

TDVI: Inteligencia Artificial

Formalidades, Intro a IA & Intro a ML

UTDT - LTD



Estructura de la clase

- Formalidades de la materia
- ¿Qué es la inteligencia artificial?
- ¿Qué es el aprendizaje automático?
- Noción de aprendizaje supervisado
- Un ejercicio ilustrativo

Estructura de la clase

- Formalidades de la materia
- ¿Qué es la inteligencia artificial?
- ¿Qué es el aprendizaje automático?
- Noción de aprendizaje supervisado
- Un ejercicio ilustrativo

Formalidades de la materia

Cuerpo docente

Teóricas:

- Ramiro Gálvez (1^{ra} mitad del curso, ambas secciones)
- Viviana Siless (2^{da} mitad del curso, ambas secciones)

Prácticas:

Sección 1

- Paula Feldman
- Federico Giorgi

Sección 2

- Ignacio Pardo
- Martin Sinnona

La comunicación será a través del campus

Formalidades de la materia

Estructura general de la materia

Hasta el primer parcial:

- Aprendizaje automático “clásico”
- Principalmente aprendizaje supervisado (veremos algo de no supervisado y self-supervised learning también)
- Fuerte énfasis en seguir un pipeline adecuado de predicción

A partir del segundo parcial:

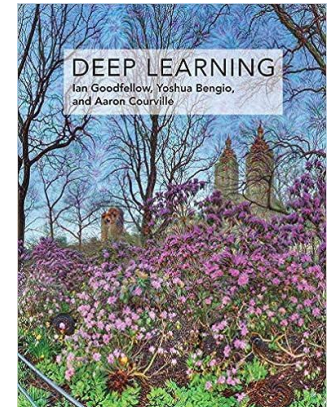
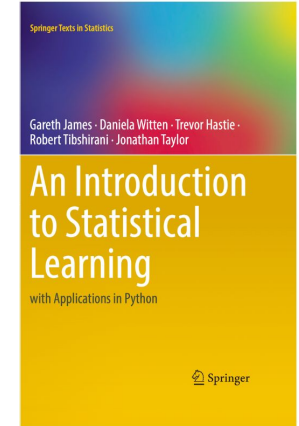
- Deep learning:
 - Perceptron
 - Perceptrón multicapa y feedforward networks
 - Redes recurrentes
 - Redes convolucionales
 - Generative adversarial Networks (GANs)
 - Encoders-decoders
 - Transformers

Formalidades de la materia

Bibliografía

Principal:

- James, Witten, Hastie & Tibshirani, "[An Introduction to Statistical Learning with Applications in Python](#)", Springer, 2023 (ISLP)
- Ian Goodfellow, Yoshua Bengio, and Aaron Courville, "[Deep learning](#)", The MIT Press, 2016



Complementaria / Avanzada:

- Capítulos sueltos de otros libros (se listarán al final de cada clase)
- Papers que se listarán al final de las sesiones en donde se vean

Formalidades de la materia

Régimen de aprobación

- Se rendirán dos exámenes parciales individuales, y se deberán resolver cuatro trabajos prácticos grupales (TPs), y eventuales recuperatorios
- Para aprobar es necesario obtener una nota mayor o igual a 60 (sobre un total de 100) en los dos parciales y en los cuatro TPs
- Quien repruebe un examen parcial en primera instancia, deberá aprobarlo en un recuperatorio específico para esa instancia
- Quien repruebe un TP en primera instancia, podrá reentregarlo/recuperarlo en un plazo estipulado. En este caso la nota final obtenida en el TP no superará los 75 puntos. En caso de recuperar un TP, los docentes podrán solicitar un coloquio de defensa en caso de considerarlo necesario
- La nota del recuperatorio reemplazará a la nota del parcial o del TP correspondiente. Quien no apruebe un recuperatorio, reprobará la materia
- Se podrán recuperar un máximo de tres (3) instancias de evaluación. Quien repruebe cuatro evaluaciones, reprobará la materia
- Está permitido rendir un recuperatorio para intentar mejorar la nota, sin embargo la nota del recuperatorio será la definitiva

IMPORTANTE: vean el programa para más detalles

Formalidades de la materia

Herramientas

Principalmente Python + Pandas + Scikit learn (se recomienda que usen la distribución [Anaconda](#)). A partir de la segunda mitad se suma Pytorch



Para el primer TP deberán utilizar R + Rstudio (ya veremos el motivo)

Aclaraciones:

- Siempre utilizaremos software libre
- Computadoras personales estándar + Google Colab debieran ser suficientes para cursar la materia

Estructura de la clase

- Formalidades de la materia
- ¿Qué es la inteligencia artificial?
- ¿Qué es el aprendizaje automático?
- Noción de aprendizaje supervisado
- Un ejercicio ilustrativo

¿Qué es la Inteligencia Artificial?

Nos llamamos a nosotros mismos **homo sapiens** (el hombre sabio), por la importancia que otorgamos a nuestra inteligencia

Por cientos de años hemos intentado entender cómo pensamos y actuamos

La Inteligencia Artificial (IA) como disciplina se ocupa no sólo de comprender cómo pensamos, sino también de **construir “entidades inteligentes”**

¿Cómo caracterizar a una entidad como “inteligente”?

Esta pregunta no tiene una respuesta simple (¿una calculadora es inteligente?)

¿Qué es la Inteligencia Artificial?

¿Qué se entiende por agentes inteligentes?

¿En dónde se manifiesta?

- Pensamiento: la inteligencia como una propiedad de los procesos y razonamientos internos del pensamiento (agentes que razonan “bien”)
- Comportamiento: la inteligencia asociada a un comportamiento “inteligente” (agentes que se comportan “bien”)

¿Qué significa “bien”?

- Agentes que replican a los humanos
- Agentes racionales

Cuatro combinaciones posibles:

	Humano	Racional
Pensamiento		
Comportamiento		

¿Qué es la Inteligencia Artificial?

	Humano	Racional
Pensamiento		
Comportamiento	X	

Actuar como un humano (*the Turing test approach*)

Test de Turing (1950): un programa pasa el test si un interrogador humano, después de plantear algunas preguntas por escrito, **no puede distinguir si las respuestas escritas provienen de una persona o de una computadora** (existen versiones “multimodales”)

Fue diseñado como un experimento mental que evita la vaguedad filosófica de la pregunta "¿Puede una máquina pensar?"

Múltiples ejemplos en la cultura popular: Ex Machina, Blade Runner, Alien, etc.

¿Qué es la Inteligencia Artificial?

	Humano	Racional
Pensamiento	X	
Comportamiento		

Pensar como un humano (*the cognitive modelling approach*)

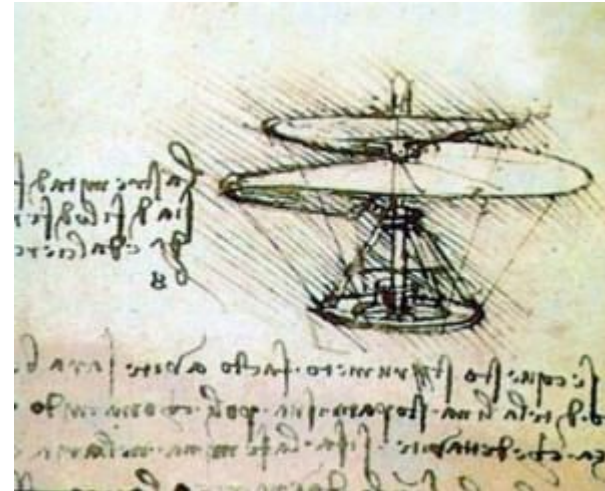
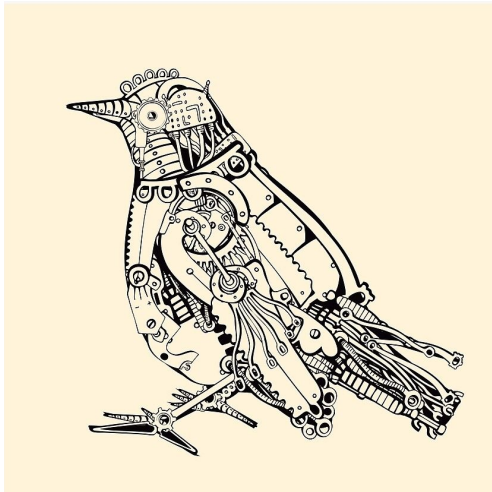
Primero se debe **entender cómo piensan/procesan información** los humanos. A través de 1) la introspección, 2) experimentos psicológicos y 3) brain imaging

Teniendo una teoría desarrollada, se la podría expresar mediante programas. Si el input-output de estos programas se comporta de acuerdo a lo que sucede en humanos, indicaría que el mecanismo propuesto opera en humanos

Fuerte relación con las **ciencias cognitivas**. Ejemplo: *computer vision* incorpora mucho conocimiento de cómo los humanos procesan información visual (el cual proviene de las ciencias cognitivas)

¿Qué es la Inteligencia Artificial?

¿Sólo se puede volar imitando a los pájaros?

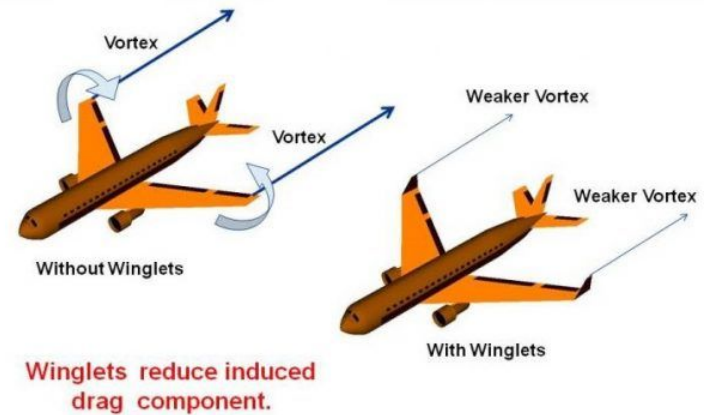


El "vuelo artificial" tuvo éxito cuando los ingenieros e inventores **dejaron de imitar a los pájaros** y comenzaron a utilizar túneles de viento y a aprender sobre aerodinámica

Los textos de ingeniería aeronáutica no definen el objetivo de su campo como hacer **máquinas que vuelan exactamente como palomas** (incluso engañando incluso a las mismas)

¿Qué es la Inteligencia Artificial?

¿Pero, no tendrá sentido inspirarse en la naturaleza?



The idea was refined for modern aircraft in the 1970s by NASA engineer Richard Whitcomb, who imagined vertical wing extensions inspired by the way birds curl up the end of their wings when in need of lift.

Links:

<https://edition.cnn.com/travel/article/airplane-winglets-cmd/index.html>

<https://appel.nasa.gov/2014/07/22/this-month-in-nasa-history-winglets-helped-save-an-industry/>

¿Qué es la Inteligencia Artificial?

	Humano	Racional
Pensamiento		X
Comportamiento		

Pensamiento racional (*the “laws of thought” approach*)

“*Sócrates es un humano*” + “*los humanos son mortales*”, entonces “*Sócrates es mortal*”

La idea inicial es desarrollar programas que puedan **resolver todo problema resoluble y describible en notación lógica**, y así crear inteligencia artificial

La lógica trabaja con conocimiento certero, la teoría de **la probabilidad permite trabajar con incertidumbre**. En principio, permitiría modelar el pensamiento humano

Este enfoque está puesto en realizar **inferencias correctas**, no “comportamiento inteligente”

¿Qué es la Inteligencia Artificial?

	Humano	Racional
Pensamiento		
Comportamiento		X

Comportamiento racional (*the rational agent approach*)

Un agente racional es uno que interactúa con el ambiente para alcanzar **el “mejor” resultado** (o el mejor resultado esperado). IMPORTANTE: hay un objetivo

No se basa necesariamente en inferencias racionales (e.g., nosotros no sacamos la mano del fuego por razonamiento)

Este enfoque tiene **dos ventajas**:

1. Inferencia racional es una de varias estrategias posibles para obtener racionalidad
2. La racionalidad está correctamente definida en términos matemáticos, lo que la vuelve adecuada para experimentar

¿Qué es la Inteligencia Artificial?

Modelo estándar (Russell & Norvig):

El enfoque del agente racional **ha prevalecido en la historia de la IA**. Inicialmente los agentes eran guiados por lógica o fuerza bruta, hoy por aprendizaje automático

“AI has focused on the study and design of agents that do the right thing” ([nota a Russell](#))

- Qué es “lo correcto” (la función objetivo) **debe ser provisto** al agente (e.g., maximizar el puntaje en juegos de Atari)
- Se llama *value alignment problem* a lograr un acuerdo entre nuestras verdaderas preferencias y el objetivo que introducimos en los agentes. Los objetivos introducidos deben estar alineados con los de los seres humanos
- A medida que uno se mueve del laboratorio al “mundo real”, se vuelve **más difícil definir estos objetivos** (el Rey Midas). Sumamente importante hoy con sistemas tipo ChatGPT, Claude, Gemini, etc.

¿Qué es la Inteligencia Artificial?

La IA como disciplina incorpora conocimientos de otras disciplinas:

- **Filosofía:** la mente como una máquina
- **Matemática:** lógica formal, teoría de la probabilidad, estadística, computación, etc.
- **Economía:** decisiones acordes a preferencias, teoría de juegos, etc.
- **Neurociencias / Ciencias Cognitivas:** cómo el cerebro procesa información
- **Psicología:** cómo los humanos piensan y actúan
- **Ingeniería Computacional:** cómo crear estas máquinas
- **Teoría del control / Cibernética:** agentes/artefactos que operan bajo su propio control
- **Lingüística:** relación entre lenguaje y pensamiento

¿Qué es la Inteligencia Artificial?

Muy breve **historia de la IA** como disciplina

- 1943 - 1956. Nacimiento de la IA (e.g., Test de Turing, Conferencia de Dartmouth, etc.)
- 1952 - 1969. Entusiasmo temprano y grandes expectativas (e.g, demostradores de teoremas, jugadores de damas, perceptrón, etc.)
- 1966 - 1973. **Una dosis de realidad**. Problemas: 1) falta de rigurosidad al analizar los problemas, 2) Intractabilidad de obtener soluciones por fuerza bruta
- 1969 - 1986. **Sistemas expertos**. A través de consultas con expertos se desarrollan listas de reglas que replican su accionar. Inicialmente un éxito comercial, luego **Winter AI**
- **1986 - presente**. El retorno de las redes neuronales
- **1987 - presente**. Razonamiento probabilístico y aprendizaje automático
- **2001 - presente**. Big data
- **2011 - presente**. Deep learning

El campo pasó de la lógica booleana al razonamiento probabilístico, del conocimiento “codeado” a mano al aprendizaje automático a partir de datos

Estructura de la clase

- Formalidades de la materia
- ¿Qué es la inteligencia artificial?
- ¿Qué es el aprendizaje automático?
- Noción de aprendizaje supervisado
- Un ejercicio ilustrativo

¿Qué es el Aprendizaje Automático?

Hay problemas que pueden resolverse/atacarse con programación “tradicional” (más o menos compleja):

- ¿N es múltiplo de 2023?
- ¿N es primo?
- Encontrar el máximo en una lista
- Ubicar 8 reinas en un tablero y que no se ataquen
- Ubicar antenas de radio para lograr que no haya interferencias

Otros casos resultan más **difíciles** de atacar:

¿Cuáles de estos son objetos sillas?



¿Qué es el Aprendizaje Automático?

¿Cómo lo podremos atacar?



Opción 1

- Llamar a un carpintero (**un experto**), preguntarle qué son las sillas y codificar esa información en una secuencia de reglas. Esto es ejemplo de un **sistema experto**. No va a escalar bien

Opción 2

- Recolectar ejemplos de objetos que sean sillas y de objetos que no lo sean. Utilizar un algoritmo que “automáticamente” descubra qué características tienen y no tienen las sillas. Esto es un ejemplo de **aprendizaje automático**

¿Qué es el Aprendizaje Automático?

Una definición de algoritmo que “aprende automáticamente”

“A computer program is said to learn from experience E with respect to some class of tasks T and performance measure P , if its performance at tasks in T , as measured by P , improves with experience E ” (Mitchell, 1997)

Ejemplo: reconocimiento de caracteres escritos (OCR)

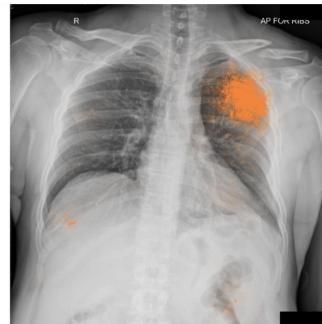
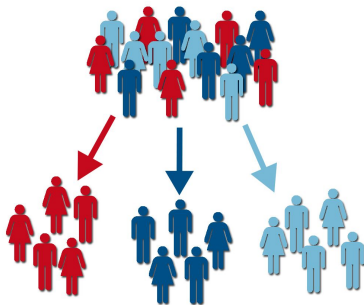
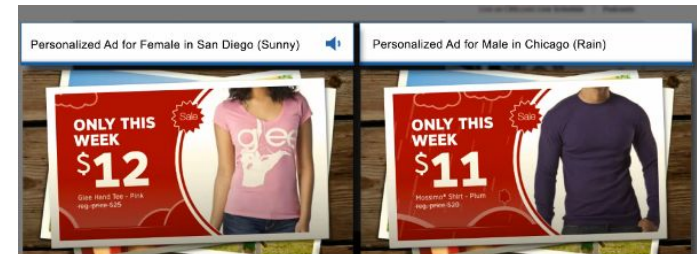
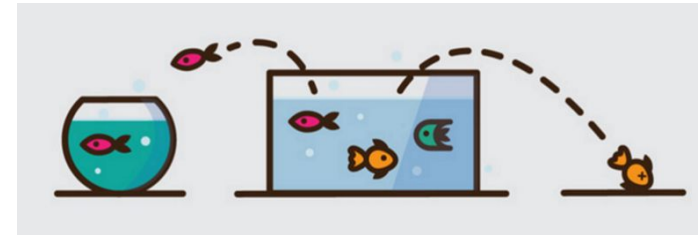
- T : reconocer caracteres escritos dentro de imágenes
- P : porcentaje de caracteres correctamente reconocidos
- E : una colección de imágenes con caracteres escritos para los cuales sabemos qué caracteres contienen

¿Cómo podría ser para un detector de correo spam?

¿Qué es el Aprendizaje Automático?

Ejemplos:

- Churn (attrition) de clientes
- Segmentación de clientes
- Recomendación de productos
- Publicidades personalizadas
- Credit scoring
- Predicción de valor de un cliente
- Diagnóstico con imágenes médicas
- Análisis de sentimiento
- Detectores de spam



Everything is a Recommendation

Ranking

Rows

Over 80% of what members watch comes from our recommendations

Recommendations are driven by Machine Learning Algorithms

NETFLIX

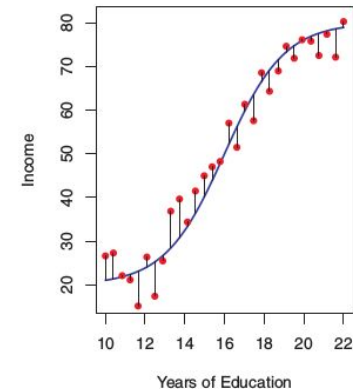
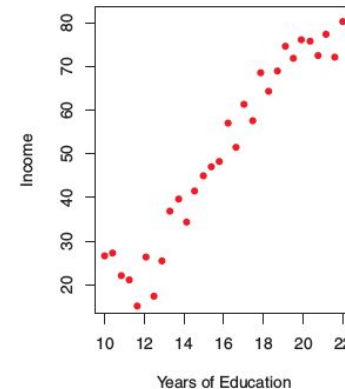
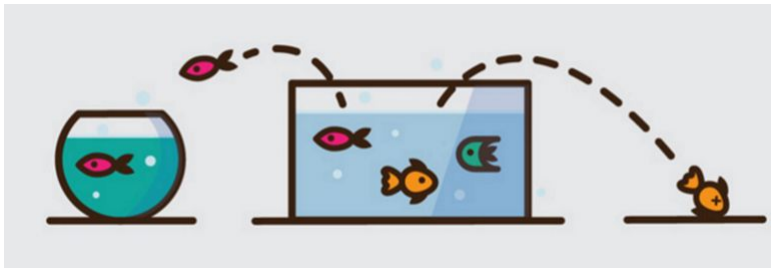
CASSANDRA SUMMIT 2016

¿Qué es el Aprendizaje Automático?

En términos generales, los algoritmos de aprendizaje automático pertenecen a tres grandes familias (supervisado, no supervisado, por refuerzos).

1. **Supervisado** (principalmente vamos a abordar este tipo):

"For each observation of the predictor measurement(s) x_i , $i = 1, \dots, n$ there is an associated response measurement y_i "

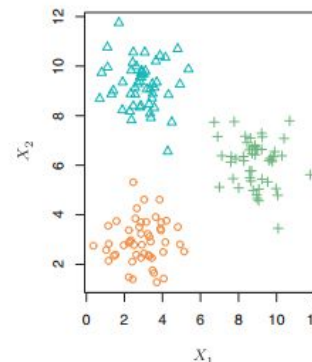
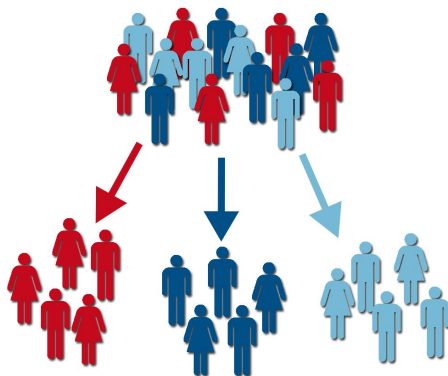


¿Qué es el Aprendizaje Automático?

En términos generales, los algoritmos de aprendizaje automático pertenecen a tres grandes familias (supervisado, no supervisado, por refuerzos).

2. No supervisado (vamos a abordarlo en menor medida):

*"Unsupervised learning describes the somewhat **more challenging situation** in which for every observation $i = 1, \dots, n$, we observe a vector of measurements x_i but **no associated response** y_i ... We can seek to understand the relationships between the variables or between the observations"*

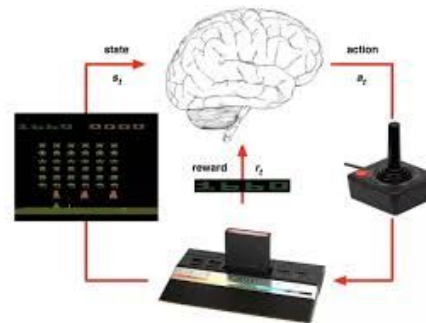
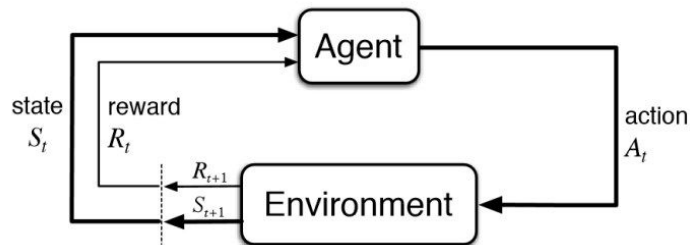


¿Qué es el Aprendizaje Automático?

En términos generales, los algoritmos de aprendizaje automático pertenecen a tres grandes familias (supervisado, no supervisado, por refuerzos).

3. **Por refuerzos** (no la vamos a abordar):

Se enfoca en desarrollar agentes que maximicen un beneficio esperado mediante la interacción con un ambiente (el cual muchas veces no conocen)

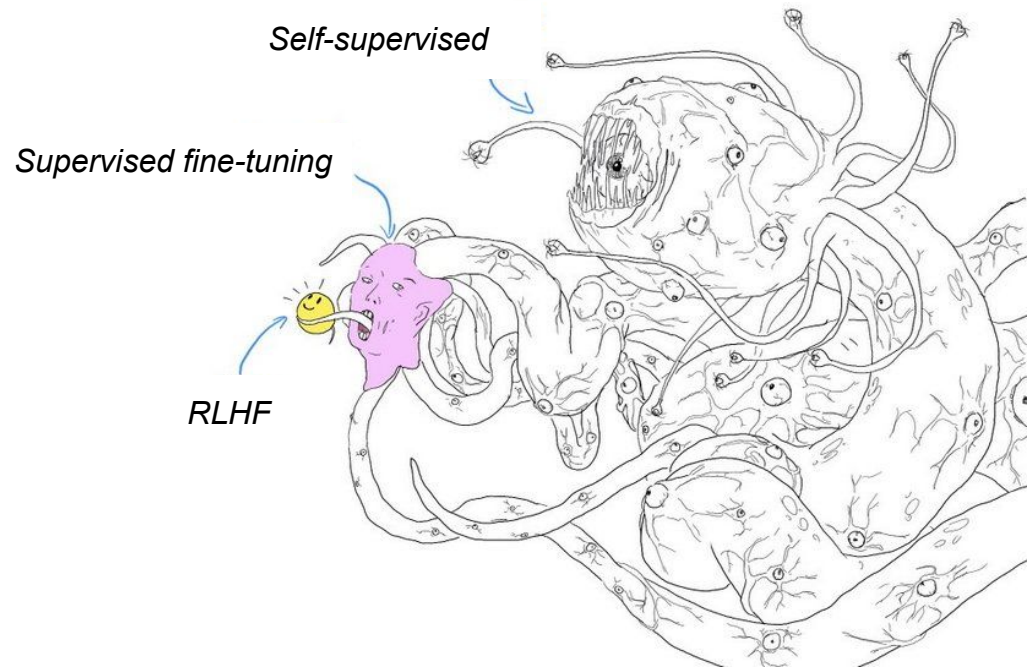


¿Qué es el Aprendizaje Automático?

La literatura generalmente clasifica a los algoritmos de aprendizaje en estas tres familias (supervisados, no supervisados, por refuerzos). Pero existen modelos que **no caen en exclusivamente una de estas categorías**

Por ejemplo:

- Semi-supervised learning
- Self-supervised learning
- Transfer learning



(meme de cómo se entrenó ChatGPT)

Estructura de la clase

- Formalidades de la materia
- ¿Qué es la inteligencia artificial?
- ¿Qué es el aprendizaje automático?
- **Noción de aprendizaje supervisado**
- Un ejercicio ilustrativo

Noción de Aprendizaje supervisado

Lo que caracteriza al aprendizaje supervisado es que se quiere **predecir una variable concreta**. En principio, hay dos tipos de variables a predecir:

Continua → regresión

$$y = f(x) + \varepsilon$$

Categorica → clasificación (binaria, multiclase)

(Igualmente, hay **otros tipos de problemas de aprendizaje supervisado**, por ejemplo: ranking - usado en buscadores, carreras de caballos, etc.)

¿Qué tipo de problema son los siguientes?

- Predecir la cantidad de ventas de líneas telefónicas en un día dado
- Predecir si un cliente en particular se dará de alta dada nuestra campaña de marketing
- Predecir si un día determinado tendremos o no más de 200 líneas nuevas vendidas

Noción de Aprendizaje supervisado

Esquema de la estrategia. Generalmente tiene **dos etapas**:

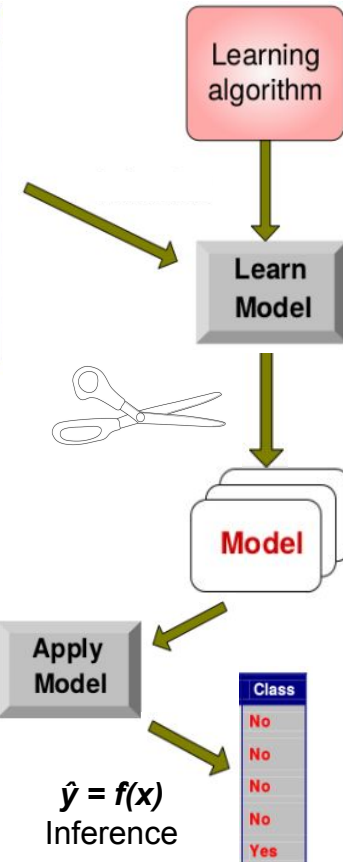
1) Aprendizaje / Entrenamiento

Tid	Attrib1	Attrib2	Attrib3	Class
1	Yes	Large	125K	No
2	No	Medium	100K	No
3	No	Small	70K	No
4	Yes	Medium	120K	No
5	No	Large	95K	Yes
6	No	Medium	60K	No
7	Yes	Large	220K	No
8	No	Small	85K	Yes
9	No	Medium	75K	No
10	No	Small	90K	Yes

Training Set

Tid	Attrib1	Attrib2	Attrib3
11	No	Small	55K
12	Yes	Medium	80K
13	Yes	Large	110K
14	No	Small	95K
15	No	Large	67K

Test Set



Predecir bien no es trivial (hay muchas formas de hacerlo mal). Veremos en detalle cómo hacerlo bien

Noción de Aprendizaje supervisado

¿Qué queremos que “arroje” el modelo en cada caso?

Regresión

- Queremos un valor predicho de y_i (i.e., \hat{y}_i) que sea lo más cercano posible a lo que efectivamente valdrá y para la observación i

Clasificación

- Idealmente, los clasificadores devolverán $P(y_i = k \mid X_i)$ para todo $k \in K$
- Es decir, una estimación de la probabilidad de que la variable y tome el valor k para una observación i con atributos X_i
- Queremos que si y_i es igual a k , entonces $P(y_i = k \mid X_i)$ sea lo más cercana a 1 (y, por ende, lo más cercana a 0 para las restantes clases)
- Sobre la base de los valores $P(y_i = k \mid X_i)$ para todo $k \in K$ decidiremos qué clase predecir para la observación i

Estructura de la clase

- Formalidades de la materia
- ¿Qué es la inteligencia artificial?
- ¿Qué es el aprendizaje automático?
- Noción de aprendizaje supervisado
- **Un ejercicio ilustrativo**

Un ejercicio ilustrativo

Animal	¿Da a luz?	¿Vuela?	¿Vive en el agua?	¿Tiene piernas?	¿Mamífero?
humano	sí	no	no	sí	mamífero
pitón	no	no	no	no	no-mamífero
salmón	no	no	sí	no	no-mamífero
ballena	sí	no	sí	no	mamífero
rana	no	no	a veces	sí	no-mamífero
dragón de komodo	no	no	no	sí	no-mamífero
murciélago	sí	sí	no	sí	mamífero
paloma	no	sí	no	sí	no-mamífero
gato	sí	no	no	sí	mamífero
tiburón leopardo	sí	no	sí	no	no-mamífero
tortuga	no	no	a veces	sí	no-mamífero
pingüino	no	no	a veces	sí	no-mamífero
puercoespín	sí	no	no	sí	mamífero
anguila	no	no	sí	no	no-mamífero
salamandra	no	no	a veces	sí	no-mamífero
monstruo de gila	no	no	no	sí	no-mamífero
ornitorrinco	no	no	no	sí	mamífero
búho	no	sí	no	sí	no-mamífero
delfín	sí	no	sí	no	mamífero
águila	no	sí	no	sí	no-mamífero

¿Da a luz?	¿Vuela?	¿Vive en el agua?	¿Tiene piernas?	¿Mamífero?
sí	no	sí	no	¿?

¿La observación faltante será o no un mamífero?

¿Sería un problema tener únicamente ejemplos de mamíferos?

Bibliografía

Básica

- Russell & Norvig, “Artificial Intelligence: A Modern Approach”, 4th ed. Capítulo 1
- Mitchell, “Machine Learning”, Capítulo 1, hasta la sección 1.1 inclusive (obviar el resto del capítulo)
- ISLP. Secciones 2.1.4 y 2.1.5