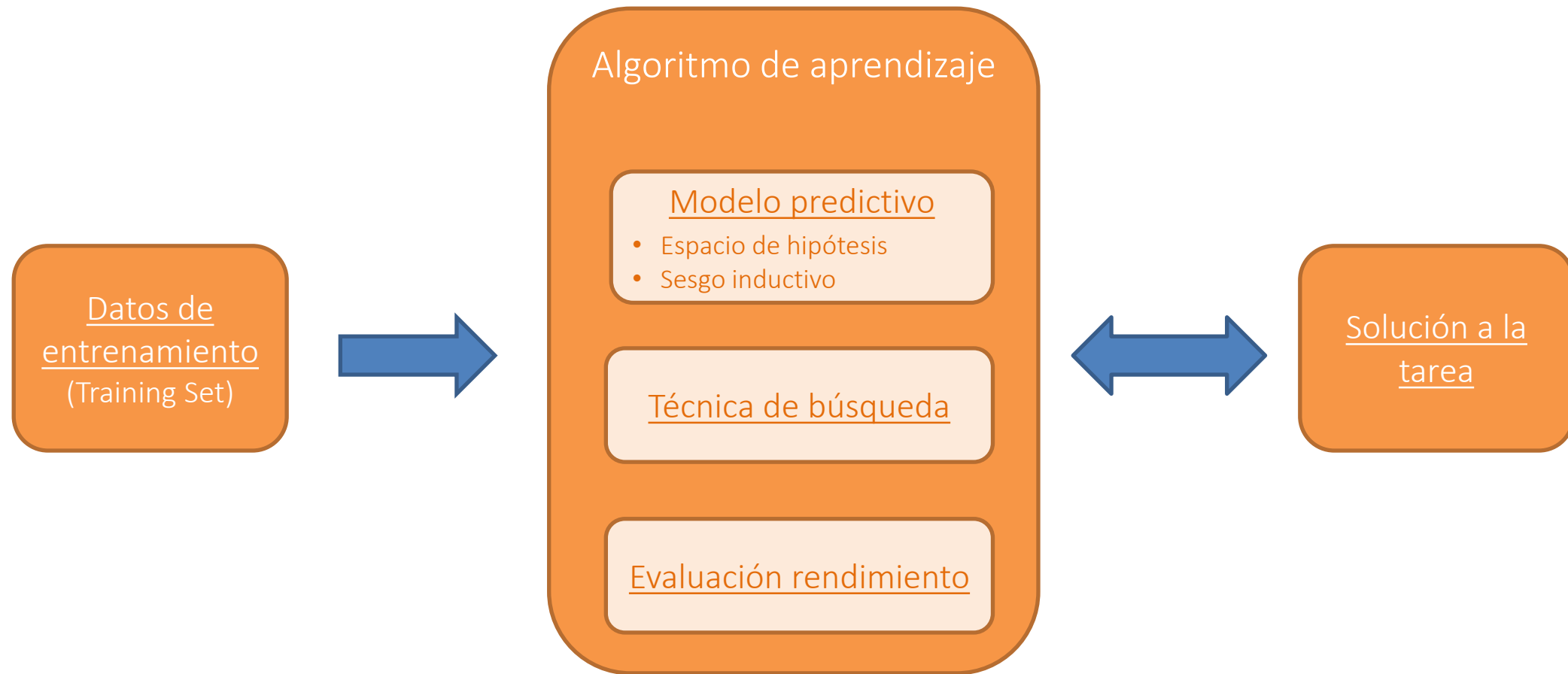


Entrada al modelo
de lenguaje



Salida del modelo de
lenguaje (mediante
muestreo o la de
mayor probabilidad)

¿Cómo podemos construir un generador de texto?



¿Qué es lo primero que necesitamos?

Algoritmo de aprendizaje: estos son los
datos y lo que debes hacer con ellos

Text Corpus

Nothing is impossible.
Even the word
impossible
says I'm possible



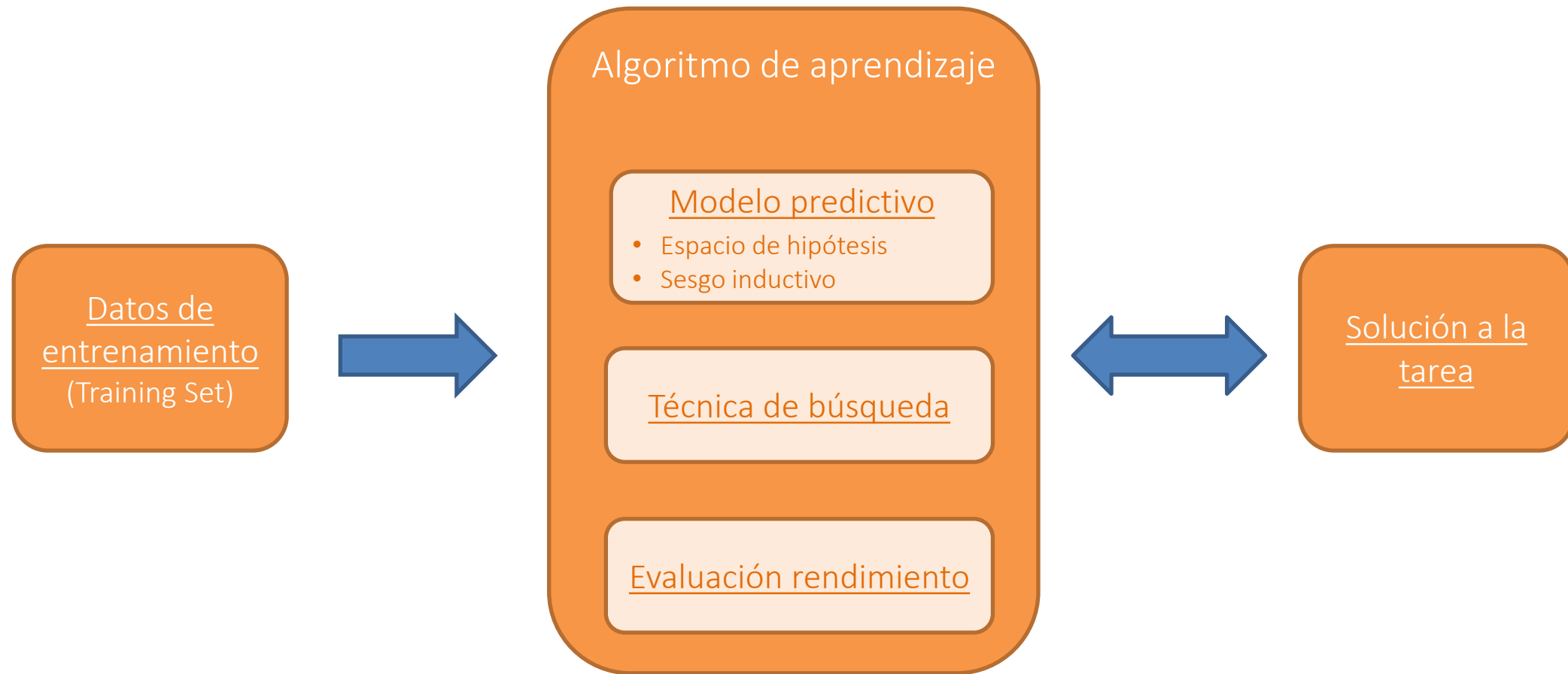
Task: Predict from past

Nothing
Nothing is
Nothing is impossible
...

Veamos un último ejemplo

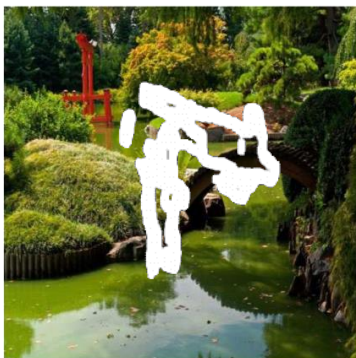


Siempre lo primero serán los datos de entrenamiento



¿Cómo generamos el conjunto de datos de entrenamiento?





Modelo
predictivo



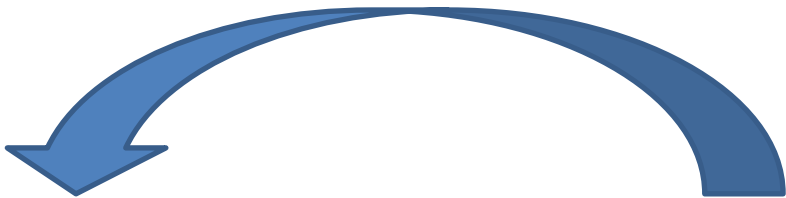
?



=



?



Veamos como funciona el sistema en la práctica
(<https://youtu.be/gg0F5JjKmhA>)



Pontificia Universidad Católica de Chile
Escuela de Ingeniería
Departamento de Ciencia de la Computación



IIC2613 - Inteligencia Artificial

Introducción a Machine Learning

Hans Löbel

Dpto. Ingeniería de Transporte y Logística
Dpto. Ciencia de la Computación