

**PARTE I**

**CLASE**

**1**

# MÓDULO 01 - CLASE 01

## INTRODUCCIÓN A LOS MODELOS IA (I)

.UBA económicas  
*Diplomaturas*

IA aplicada a  
Entornos Digitales  
de Gestión



# 1940 - 1950

## Nacen los gemelos



# 1943

## Primera aproximación a la neurona artificial.

En 1943, los neurocientíficos Warren McCulloch y Walter Pitts publicaron un artículo fundamental titulado "A Logical Calculus of Ideas Immanent in Nervous Activity". Introducieron el concepto de neurona artificial y sentaron las bases para el desarrollo de las redes neuronales.

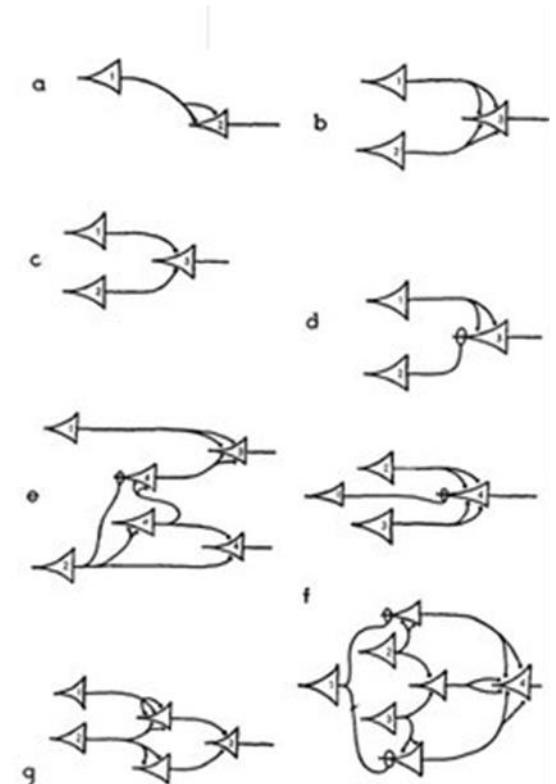
BULLETIN OF  
MATHEMATICAL BIOPHYSICS  
VOLUME 5, 1943

### A LOGICAL CALCULUS OF THE IDEAS IMMANENT IN NERVOUS ACTIVITY

WARREN S. McCULLOCH AND WALTER PITTS

FROM THE UNIVERSITY OF ILLINOIS, COLLEGE OF MEDICINE,  
DEPARTMENT OF PSYCHIATRY AT THE ILLINOIS NEUROPSYCHIATRIC INSTITUTE,  
AND THE UNIVERSITY OF CHICAGO

Because of the "all-or-none" character of nervous activity, neural events and the relations among them can be treated by means of propositional logic. It is found that the behavior of every net can be described in these terms, with the addition of more complicated logical means for nets containing circles; and that for any logical expression satisfying certain conditions, one can find a net behaving in the fashion it describes. It is shown that many particular choices among possible neurophysiological assumptions are equivalent, in the sense that for every net behaving under one assumption, there exists another net which behaves under the other and gives the same results, although perhaps not in the same time. Various applications of the calculus are discussed.



1943. Primera aproximación al concepto de Neurona Artificial

MODULO 01. NIVELATORIO

CLASE 01. INTRODUCCIÓN A LOS MODELOS IA

02

.UBA económicas  
Diplomaturas

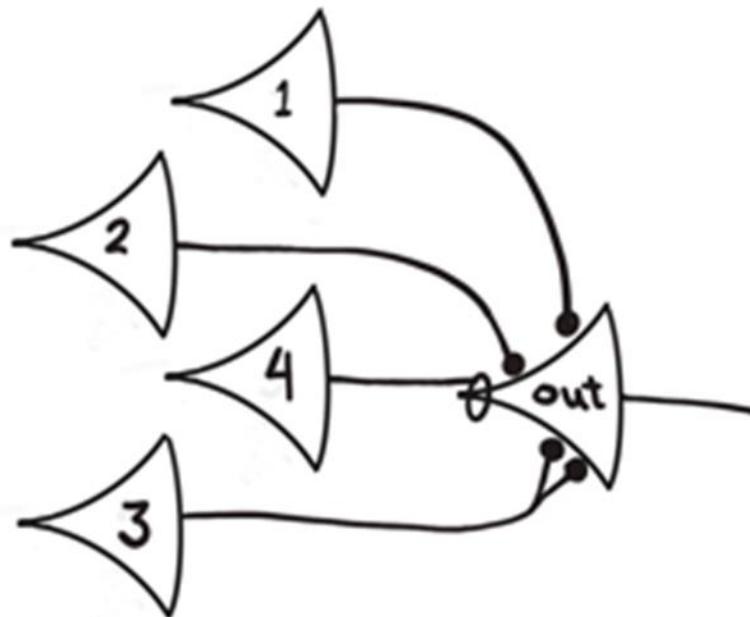
IA  
aplicada a  
Entornos Digitales  
de Gestión



# 1943

Primera aproximación a la neurona artificial.

Se propone un modelo matemático para describir cómo las neuronas podrían funcionar en el cerebro, lo que más tarde influyó en el campo de la inteligencia artificial y el desarrollo de computadoras y algoritmos inspirados en el funcionamiento del cerebro humano.



	$t_1$	$t_2$	$t_3$	$t_4$	$t_{i+1}$
F	F	F	F	F	F
T	F	F	F	F	F
F	T	F	F	F	F
T	T	F	F	T	T
F	F	T	F	T	T
T	F	T	F	T	T
F	T	T	F	T	T
T	T	T	F	T	T
X	X	X	T	F	F

$$N_{out}(t+1) = ((N_1(t) \cdot N_2(t)) \vee N_3(t)) \cdot \sim N_4(t)$$

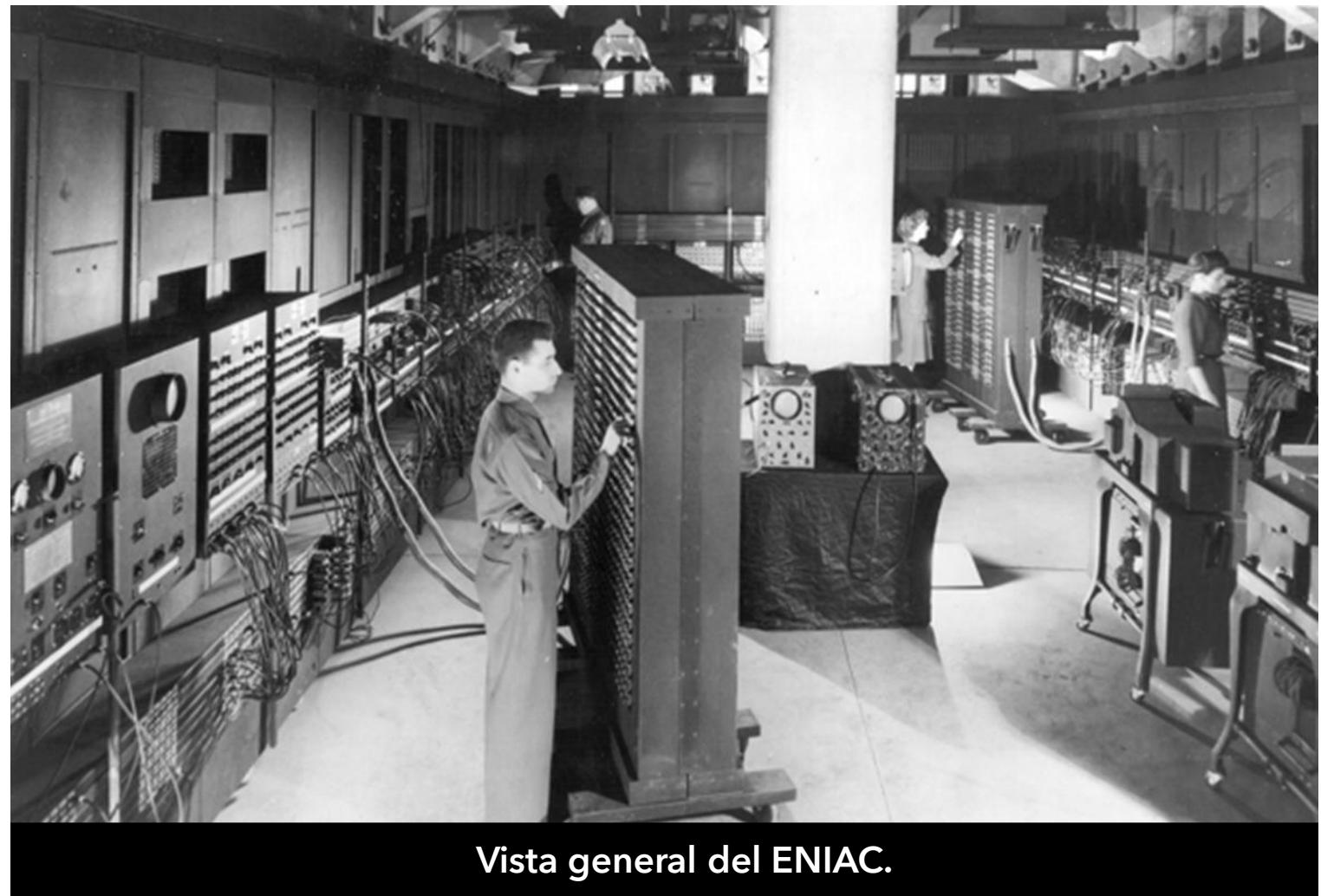
Red neuronal simple en notación McCulloch - Pitts, con tabla de verdad.



# 1945

## ENIAC. Primera computadora digital de uso general.

ENIAC (Electronic Numerical Integrator and Computer) fue la primera computadora electrónica de propósito general, creada en 1945 por John W. Mauchly y J. Presper Eckert en la Universidad de Pensilvania. Marca el nacimiento de la 1ra Generación de Computadoras.



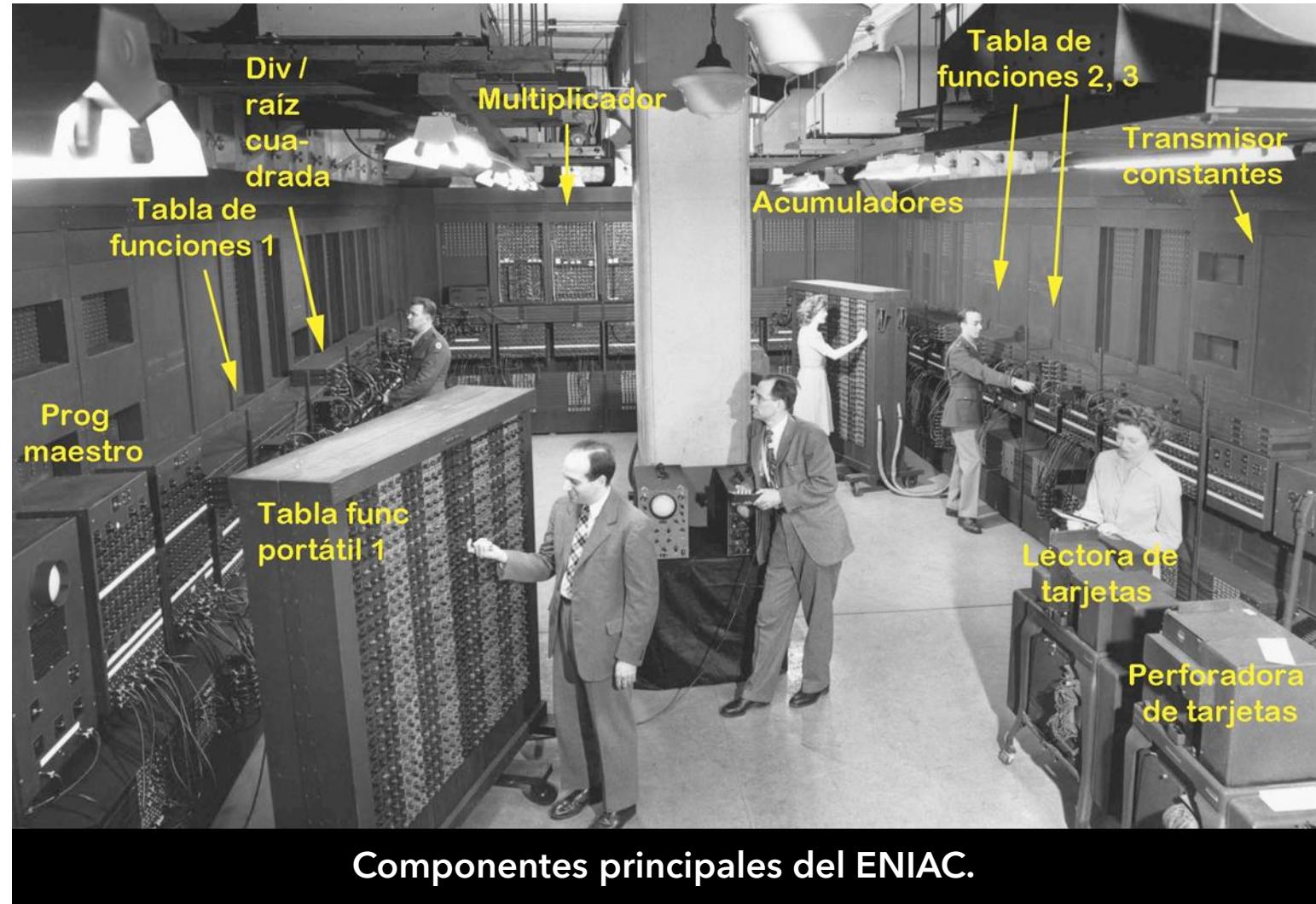
Vista general del ENIAC.



# 1945

**ENIAC.** Primera computadora digital de uso general.

Ocupaba 167 m<sup>2</sup> y utilizaba 17,468 tubos de vacío, realizando 5,000 operaciones por segundo. Fue desarrollada para calcular tablas de artillería para el ejército de EE. UU., marcando el inicio de la era de la computación moderna.



Componentes principales del ENIAC.



# 1950 - 1960

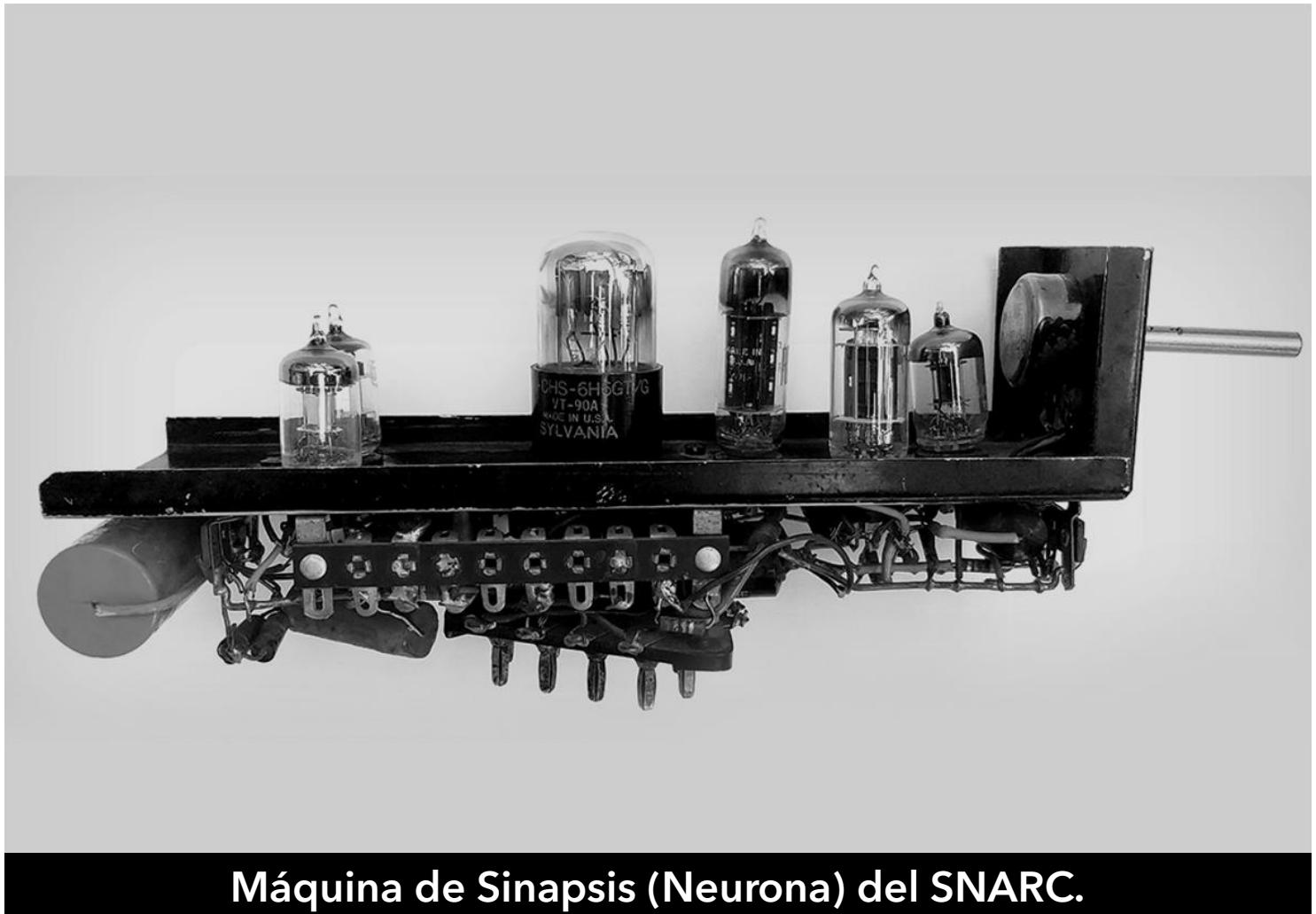
## Los gemelos crecen



# 1951

**SNARC.** Primer intento de una red neuronal artificial operativa.

El SNARC (*Stochastic Neural Analog Reinforcement Calculator*) fue creado en 1951 por Marvin Minsky y Dean Edmonds en la Universidad de Harvard. Con tubos de vacío y piezas electromecánicas simulaba una red de 40 máquinas de sinapsis (neuronas).



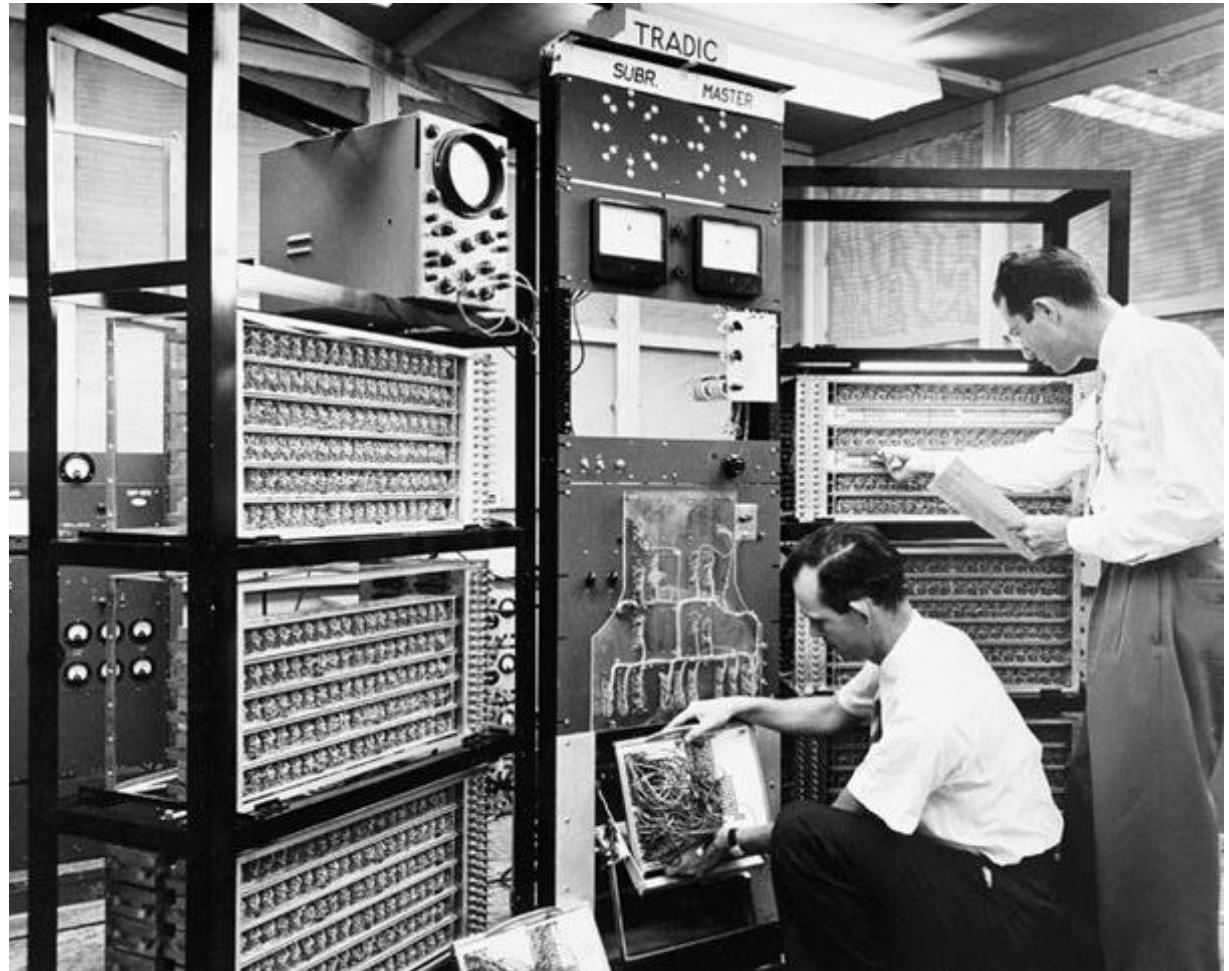
Máquina de Sinapsis (Neurona) del SNARC.



# 1954

## TRADIC. Primera computadora totalmente transistorizada.

El TRADIC en inglés: "TRansistorized DIgital Computer" fue el primer ordenador de la 2.<sup>a</sup> generación de computadores, siendo elaborado en los Laboratorios Bell. Se fabricó en el año 1955 y se creó la primera unidad para la venta en 1958. Consumía muchísimo menos electricidad y era más estable.



# 1956

La Conferencia de Dartmouth marcó el nacimiento de la IA como campo de estudio.

“Dartmouth Summer Research Project on Artificial Intelligence” es el nombre del encuentro que tuvo lugar en el verano de 1956 en la universidad Dartmouth College, ubicada en Hanover, Nuevo Hampshire (Estados Unidos).



Dartmouth Summer Research Project on Artificial Intelligence, 1956.



# 1956

La Conferencia de Dartmouth marcó el nacimiento de la IA como campo de estudio.

En esta conferencia, un grupo de científicos y expertos en computación y ciencia cognitiva se reunieron para discutir la posibilidad de crear máquinas capaces de pensar y aprender como los seres humanos.



John McCarthy



Marvin Minsky



Claude Shannon



Ray Solomonoff



Alan Newell



Herbert Simon



Arthur Samuel



Oliver Selfridge



Nathaniel Rochester



Trenchard More

Organizadores de la conferencia de Dartmouth en 1956



# 1956

La Conferencia de Dartmouth marcó el nacimiento de la IA como campo de estudio.

La primera definición de Inteligencia Artificial fue postulada por John McCarthy como “La ciencia e ingenio de hacer máquinas inteligentes, especialmente programas de cómputo inteligentes”.



John McCarthy (1927-2011)



# 1957

## Perceptrón. Primera implementación

El perceptrón, desarrollado por Frank Rosenblatt en 1957, es más simple y eficiente que el SNARC. Utiliza aprendizaje supervisado para clasificar patrones linealmente separables, siendo más práctico y accesible. Su implementación se dio en forma de programa en una IBM 704.



Perceptrón. IBM 704 en la Universidad de Cornell.



# 1958

## Perceptrón. Pionero de las Redes Neuronales y la IA Supervisada

El perceptrón se presenta formalmente en 1958, en el paper "A Probabilistic Model for Information Storage and Organization in the Brain". Este trabajo lo describe como un modelo matemático inspirado en el cerebro humano, diseñado para aprender y clasificar patrones.

*Psychological Review*  
Vol. 65, No. 6, 1958

### THE PERCEPTRON: A PROBABILISTIC MODEL FOR INFORMATION STORAGE AND ORGANIZATION IN THE BRAIN<sup>1</sup>

F. ROSENBLATT

*Cornell Aeronautical Laboratory*

If we are eventually to understand the capability of higher organisms for perceptual recognition, generalization, recall, and thinking, we must first have answers to three fundamental questions:

1 How is information about the

and the stored pattern. According to this hypothesis, if one understood the code or "wiring diagram" of the nervous system, one should, in principle, be able to discover exactly what an organism remembers by reconstructing the original sensory patterns from

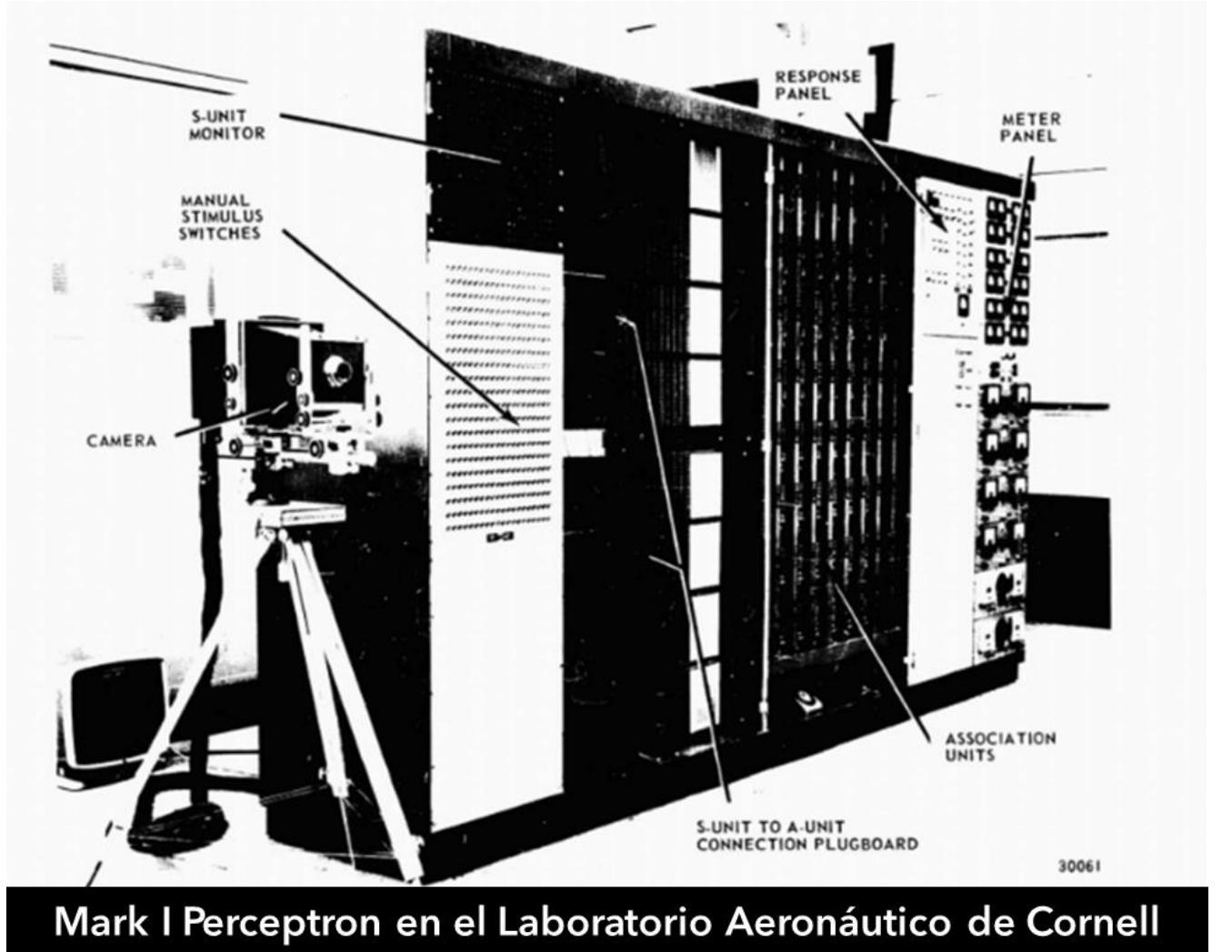
El Perceptrón es un modelo neuronal de clasificación binaria.



# 1959

## Perceptrón. Nace el "Mark I Perceptrón".

El Mark I fue la primera implementación del perceptrón por hardware. Funcionaba utilizando sensores ópticos para recibir entradas visuales, como imágenes de letras. El perceptrón ajustaba estos pesos a través de un algoritmo de aprendizaje supervisado, permitiéndole aprender a reconocer y clasificar patrones visuales simples.



Mark I Perceptron en el Laboratorio Aeronáutico de Cornell



# 1960 - 1970

## Lucha de Enfoques



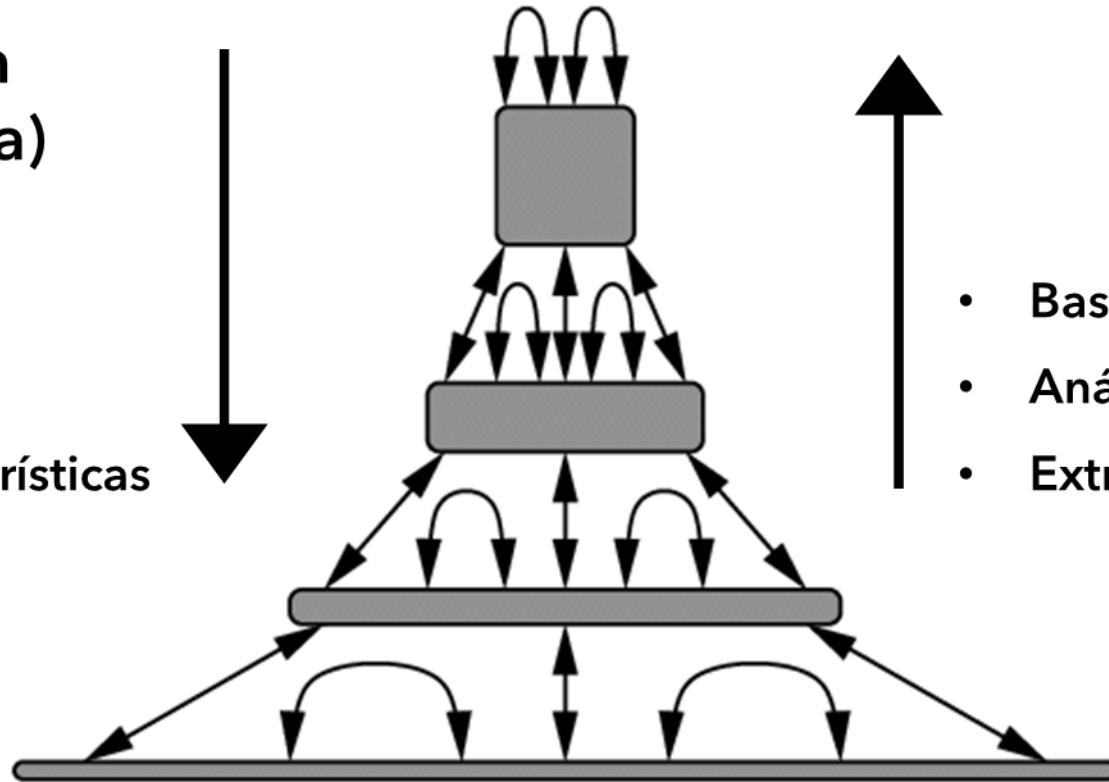
# Enfoques IA: Top-Down vs Bottom-Up

**Top - Down  
(IA Simbólica)**

- Basado en Hipótesis
- Síntesis
- Expansión de Características

**Bottom - Up  
(IA Neuronal)**

- Basado en Datos
- Análisis
- Extracción de Características



# 1964

**IBM S/360.** Nace la tercera generación de computadoras.

Fue una serie revolucionaria de computadoras mainframe que ofrecía compatibilidad entre modelos y capacidad de ejecutar el mismo software. Esta innovación permitió a las empresas actualizar su hardware sin reescribir programas, estableciendo nuevos estándares en la industria de la computación. Fue el primero en utilizar circuitos integrados.



# 1966

## ELIZA. La IA Psicoterapeuta.

Fue creada por Joseph Weizenbaum en el MIT, se presentó como uno de los primeros programas de procesamiento de Lenguaje Natural (NLP). Eliza simulaba la conversación humana con un enfoque simbólico basado en reglas y patrones predefinidos para interactuar con los usuarios.

Welcome to

EEEEEE	LL	IIII	ZZZZZ	AAAAA
EE	LL	II	ZZ	AA AA
EEEEEE	LL	II	ZZZ	AAAAAAA
EE	LL	II	ZZ	AA AA
EEEEEE	LLLLL	IIII	ZZZZZ	AA AA

Eliza is a mock Rogerian psychotherapist.  
The original program was described by Joseph Weizenbaum in 1966.  
This implementation by Norbert Landsteiner 2005.

ELIZA: Is something troubling you ?  
YOU: Men are all alike.  
ELIZA: What is the connection, do you suppose ?  
YOU: They're always bugging us about something or other.  
ELIZA: Can you think of a specific example ?  
YOU: Well, my boyfriend made me come here.  
ELIZA: Is it important to you that your boyfriend made you come here ?  
YOU: He says I'm depressed much of the time.  
ELIZA: I am sorry to hear that you are depressed.  
YOU: It's true. I am unhappy.  
ELIZA: Can you explain what made you unhappy ?  
YOU:

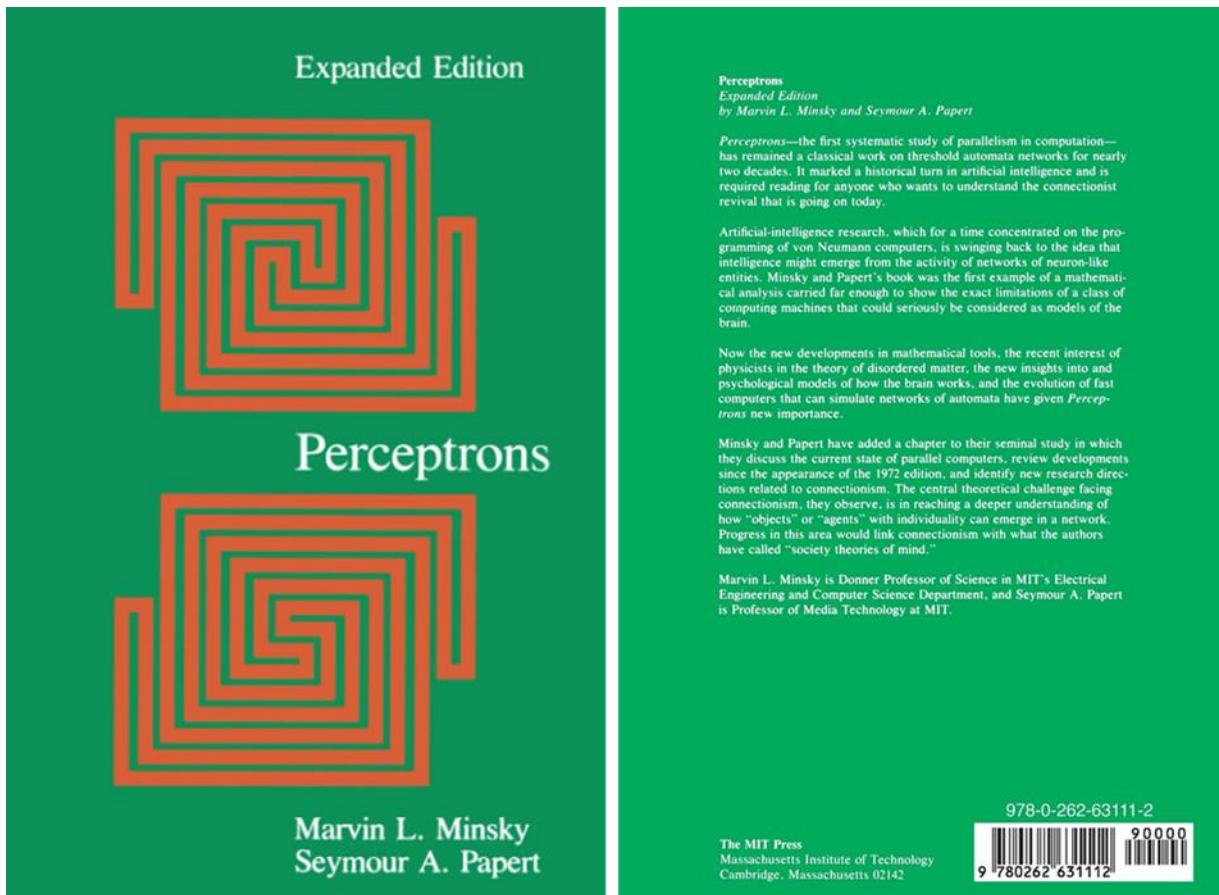
El nombre ELIZA proviene del personaje Eliza Doolittle de la obra teatral "Pigmalión", de George Bernard Shaw.



# 1969

## Perceptrons. El libro que adelantó el Invierno.

En 1969 Marvin Minsky y Seymour Papert, publicaron un libro ("Perceptrons") en el que mostraban cómo este tipo de redes neuronales eran incapaces de resolver problemas de discriminación no lineal. La publicación de este libro supuso un notable descenso en el interés y la financiación disponible para la investigación de redes neuronales.



Perceptrons. Tapa y contratapa de la edición explicada.



# 1970 - 1980

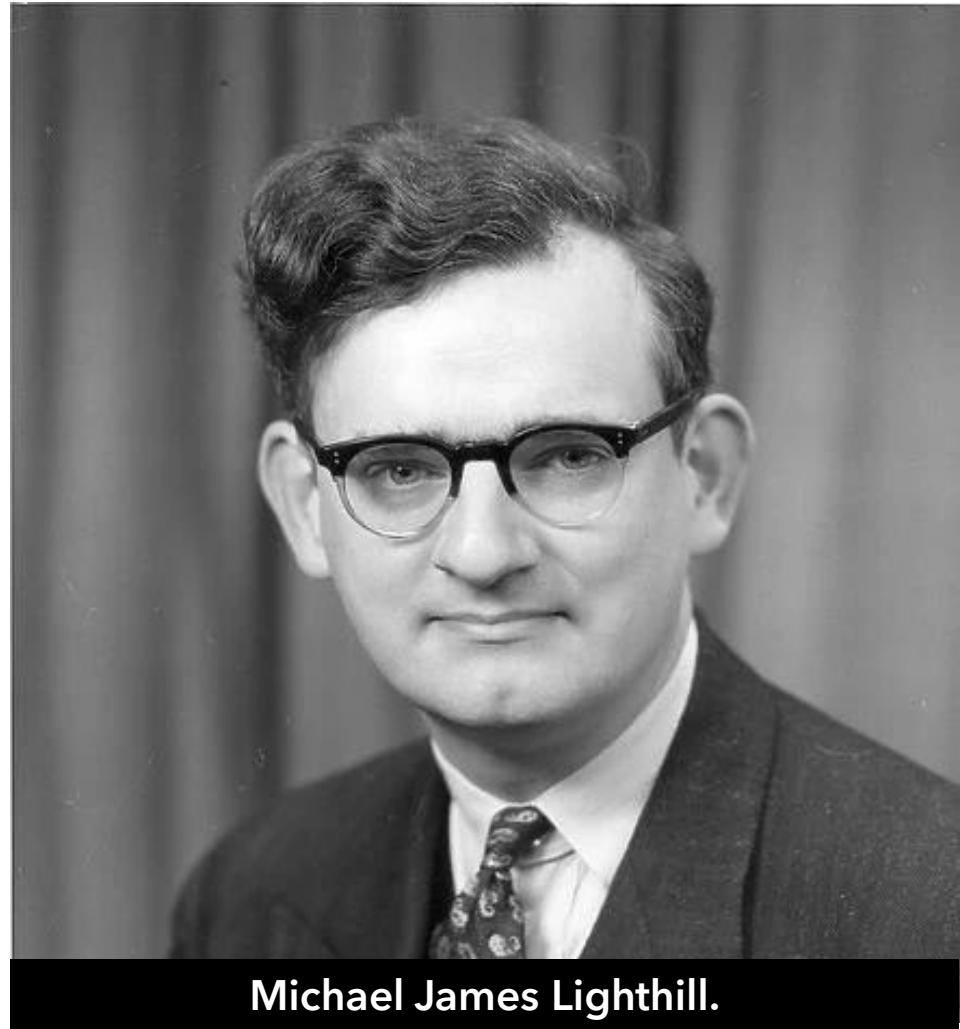
## Invierno de IA Verano de Computación



# 1973

Se acabó la plata.

El informe Lighthill emite un veredicto en su mayoría negativo sobre la investigación de IA en Gran Bretaña y sirve de base para la decisión del gobierno británico de discontinuar el apoyo a la investigación de IA en todas menos dos universidades.



Michael James Lighthill.



# 1974

Nace la cuarta generación de computadoras.

La Altair 8800, diseñada por Ed Roberts en el año 1974, fue la primera computadora personal usada en la cuarta generación de computadoras con un chip de CPU Intel 8080 y con el microprocesador de 16 bits.



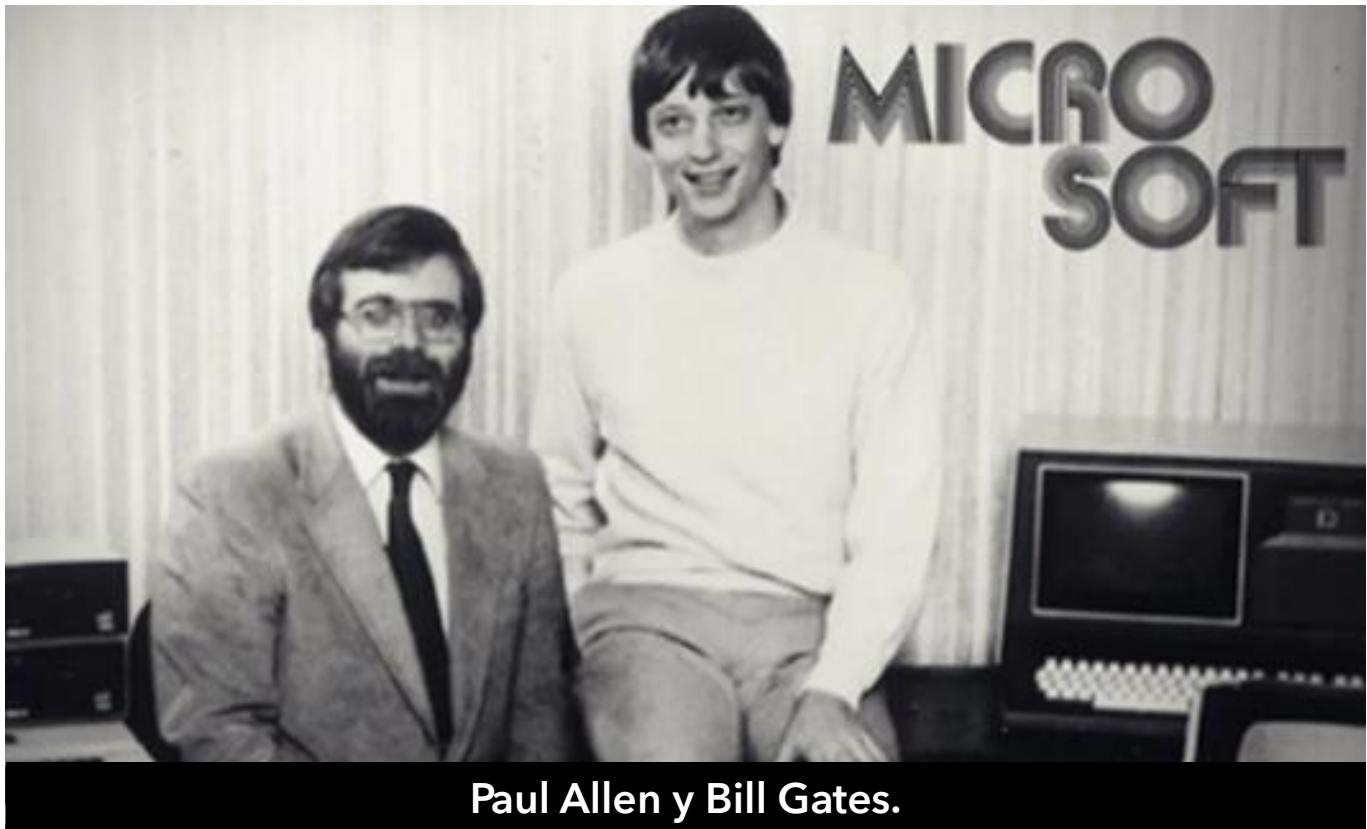
ALTAIR 8800, en su versión sin teclado ni monitor



# 1975

Nace Microsoft.

Paull Allen y Bill Gates, dos jóvenes apasionados de la tecnología, creen que la informática personal puede llegar a ser el futuro y crean el 4 de abril una pequeña compañía llamada Microsoft en Albuquerque, Nuevo México.



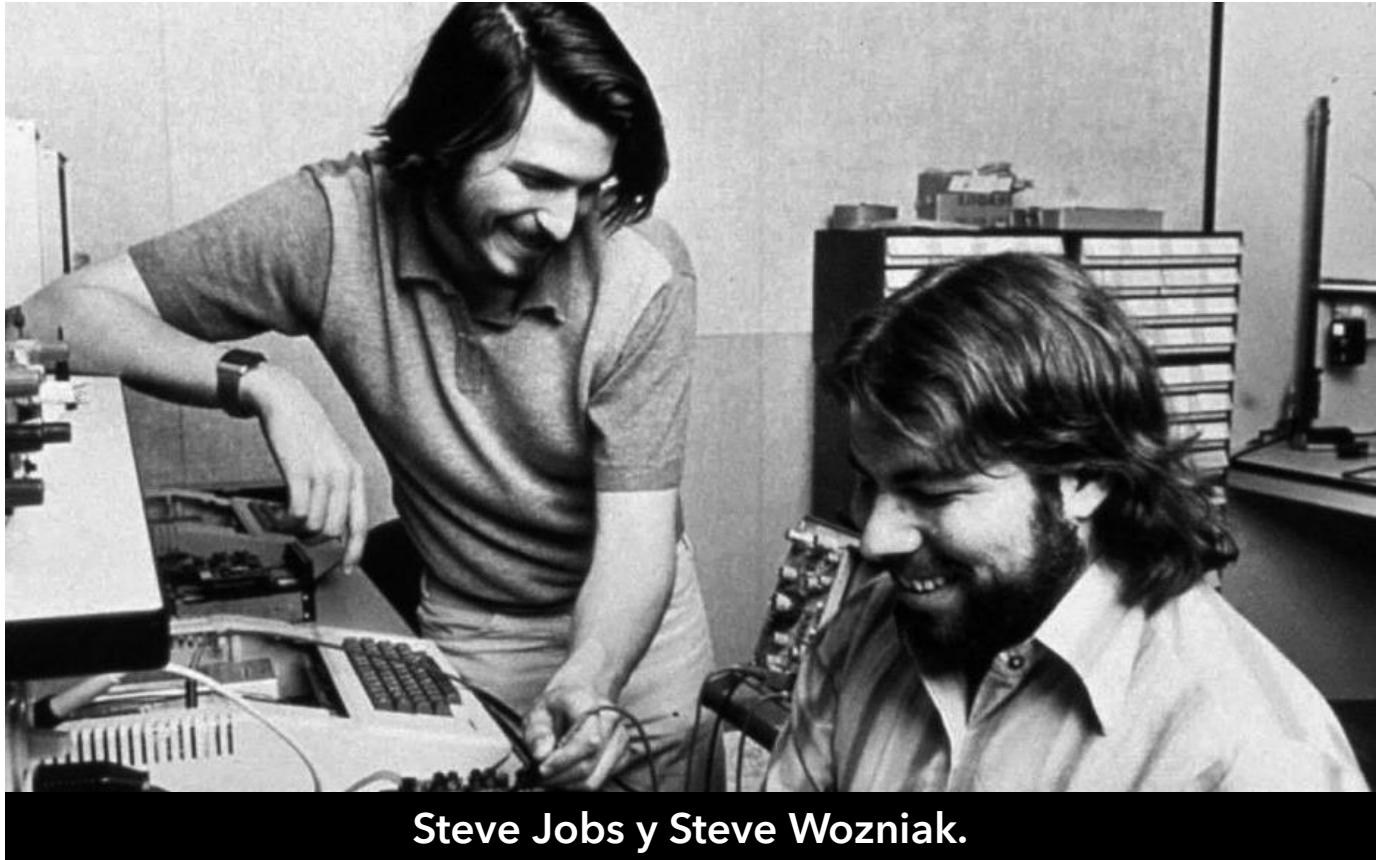
Paul Allen y Bill Gates.



# 1976

Nace Apple.

Steve Jobs fundó Apple el 1º de abril junto a su amigo Steve Wozniak. Ambos compartían la pasión por crear productos tecnológicos sencillos e intuitivos para el usuario común. En el garaje de Jobs diseñaron la primera computadora personal de Apple: el Apple I.



Steve Jobs y Steve Wozniak.



# 1977

## Inicia el furor de la micro computación.

En este año se introdujeron las microcomputadoras, que iniciaron la historia de las computadoras personales como dispositivos de consumo para el mercado masivo. Los primeros microcomputadores fueron el Apple II de Apple Computer, el TRS 80 modelo I de Tandy y el Commodore PET de Commodore.



Commodore PET.



# 1980 - 1990

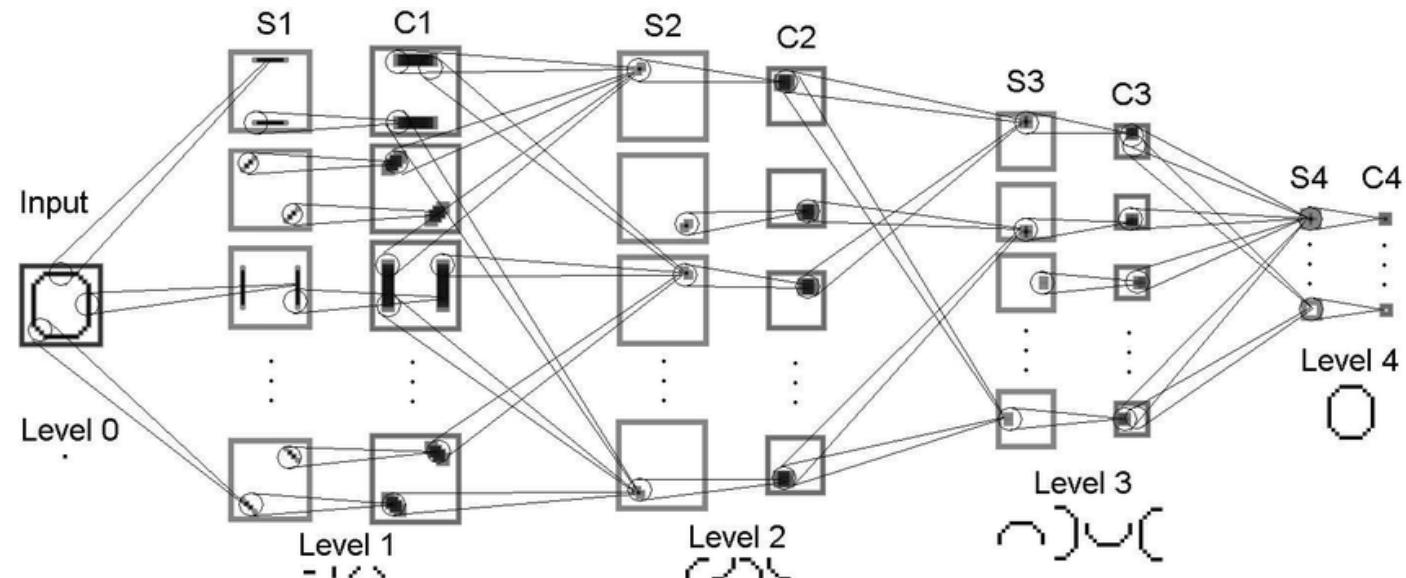
Vuelven las redes Neuronales,  
Comienza el reinado de las PC.



# 1980

## Neocognitron. Modelando la visión por computadora

El neocognitron es un tipo de red neuronal introducida por Kunihiko Fukushima en 1980. Diseñado para el reconocimiento de patrones visuales, este modelo está inspirado en el sistema visual de los mamíferos y es capaz de aprender y reconocer formas y objetos a través de una estructura jerárquica de capas, donde cada capa extrae características más complejas de la imagen.



Neocognitron. Representación de imágenes por capas



# 1981

Nace la PC IBM.

El lanzamiento del IBM 5150 en agosto de 1981 cambió el rumbo de la informática doméstica. No fue el primer ordenador personal como tal, pero sí el que transformó la industria por completo ya que su éxito motivó a los demás fabricantes para que construyeran piezas y software compatibles con el dispositivo.



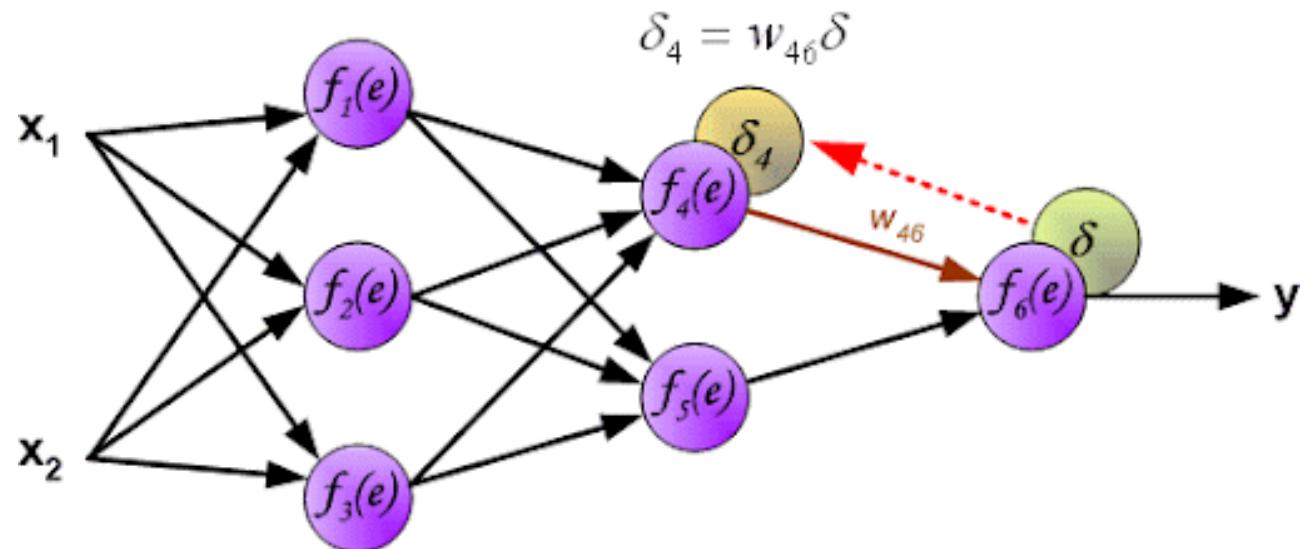
IBM PC 5150.



# 1986

## Algoritmo de Retropropagación.

David Rumelhart, Geoffrey Hinton y Ronald Williams describen en un trabajo académico, la aplicabilidad del algoritmo de *backpropagation* o retropropagación de errores al entrenamiento de redes neuronales multicapa. Solucionando así el problema planteado en 1969, respecto al ajuste de los pesos en las redes neuronales multicapa.



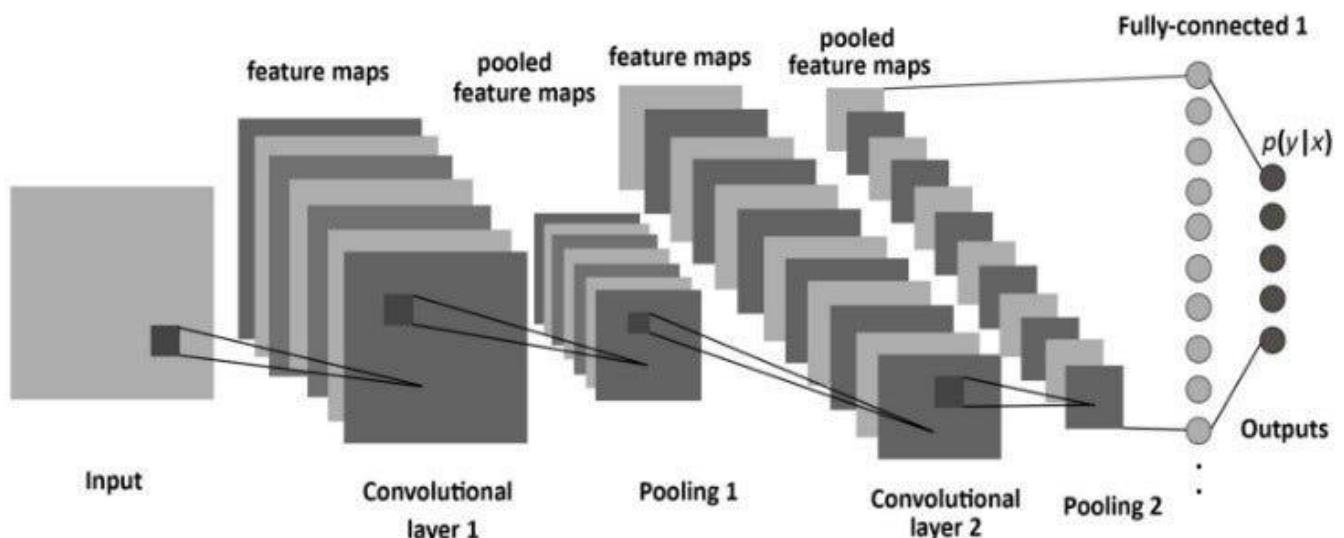
Representación del Algoritmo Backpropagation.



# 1989

## Redes Neuronales Convolucionales.

En 1989, Yann LeCun y sus colegas introdujeron las redes neuronales convolucionales (CNNs) modernas, aplicándolas exitosamente al reconocimiento de caracteres escritos a mano, como los dígitos en cheques bancarios. Estas redes combinan capas convolucionales, que capturan características locales de las imágenes, con capas de pooling para reducir la cantidad de parámetros.



Representación de una Red Neuronal Convolucional.



# 1990 - 2000

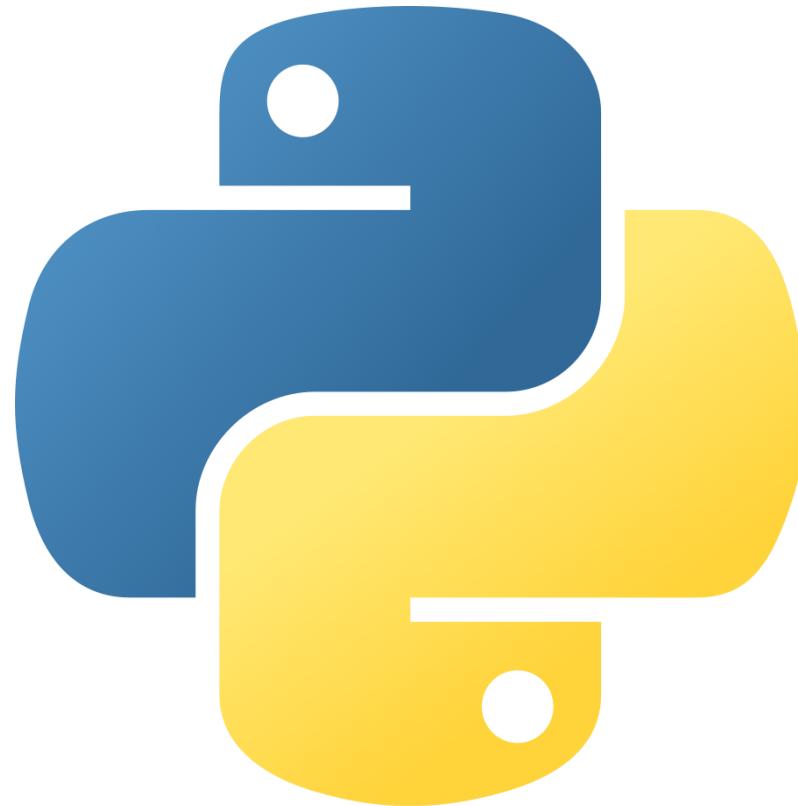
## Aparece INTERNET (y mil cosas más)



# 1991

Se libera el lenguaje Python.

Python es un lenguaje de alto nivel de programación interpretado cuya filosofía hace hincapié en la legibilidad de su código. Guido Van Rossum publicó la primera versión del código Python (versión 0.9.0) en 1991. Dicha versión ya incluía buenas características, como algunos tipos de datos y funciones para la gestión de errores.



Logo actual de Python.

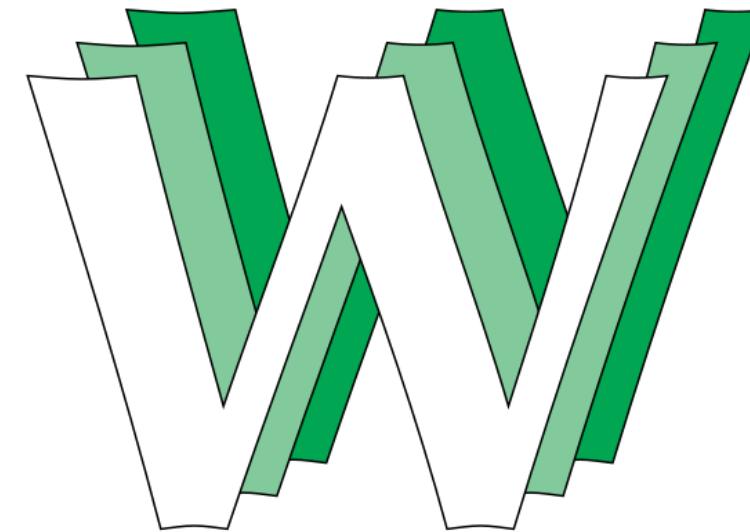


# 1993

Se libera la World Wide Web.

La World Wide Web (WWW) fue creada en 1989 por el científico británico Tim Berners-Lee, quien trabajaba en el CERN en Ginebra, Suiza. El 30 de abril de 1993, el CERN anunció que la WWW sería gratuita para todo el mundo, lo que contribuyó a su enorme crecimiento.

Let's Share What We Know



## World Wide Web

Primer logo de WWW creado por Robert Cailliau.



# 1994

**Se libera Linux.**

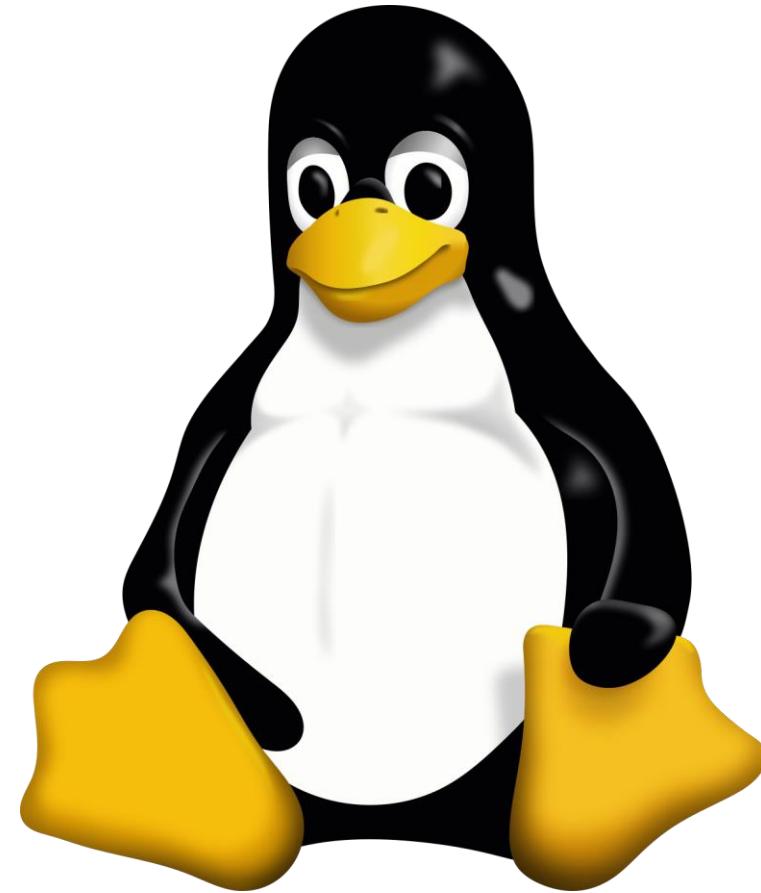
En marzo de 1994, Torvalds lanzó la versión 1.0 de Linux, que fue considerada la primera versión completa del sistema operativo. Linux comenzó como un proyecto personal de Torvalds inspirado en el sistema operativo Unix, pero con el tiempo se convirtió en el kernel del sistema operativo de GNU y uno de los sistemas operativos más importantes del mundo.

.

**MODULO 01. NIVELATORIO**

**CLASE 01. INTRODUCCIÓN A LOS MODELOS IA**

**34**



**Logo de Linux.**

**.UBA** económicas  
*Diplomaturas*

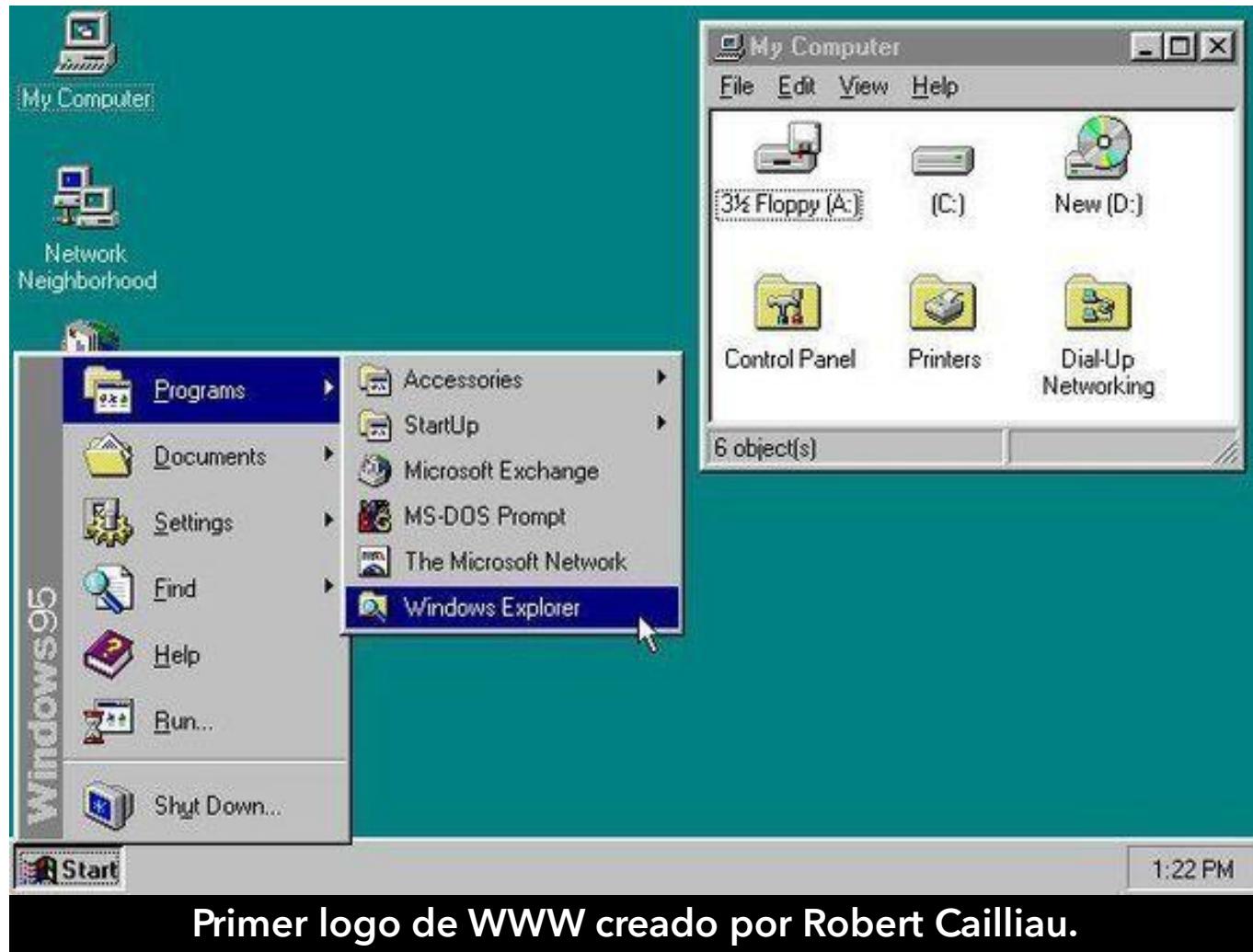
**IA** aplicada a  
Entornos Digitales  
de Gestión



# 1995

Aparece Windows 95.

Windows 95 fue un éxito en el mercado, vendiendo más de 40 millones de copias en el primer año. Fue elogiado por su facilidad de uso, rendimiento y compatibilidad. También provocó una ola de innovación y competencia en la industria del software, e influyó en el diseño y desarrollo de futuros sistemas..



# 1996

Aparece Clippy.

Clippy, oficialmente llamado "Clippit," fue un asistente virtual en forma de clip de papel introducido por Microsoft en Office 97 para ayudar a los usuarios con tareas comunes. Sin embargo, se volvió notorio por ser intrusivo y molesto, lo que llevó a su retirada en versiones posteriores de Office.



Primera versión de Clippy.



# 1997

Deep Blue gana al Ajedrez.

Deep Blue es un supercomputador desarrollado por IBM que se enfrentó al entonces campeón mundial de ajedrez, Garry Kasparov, en 1996 y 1997. El 11 de mayo de 1997, Deep Blue venció a Kasparov en un partido de seis partidas, marcando un hito en la historia del ajedrez y de la inteligencia artificial..



Enfrentamiento entre Deep Blue y Garry Kasparov.



# 1999

## El iMac se convierte en Furor .

El iMac G3 fue una serie de ordenadores personales vendidos por Apple entre 1998 y 2003. El iMac G3 de 1999 tenía un procesador PowerPC G3 de 400 MHz y 64 MB de RAM. Fue el primer Macintosh en prescindir de la disquetera y del Apple Desktop Bus, e incluía puertos USB. También fue el único ordenador de sobremesa de Apple disponible en una amplia gama de colores.



Vista trasera de las iMac de 1999.



# Seguimos la próxima



# 2000 - 2010

## El BIG DATA y las redes de alta velocidad



# 2010 - 2020

## La era de los smartphones y las redes sociales



# 2020 - 2024

## ¡La IA en todos lados!

MODULO 01. NIVELATORIO

CLASE 01. INTRODUCCIÓN A LOS MODELOS IA

42

.UBA económicas  
*Diplomaturas*

IA  
aplicada a  
Entornos Digitales  
de Gestión



# MÓDULO 01 - CLASE 01

# INTRODUCCION A GOOGLE COLAB Y PYTHON

.UBA económicas  
*Diplomaturas*

IA aplicada a  
Entornos Digitales  
de Gestión



## ¿Por qué PYTHON?

Python es un lenguaje que todo el mundo debería conocer. Su **sintaxis simple, clara y sencilla**; el tipado dinámico, el gestor de memoria, la **gran cantidad de librerías disponibles** y la potencia del lenguaje, entre otros, hacen que desarrollar una aplicación en Python sea sencillo, muy rápido y, lo que es más importante, divertido.



# Sintaxis

La **sintaxis en programación** se refiere a las reglas y convenciones que deben seguirse al escribir código para que sea válido y entendible por la computadora.



## Sintaxis en PYTHON

La sintaxis de Python es tan **sencilla y cercana al lenguaje natural** que los programas elaborados en Python **parecen pseudocódigo**. Por este motivo se trata además de uno de los mejores lenguajes para comenzar a programar.



# Sintaxis en PYTHON

```
print()
```

```
print("BIENVENIDOS a la Diplomatura")
```

```
print(objeto1, objeto2, ..., sep='', end='\n', file=sys.stdout,  
      flush=False)
```



## Pseudocódigo

Pseudocódigo es el empleo de un lenguaje cotidiano para plantear la estrategia que debe emplearse para resolver un problema. Como el nombre lo dice, es un código falso, pero representa el puente perfecto, para llegar al punto de iniciar a programar.



# Google colab

---

## Google Colab

Google Colab, o Colaboratory, es un entorno de desarrollo en línea gratuito que permite escribir y ejecutar código en Python. Está basado en la plataforma de Google Cloud y proporciona un entorno de ejecución de Jupyter Notebooks alojado en la nube.



# Google colab

---

Acceso Gratis

Recursos Potentes

Soporte Jupyter  
Notebooks

Integración con  
Google Drive

Compatible con  
Bibliotecas populares

MODULO 01. NIVELATORIO

CLASE 01. Introducción a Google Colab y Python

07

.UBA económicas  
*Diplomaturas*

**IA** aplicada a  
Entornos Digitales  
de Gestión



# Google colab

---

Google Colab

<https://colab.google/>



MODULO 01. NIVELATORIO

CLASE 01. Introducción a Google Colab y Python

08

.UBA económicas  
*Diplomaturas*

IA  
aplicada a  
Entornos Digitales  
de Gestión



# Google colab

---

## Google Colab Clase 1

[acortar.link/clase1python](https://acortar.link/clase1python)



MODULO 01. NIVELATORIO

CLASE 01. Introducción a Google Colab y Python

09

.UBA económicas  
*Diplomaturas*

**IA** aplicada a  
Entornos Digitales  
de Gestión





# Vamos a CODEAR

MODULO 01. NIVELATORIO

CLASE 01. Introducción a Google Colab y Python

10

.UBA económicas  
*Diplomaturas*

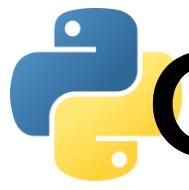
IA  
aplicada a  
Entornos Digitales  
de Gestión



# Operadores Aritmeticos

	Simbolo	Expresión algebraica	Expresión en Python	Ejemplo
Suma	+	$a + b$	<code>a + b</code>	In [1] : 9+2 Out[1]: 11
Resta	-	$a - b$	<code>a - b</code>	In [1] : 9-2 Out[1]: 7
Multiplicación	*	$a * b$	<code>a * b</code>	In [1] : 9*2 Out[1]: 18
División	/	$\frac{a}{b}$	<code>a / b</code>	In [1] : 9/2 Out[1]: 4,5
División parte entera	//	$\left[ \frac{a}{b} \right]$	<code>a // b</code>	In [1] : 9//2 Out[1]: 4
Módulo	%	$a \ mod \ b$	<code>a % b</code>	In [1] : 8%2 Out[1]: 0
Exponente	**	$a^b$	<code>a ** b</code>	In [1] : 8**2 Out[1]: 64





# Operadores Aritmeticos

---

## División Entera : a//b

La división entera es una operación matemática que consiste en dividir dos números y obtener como resultado un cociente entero, descartando cualquier parte fraccionaria o decimal.

## Parte Residual en una División: a%b

La división  $10 \div 3$ , el cociente es 3 y la parte residual es 1. Esto significa que 3 cabe en 10 tres veces y queda un residuo de 1.





# Variables

---

Una variable es un contenedor que almacena datos que pueden ser modificados durante la ejecución del programa. Las variables en Python son utilizadas para almacenar valores como números, cadenas de texto, listas, diccionarios, entre otros tipos de datos.





# Algoritmos

Los libros de programación definen a un algoritmo como una secuencia ordenada de pasos, que han de ejecutarse uno después de otro.

Definición más actual, por ejemplo, Deitel, señala que un algoritmo es un procedimiento realizado para resolver un problema, en términos de:

1. Las acciones a ejecutar, y
2. El orden en el cual se realizan estas acciones.



## Algoritmo para preparar un Café:

1. Calientas agua
2. Agregas café
3. Agregas Azucar
4. Lo bebes





# Sentencias

---

**if ... elif ... else**



Esta sentencia realiza una acción, cuando la condición es verdadera;  
y la ignora, cuando es falsa.

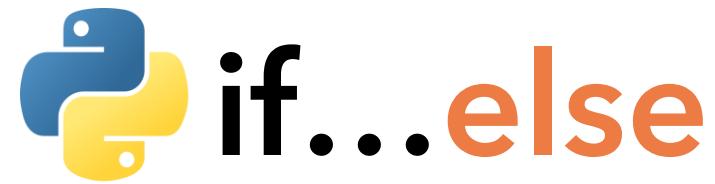




# Operadores

	Operador algebraico	Operador en Python	Descripcion
>	$a > b$	$a > b$	a es mayor que b
<	$a < b$	$a < b$	a es menor que b
$\geq$	$a \geq b$	$a \geq b$	A es mayor o igual que b
$\leq$	$a \leq b$	$a \leq b$	a es menor o igual que b
=	$a = b$	$a == b$	a es igual que b
$\neq$	$a \neq b$	$a != b$	a es diferente que b





Esta sentencia realiza una acción, cuando la condición es verdadera; y realiza la otra acción cuando es falsa.





# if...elif ...else

**Esta sentencia realiza una de diferentes acciones dependiendo como hayan sido calificadas las otras condiciones.**

**Se utiliza en estructuras condicionales para evaluar múltiples condiciones secuenciales después de una instrucción if, pero antes de la instrucción else.**





# Mejoramos el código

`input()` es una función integrada en Python que se utiliza para solicitar al usuario que ingrese datos desde la consola durante la ejecución del programa.

`int` es un tipo de datos que representa números enteros, es decir, números sin parte fraccionaria o decimal.

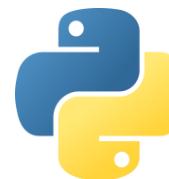




## Clasificador de estatura:

Solicita al usuario que ingrese su estatura en centímetros y clasifica su estatura en una de las siguientes categorías: "Baja" (menos de 150 cm), "Media" (entre 150 cm y 170 cm), "Alta" (entre 170 cm y 190 cm) o "Muy alta" (más de 190 cm).





## Clasificador de rendimiento académico:

Solicita al usuario que ingrese su calificación final en una asignatura y clasifica su rendimiento académico en una de las siguientes categorías: "Excelente" (9 o más), "Bueno" (7 o más), "Suficiente" (4 o más) o "Insuficiente" (menos de 4).





## Calcular el índice de masa corporal (IMC):

Solicita al usuario que ingrese su peso en kilogramos y su altura en metros. Calcula su índice de masa corporal (IMC) utilizando la fórmula **IMC = peso / altura<sup>2</sup>**. Luego, muestra un mensaje que indique si la persona está bajo peso, peso normal, sobrepeso o obeso, según el valor de su IMC.



**PARTE I**

**CLASE**

**2**

# MÓDULO 01 - CLASE 02

## INTRODUCCIÓN A LOS MODELOS IA (II)

.UBA económicas  
*Diplomaturas*

IA aplicada a  
Entornos Digitales  
de Gestión



# 2000 - 2010

## El BIG DATA y las redes de alta velocidad

MODULO 01. NIVELATORIO

CLASE 02. INTRODUCCIÓN A LOS MODELOS IA (II)

01

.UBA económicas  
*Diplomaturas*

**IA** aplicada a  
Entornos Digitales  
de Gestión



# 2000

Aparecen los pendrives.

IBM y Trek Technology lanzan al mercado los primeros pendrives con un almacenamiento de 8 megabytes. Esta tecnología generó la posterior desaparición de los diskettes como fuente principal de transporte de archivos.

(Actualmente los pendrives tienen hasta 2 terabytes de capacidad, es decir 2 millones de megabytes)



Imagen de uno de los primeros pendrives.



# 2001

## Aparecen el iPod y Wikipedia.

El 15 de Enero de 2001 se presenta Wikipedia, una enciclopedia colaborativa que generó una revolución en la denominada democratización de la generación de conocimiento.

El 23 de Octubre se lanza el iPod, el reproductor de música que permitía “tener 1000 canciones en el bolsillo”.



Crear una cuenta Acceder ...

## iPod

Artículo Discusión

Leer Editar Ver histórico Herramientas

El **iPod** fue una línea de [reproductores de audio digital](#) portátiles diseñados y comercializados por [Apple Inc.](#) Fue presentado por primera vez el 23 de octubre de 2001 por [Steve Jobs](#) en el Apple Town Hall en Cupertino, California con un costo de 399 dólares<sup>2</sup> y descontinuado el 10 de mayo de 2022<sup>3</sup>.

En sus inicios, el dispositivo de almacenamiento musical pesaba 184 gramos, contaba con [5G](#) de capacidad y una pantalla de cristal líquido de 2 pulgadas. Su batería permitía a los usuarios escuchar música por hasta 10 horas<sup>4</sup> y tenía la capacidad de pasar más de 10 gigabytes de



Ipod en el sitio web de wikipedia.



# 2002

El wifi comienza a popularizarse.

El año 2002 fue clave para el Wi-Fi debido a la aprobación del estándar 802.11g, que ofrecía velocidades significativamente mayores de hasta 54 Mbps y era compatible con el estándar anterior 802.11b. Esto facilitó una adopción más amplia del Wi-Fi en entornos domésticos y comerciales. En enero 2024 se lanzó Wi-Fi 7 que alcanza una velocidad máxima de 46 Gbps.



Logo del Wifi.



# 2003

## Nace Tesla, Inc.

Tesla Motors fue fundada en julio de 2003 por los ingenieros Martin Eberhard y Marc Tarpenning. Posteriormente, Ian Wright, Elon Musk y JB Straubel se unieron y fueron reconocidos como cofundadores. Elon Musk, quien ya era un empresario exitoso, invirtió \$6.5 millones en Tesla, asumiendo el rol de presidente, mientras que Eberhard se convirtió en director ejecutivo y Straubel en director de tecnología.



TESLA

Logo de Tesla



# 2004

## Se lanza Gmail.

Gmail se originó como un proyecto destinado únicamente a los empleados de Google, funcionando inicialmente como un sistema de mensajería instantánea. Originalmente denominado Google Groups, fue lanzado al público en 2004 bajo el nombre de Gmail.

Inicialmente, el acceso estaba restringido y solo se podía obtener mediante un enlace de invitación de un usuario ya registrado.



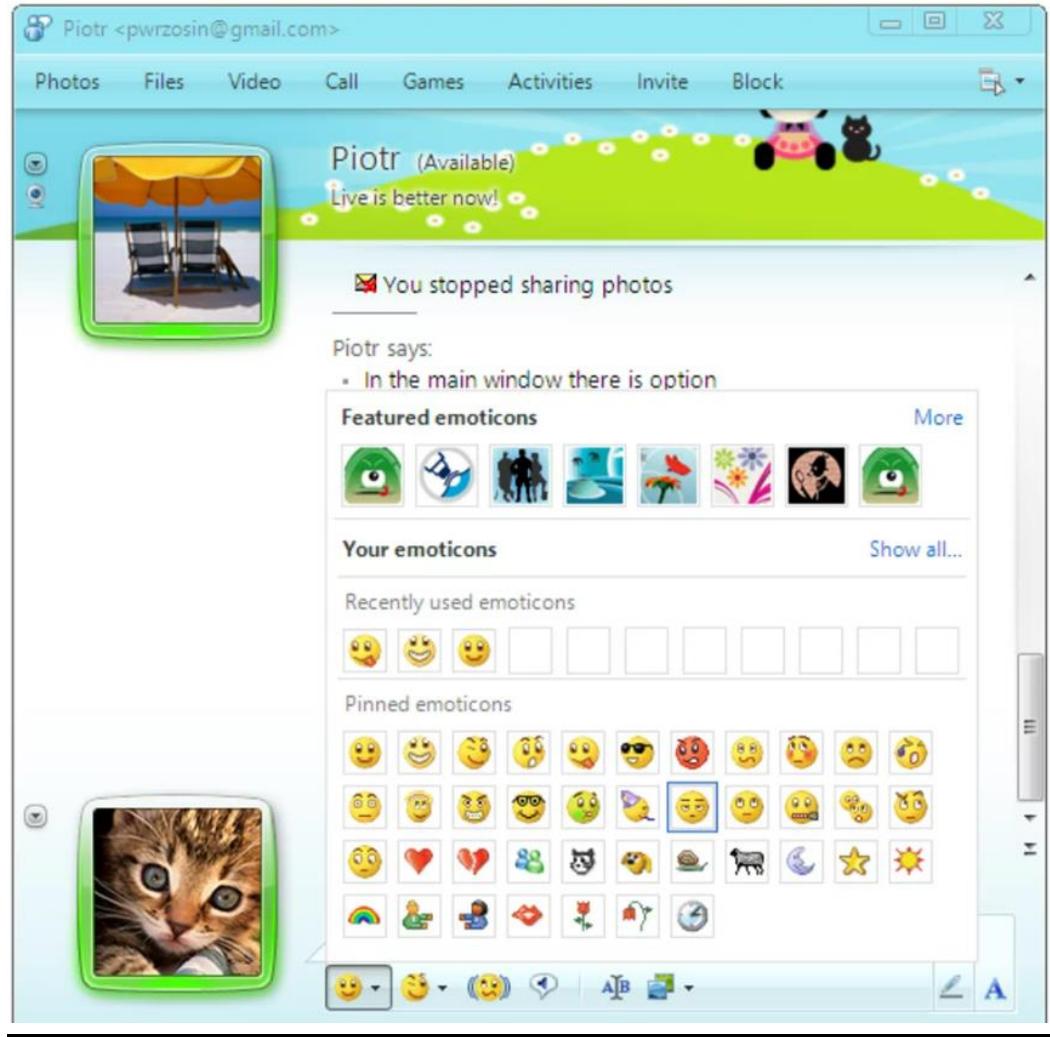
Primer logo de Gmail.



# 2005

## Comienza el reinado del Messenger.

En 2005, Windows Live Messenger (WLM) se lanzó como una actualización del MSN Messenger, incorporando nuevas características. Este lanzamiento incluyó funciones avanzadas como la personalización del usuario y la capacidad de enviar mensajes a contactos desconectados, fortaleciendo su posición como una plataforma líder de mensajería instantánea.



Interfaz de chat del WLM.



# 2006

## Se consolida el concepto de BIG DATA.

En 2006, se lanzó Hadoop, una plataforma de código abierto diseñada para almacenar y analizar grandes volúmenes de datos. Simultáneamente, las bases de datos NoSQL, que son sistemas de almacenamiento no relacionales, comenzaron a ganar popularidad. Estas tecnologías emergieron como soluciones clave para manejar la creciente cantidad y complejidad de los datos (BIG DATA) en diversos sectores.



Logo de Hadoop.



# 2007

Aparece el iPhone.

El iPhone original, también conocido como iPhone 1 o iPhone 2G, fue presentado por Steve Jobs el 9 de enero de 2007 en la conferencia anual de Macworld y salió a la venta el 29 de junio de ese mismo año. Este teléfono inteligente fue el primero de la serie iPhone y combinaba un iPod y un teléfono.



Sitio web de Twitter en el 2007.



# 2008

**Google lanza Chrome y el primer teléfono Android.**

En 2008, Google lanzó Chrome, un navegador destacado por su rapidez y simplicidad, y se introdujo el primer teléfono Android, el T-Mobile G1, marcando el comienzo de Android como una importante plataforma móvil. Ambos lanzamientos tuvieron un impacto significativo en el ámbito tecnológico.



Imagen del primer HTC Dream o T-Mobile G1 y logo de Chrome.

# 2009

## El Imac se convierte en Furor .

La tecnología 4G, también conocida como LTE, fue introducida en 2009. En marzo de 2009, la empresa lituana LRTC anunció la primera red 4G operativa en los países bálticos. En diciembre de 2009, Telia-Sonera anunció la primera implementación comercial de LTE en Estocolmo y Oslo. Esta red cuenta con una velocidad (teórica) máxima de descarga de 150 Mbps y de subida de 50 Mbps. (1Gbits en ubicaciones fijas)



LTE Significa "Long Term Evolution" siendo conocida como 4G.



# 2010 - 2020

## La era de los smartphones y las redes sociales



# 2010

Se lanza el iPad.

El iPad 2010, también conocido como iPad 1G o iPad original, fue anunciado por Steve Jobs el 27 de enero de 2010 en un evento de prensa de Apple. El iPad 2010 estaba disponible en dos versiones: Wi-Fi y Wi-Fi + 3G. El modelo Wi-Fi pesaba 680 gramos y el modelo Wi-Fi + 3G pesaba 730 gramos.



Imagen de Steve Jobs y el iPad 1G.



# 2011

**Se presenta Siri integrado al iPhone.**

Siri es un asistente personal inteligente de Apple, disponible en dispositivos como iPhone, iPad, Mac, Apple Watch y HomePod. Utiliza tecnologías de voz y procesamiento de lenguaje natural para responder a solicitudes de los usuarios, ejecutando tareas como enviar mensajes, hacer llamadas, establecer recordatorios, proporcionar direcciones, y controlar la automatización del hogar, entre otras funciones.



Presentación de Siri integrado al iPhone 4S.

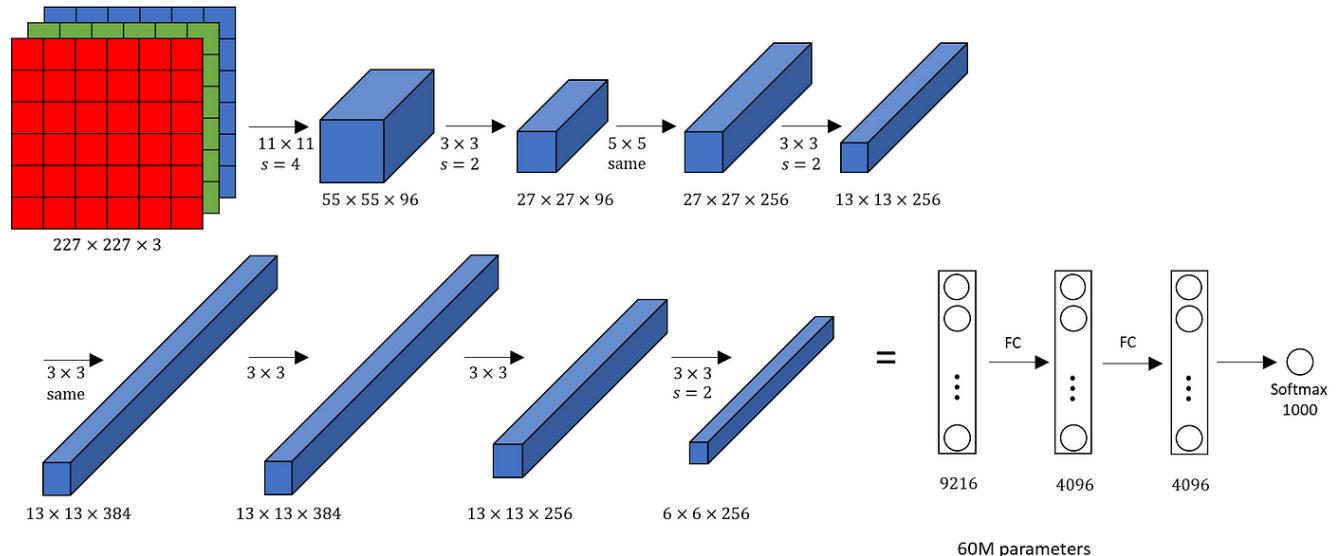


# 2012

## El Imac se convierte en Furor .

En 2012, AlexNet, una red neuronal convolucional, gana la competencia ImageNet, reduciendo del 26% al 15% la tasa de errores en la clasificación.

Ese mismo año, el proyecto Google Brain destacó al entrenar una red neuronal que identificaba gatos en videos de YouTube, mostrando la capacidad de aprender de grandes cantidades de datos no estructurados.



Estructura de la red neuronal AlexNet.



# 2013

Una buena idea que fracasó.

En abril de 2013, Google presentó las Google Glass, unas gafas inteligentes de realidad aumentada. El 15 de abril de 2013, la edición Explorer Edition de las Google Glass salió a la venta para desarrolladores cualificados por 1.500 dólares. Esta edición incluía los componentes de hardware necesarios para crear las primeras aplicaciones.



Imagen de las Google Glass



# 2014

## El DeepLearning juega a la Atari.

Se da a conocer un trabajo denominado "*Playing Atari with Deep Reinforcement Learning*" publicado por DeepMind, ya bajo el ala de Google, en el que demostraban que su algoritmo de aprendizaje profundo podía aprender a jugar videojuegos de Atari solo a partir de la entrada visual de píxeles, una demostración impresionante de aprendizaje por refuerzo.

## Playing Atari with Deep Reinforcement Learning

Volodymyr Mnih   Koray Kavukcuoglu   David Silver   Alex Graves   Ioannis Antonoglou

Daan Wierstra   Martin Riedmiller

DeepMind Technologies

{vlad,koray,david,alex.graves,ioannis,daan,martin.riedmiller} @ deepmind.com

### Abstract

We present the first deep learning model to successfully learn control policies directly from high-dimensional sensory input using reinforcement learning. The model is a convolutional neural network, trained with a variant of Q-learning, whose input is raw pixels and whose output is a value function estimating future rewards. We apply our method to seven Atari 2600 games from the Arcade Learning Environment, with no adjustment of the architecture or learning algorithm. We find that it outperforms all previous approaches on six of the games and surpasses a human expert on three of them.

El paper se publicó en diciembre del 2013, pero se conoció en el 2014.



# 2015

## Facebook lanza DeepFace.

DeepFace es un sistema de reconocimiento facial de aprendizaje profundo creado por un equipo de desarrollo en Facebook. Este define rostros humanos en imágenes digitales. Emplea una red neuronal de 9 capas con más de 120 millones de conexiones, y ha sido entrenado en 4 millones de imágenes subidas por los usuarios de Facebook. Se ha mencionado que el sistema tiene un 97% de aciertos.

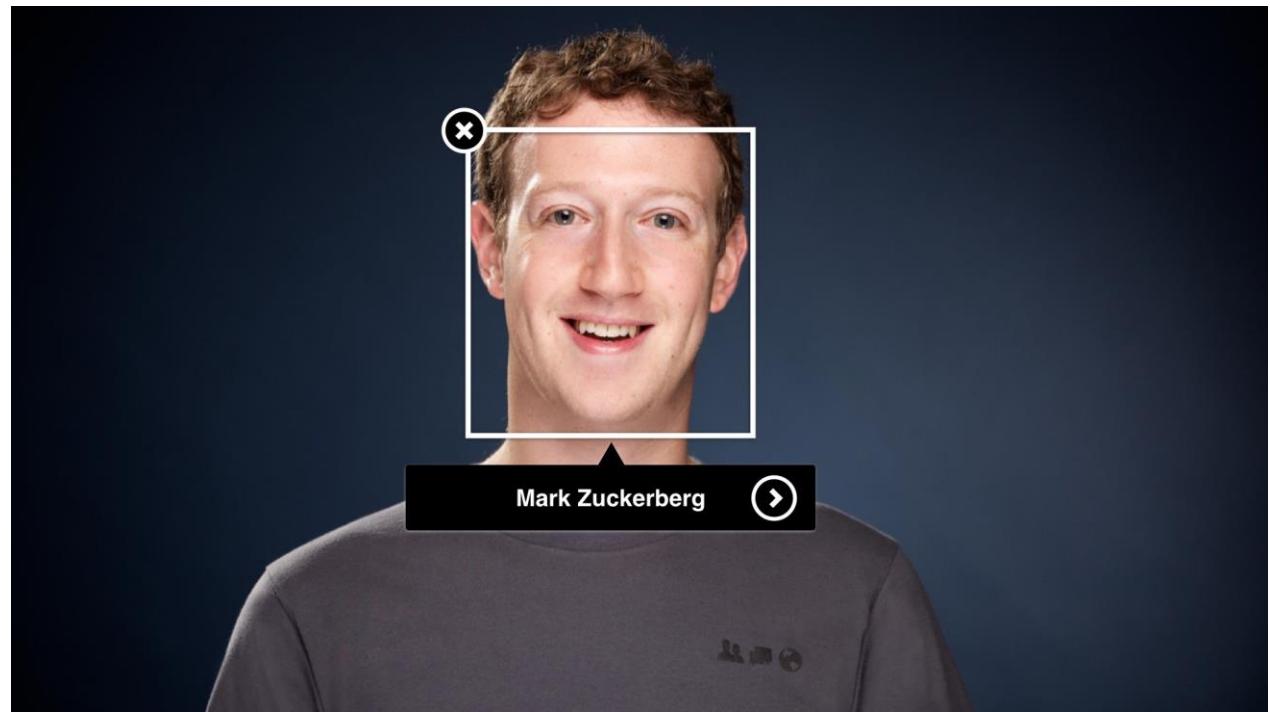


Imagen de presentación de DeepFace.



# 2016

**AlphaGo derrota al mejor .**

En marzo de 2016, AlphaGo logró una hazaña histórica al derrotar a Lee Sedol, uno de los mejores jugadores de Go del mundo, en una serie de cinco partidas, ganando cuatro de ellas. Este evento marcó un hito en la historia de la inteligencia artificial, demostrando que los sistemas de aprendizaje profundo pueden superar a los humanos en tareas complejas que requieren intuición y creatividad.



**AlphaGo derrota a Lee Sedol , el mejor jugador de GO**



# 2017

## Transformando la atención.

Las redes Transformer (TNN) fueron introducidas por primera vez en el artículo "*Attention is All You Need*", publicado por investigadores de Google en junio de 2017. Este trabajo presentó una nueva arquitectura basada en mecanismos de atención, eliminando la necesidad de recurrir a redes recurrentes (RNN) o convolucionales (CNN) para manejar secuencias de datos.

### Attention Is All You Need

Ashish Vaswani\*  
Google Brain  
avaswani@google.com

Noam Shazeer\*  
Google Brain  
noam@google.com

Niki Parmar\*  
Google Research  
nikip@google.com

Jakob Uszkoreit\*  
Google Research  
usz@google.com

Llion Jones\*  
Google Research  
llion@google.com

Aidan N. Gomez\* †  
University of Toronto  
aidan@cs.toronto.edu

Lukasz Kaiser\*  
Google Brain  
lukaszkaiser@google.com

Illia Polosukhin\* ‡  
illia.polosukhin@gmail.com

#### Abstract

The dominant sequence transduction models are based on complex recurrent or convolutional neural networks that include an encoder and a decoder. The best performing models also connect the encoder and decoder through an attention mechanism. We propose a new simple network architecture, the Transformer, based solely on attention mechanisms, dispensing with recurrence and convolutions entirely. Experiments on two machine translation tasks show these models to be superior in quality while being more parallelizable and requiring significantly less time to train. Our model achieves 28.4 BLEU on the WMT 2014 English-to-German translation task, improving over the existing best results, including ensembles, by over 2 BLEU. On the WMT 2014 English-to-French translation task, our model establishes a new single-model state-of-the-art BLEU score of 41.8 after training for 3.5 days on eight GPUs, a small fraction of the training costs of the best models from the literature. We show that the Transformer generalizes well to other tasks by applying it successfully to English constituency parsing both with large and limited training data.

Las TNN analizan todas las conexiones en simultaneo.



# 2018

Nace GPT-1.

GPT-1, desarrollado por OpenAI y lanzado en junio de 2018, fue uno de los primeros modelos en usar la arquitectura Transformer para generar texto, aprendiendo de un amplio corpus para predecir la próxima palabra en secuencias de texto. Este enfoque marcó un cambio significativo en el desarrollo de modelos de lenguaje, influyendo en la creación de modelos posteriores más avanzados.



GPT-1 tenía unos 117 millones de parámetros.

MODULO 01. NIVELATORIO

CLASE 02. INTRODUCCIÓN A LOS MODELOS IA (II)

21

.UBA económicas  
*Diplomaturas*

**IA** aplicada a  
Entornos Digitales  
de Gestión



# 2019

## El Imac se convierte en Furor .

La tecnología 5G comenzó a operar en varias ciudades del mundo. El 5G es la quinta generación de tecnología de redes móviles, diseñada para ofrecer velocidades de conexión mucho más rápidas, una mayor capacidad y una menor latencia en comparación con las redes 4G anteriores. La velocidad máxima teórica del 5G es de 10 Gbps.



5g es hasta 10 veces más rápida que 4g.



# Seguimos la próxima



# 2020 - 2024

## ¡La IA en todos lados!

MODULO 01. NIVELATORIO

CLASE 02. INTRODUCCIÓN A LOS MODELOS IA (II)

24

.UBA económicas  
*Diplomaturas*

**IA** aplicada a  
Entornos Digitales  
de Gestión



## Documentación de las Bibliotecas

- NumPy: <https://numpy.org/doc/>
- Pandas: <https://pandas.pydata.org/docs/>
- Matplotlib: <https://matplotlib.org/stable/contents.html>
- Seaborn: <https://seaborn.pydata.org/tutorial.html>
- Requests: <https://docs.python-requests.org/en/latest/>
- Django: <https://docs.djangoproject.com/en/stable/>
- Flask: <https://flask.palletsprojects.com/en/2.0.x/>
- TensorFlow: [https://www.tensorflow.org/api\\_docs](https://www.tensorflow.org/api_docs)
- PyTorch: <https://pytorch.org/docs/stable/index.html>
- Tkinter: <https://docs.python.org/3/library/tkinter.html>
- PyQt: <https://doc.qt.io/qtforpython/>
- Os: <https://docs.python.org/3/library/os.html>
- DateTime: <https://docs.python.org/3/library/datetime.html>
- Random: <https://docs.python.org/3/library/random.html>

# MÓDULO 01 - CLASE 02

Python:  
(Sentencias - Bucles - Operadores  
- Funciones - Bibliotecas)



# REPASO Clase 1

MODULO 01. NIVELATORIO

CLASE 02. Python (II)

01

.UBA económicas  
*Diplomaturas*

**IA** aplicada a  
Entornos Digitales  
de Gestión



# Google colab

---

## Google Colab Clase 2

[acortar.link/clase2python](https://acortar.link/clase2python)



MODULO 01. NIVELATORIO

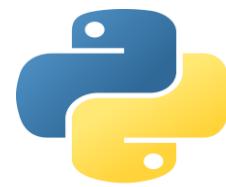
CLASE 02. Python (II)

02

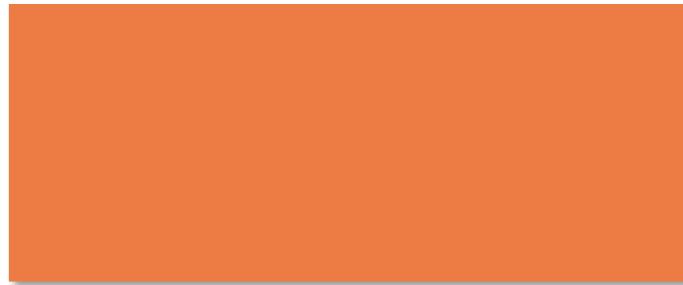
.UBA económicas  
*Diplomaturas*

**IA** aplicada a  
Entornos Digitales  
de Gestión

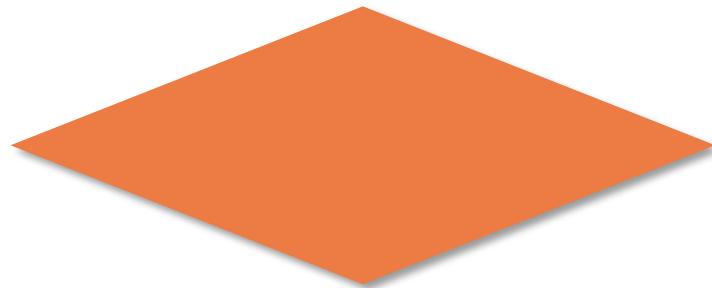




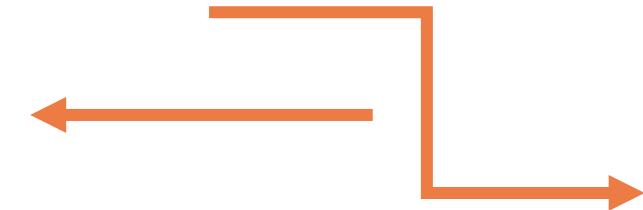
# Sentencias de Control



Indica Acción o Proceso



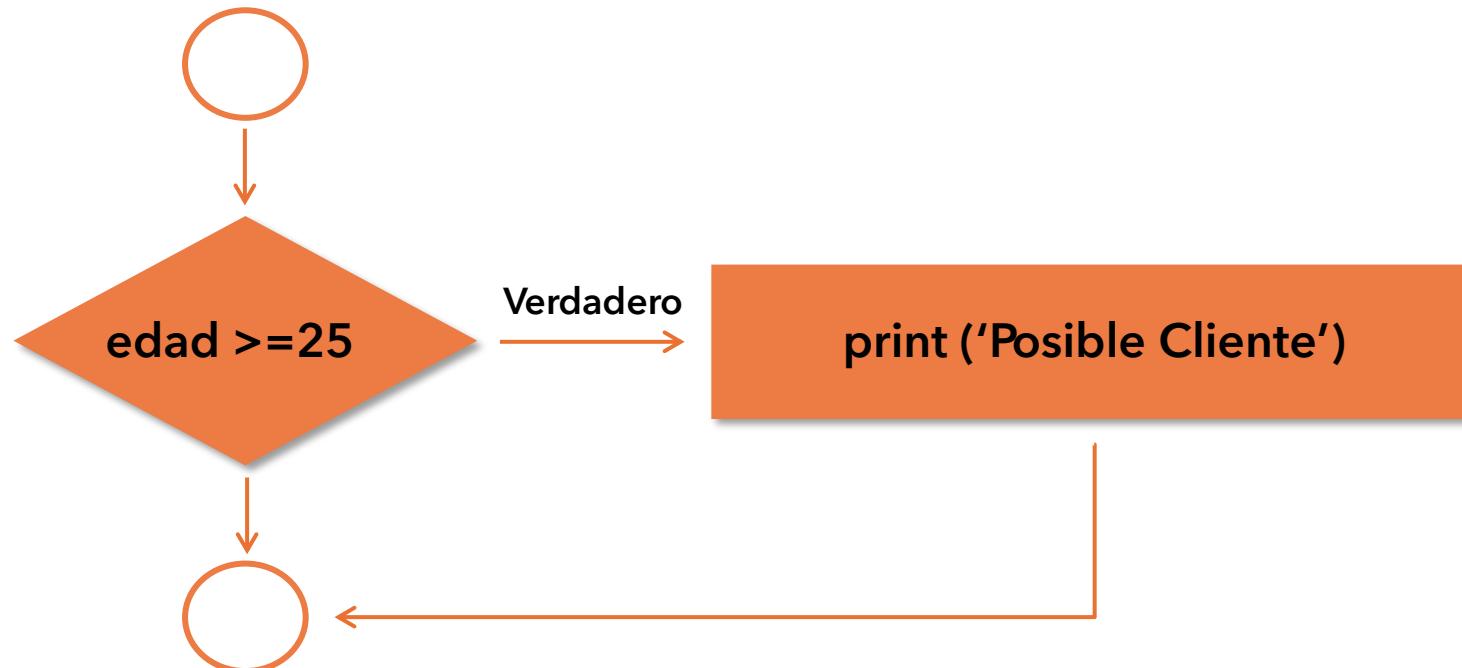
Toma de Decisiones



Flecha que indica la dirección del Flujo a Ejecutarse

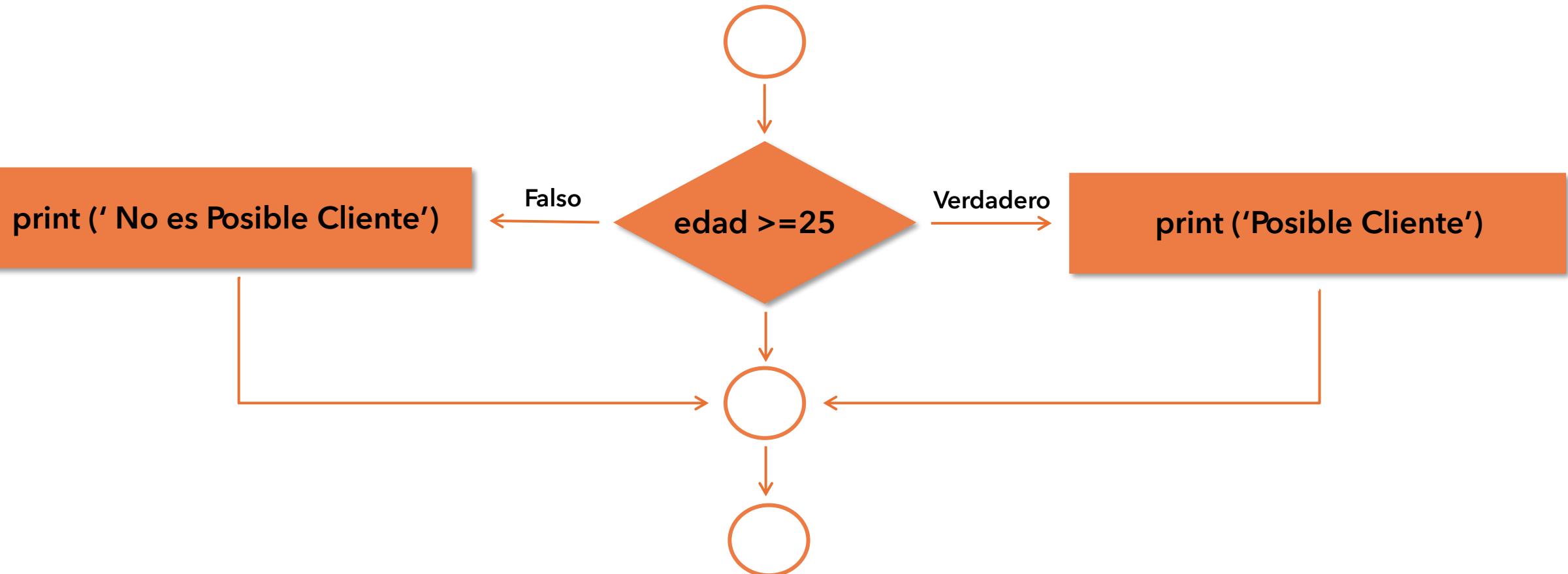


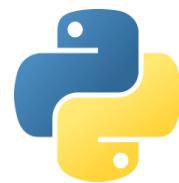
Esta sentencia realiza una acción, cuando la condición es verdadera; y la ignora, cuando es falsa.



# if...else

Esta sentencia realiza una acción, cuando la condición es verdadera; y realiza la otra acción cuando es falsa.



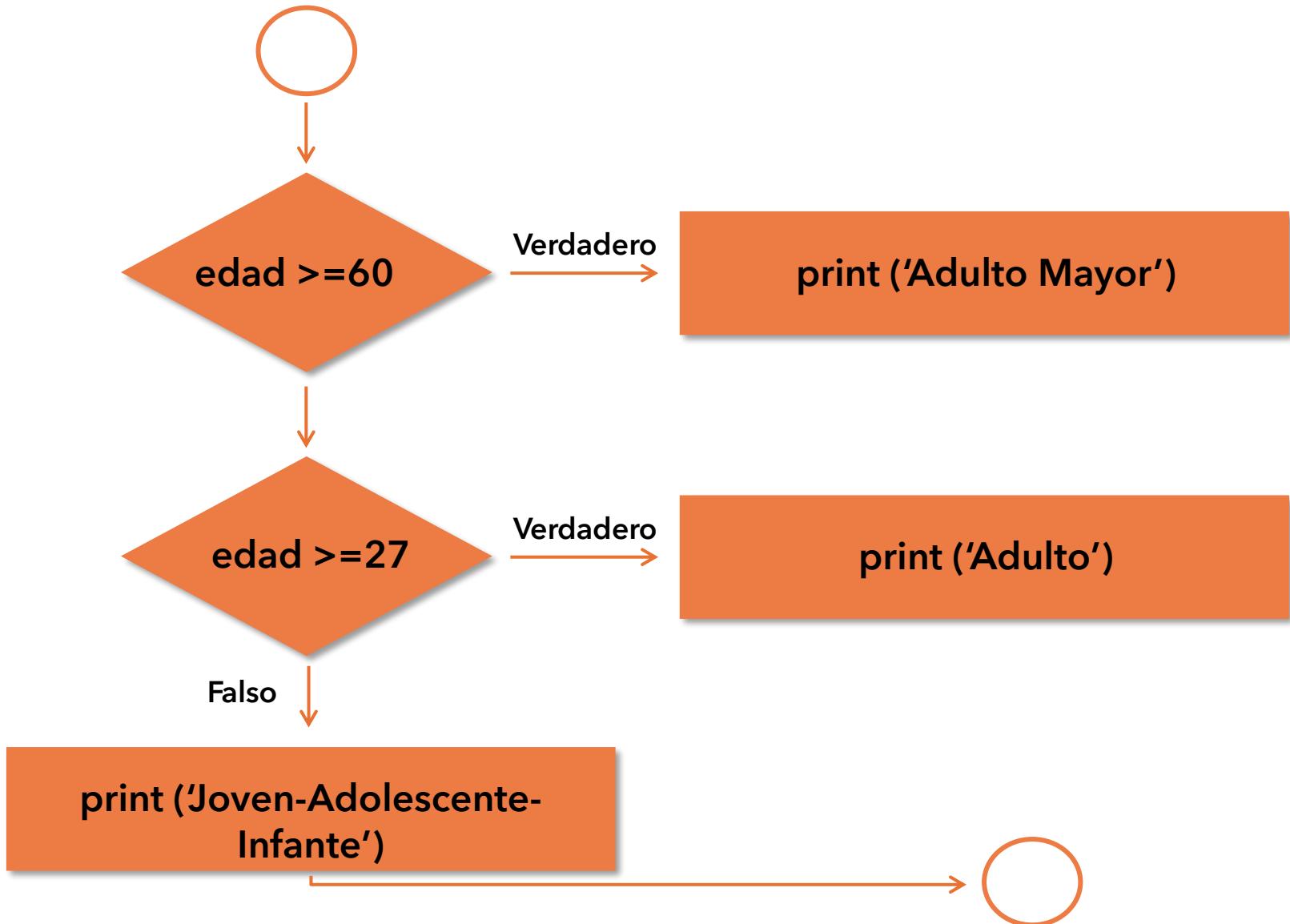


# if...elif ...else

Esta sentencia realiza una de diferentes acciones dependiendo como hayan sido calificadas las otras condiciones.

Se utiliza en estructuras condicionales para evaluar múltiples condiciones secuenciales después de una instrucción if, pero antes de la instrucción else.







miro.com



MODULO 01. NIVELATORIO

CLASE 02. Python (II)

08

.UBA económicas  
*Diplomaturas*

IA aplicada a  
Entornos Digitales  
de Gestión



# Ejercicios Clase 1

MODULO 01. NIVELATORIO

CLASE 02. Python (II)

09

.UBA económicas  
*Diplomaturas*

**IA** aplicada a  
Entornos Digitales  
de Gestión





# Sentencia while

---

Esta sentencia permite realizar una acción mientras que una condición se mantenga como verdadera.

Con esta sentencia se forma un lazo, o bucle, o ciclo (loop) que repite las instrucciones hasta cumplir cierta condición.





# Sentencia for

Esta sentencia repite una o varias acciones.

La sentencia for realiza la acción para cada elemento en una sucesión de elementos.

Tal sucesión de elementos puede ser una lista de valores, que no necesariamente debe estar ordenada o pueden ser caracteres, ya que una palabra es una lista de caracteres.





# Sintaxis for

**for variable in iterable:**

```
# Bloque de código a ejecutar en cada iteración  
# Podes usar 'variable' dentro de este bloque
```

- **variable:** Es el nombre de la variable que tomará el valor de cada elemento en el iterable en cada iteración del bucle. Suele llamarse "i"
- **iterable:** Es una secuencia de elementos que pueden ser recorridos. Puede ser una lista, tupla, cadena de texto, conjunto, diccionario, etc.
- **Bloque de código a ejecutar en cada iteración:** Es el conjunto de instrucciones que se ejecutarán en cada iteración del bucle. Estas instrucciones deben estar indentadas respecto a la declaración for.





# Asignadores de Aumento

Operacion	Con el asignador
<code>a=a+8</code>	<code>a+=8</code>
<code>b=b-1</code>	<code>b-=1</code>
<code>c=c*4</code>	<code>c*=4</code>
<code>d=d**2</code>	<code>d**=2</code>
<code>e=e/5</code>	<code>e/=5</code>
<code>f=f//2</code>	<code>f//=2</code>
<code>g=g%3</code>	<code>g%=3</code>





# range()

---

Es una función incorporada en Python que genera una secuencia de números enteros en un rango especificado. Puede tomar uno, dos o tres argumentos.





# Sintaxis range()

**range(stop)**

**range(start, stop)**

**range(start, stop, step)**

- **start (opcional):** El valor inicial de la secuencia. Por defecto es 0.
- **stop (requerido):** El valor final de la secuencia. El rango no incluye este valor.
- **step (opcional):** El tamaño del paso entre cada número en la secuencia. Por defecto es 1.



# Operadores Booleanos





## Tabla de verdad para el operador lógico and

Proposición 1	Proposición 2	Proposición 1 y Proposición 2
V	V	V
V	F	F
F	V	F
F	F	F





or

## Tabla de verdad para el operador lógico or

Proposición 1	Proposición 2	Proposición 1 y Proposición 2
V	V	V
V	F	V
F	V	V
F	F	F



# Funciones



```
def nombre_de_la_funcion(argumento1, argumento2, ...):  
    # Bloque de código que define la función  
    # Puede incluir cualquier número de declaraciones  
    # Puede tener un retorno opcional utilizando la palabra clave "return"
```

- **nombre\_de\_la\_funcion:** Es el nombre que le das a tu función. Debe seguir las mismas reglas que los nombres de las variables en Python.
- **argumento1, argumento2, ...:** Son los argumentos que la función puede aceptar opcionalmente. Puedes tener cero o más argumentos. Estos se utilizan para pasar valores a la función que la función puede usar internamente.
- **return:** La palabra clave return se utiliza para devolver un valor desde la función. Esto es opcional; no todas las funciones necesitan devolver un valor.



# Biblioteca/Libreria/Modulo en Python





# Biblioteca

En Python, una "biblioteca" (también conocida como "librería" o "módulo") se refiere a un conjunto de funciones y métodos predefinidos que se pueden utilizar para realizar tareas específicas. Estas bibliotecas están diseñadas para ser reutilizables y pueden incluir una variedad de funcionalidades, desde operaciones matemáticas hasta manipulación de archivos, acceso a bases de datos, creación de interfaces gráficas de usuario y mucho más.





# Biblioteca

## Las Bibliotecas más populares:

- ❑ **NumPy**: Para cálculos numéricos y manipulación de matrices.
- ❑ **Pandas**: Para el análisis y manipulación de datos estructurados.
- ❑ **Matplotlib y Seaborn**: Para la creación de gráficos y visualización de datos.
- ❑ **Requests**: Para realizar solicitudes HTTP a través de la web.
- ❑ **Django y Flask**: Para el desarrollo web.
- ❑ **TensorFlow y PyTorch**: Para el aprendizaje profundo y la inteligencia artificial.
- ❑ **Tkinter y PyQt**: Para crear interfaces gráficas de usuario (GUI).
- ❑ **Os**: Para funciones relacionadas con el sistema operativo y archivos.
- ❑ **DateTime**: Para manipulación de fechas y horas.
- ❑ **Random**: Para generación de números aleatorios.

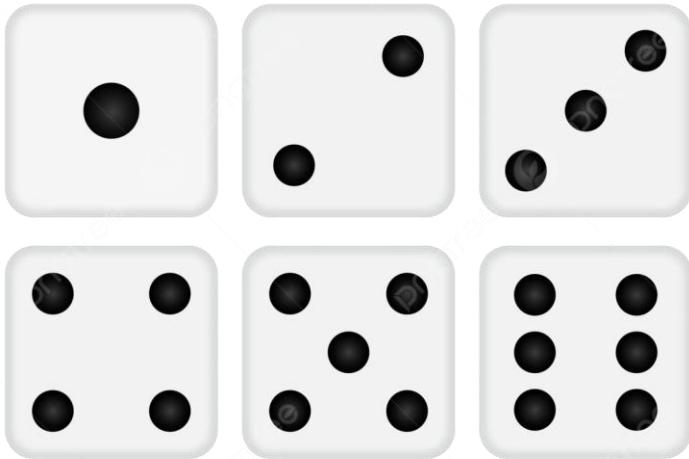




# RANDOM

La biblioteca *random* de Python proporciona herramientas para trabajar con números aleatorios y realizar operaciones relacionadas con la aleatoriedad, como la generación de números aleatorios, la selección de elementos al azar y la mezcla de secuencias.





# Visualización de datos estáticos





# Lanzar 1 dado

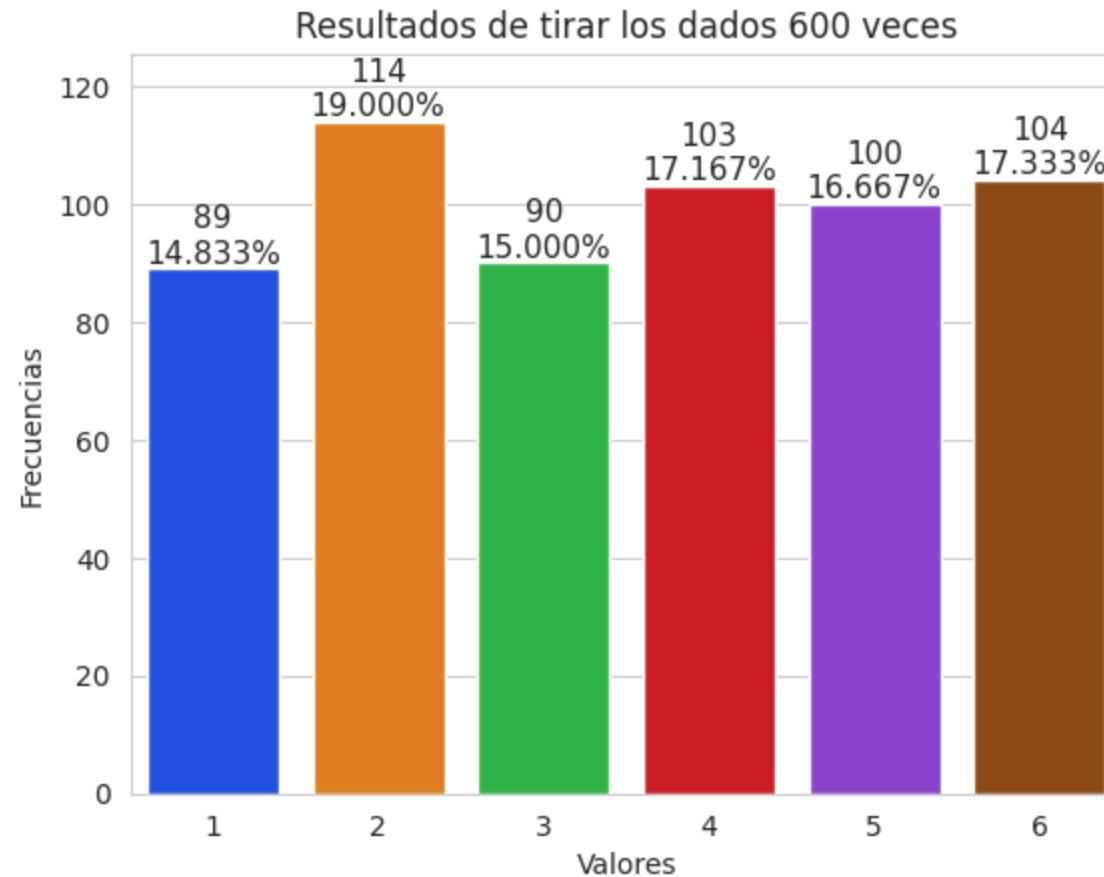
$$\text{Probabilidad (Evento)} = \frac{\text{Sucesos Favorables}}{\text{Sucesos Posibles}}$$

Evento (Resultado al Lanzar el Dado)	Probabilidad
1	$\frac{1}{6} = 0,1\bar{6} \rightarrow \text{Aprox } 16,6\%$
2	$\frac{1}{6} = 0,1\bar{6} \rightarrow \text{Aprox } 16,6\%$
3	$\frac{1}{6} = 0,1\bar{6} \rightarrow \text{Aprox } 16,6\%$
4	$\frac{1}{6} = 0,1\bar{6} \rightarrow \text{Aprox } 16,6\%$
5	$\frac{1}{6} = 0,1\bar{6} \rightarrow \text{Aprox } 16,6\%$
6	$\frac{1}{6} = 0,1\bar{6} \rightarrow \text{Aprox } 16,6\%$





# Lanzar 1 dado 600 veces





# Para pensar

**Comparación de Números:** Pide al usuario que ingrese dos números y luego imprime el mayor de los dos.





# Para pensar

**Tabla de Multiplicar:** Pide al usuario ingresar un número y muestra su tabla de multiplicar del 1 al 10





# Para pensar

**Contador de Vocales:** Sigue el siguiente código para que el usuario ingrese una palabra y cuente cuántas vocales tiene.





# Para pensar

**Validación de Contraseña:** Crea un programa que solicite al usuario ingresar una contraseña. Si la longitud de la contraseña es menor que 8 caracteres, imprime un mensaje de error. Si la longitud es adecuada, imprime un mensaje de confirmación.



**PARTE I**

**CLASE**

**3**

# MÓDULO 01 - CLASE 03

## INTRODUCCIÓN A LOS MODELOS IA (III)

.UBA económicas  
*Diplomaturas*

IA aplicada a  
Entornos Digitales  
de Gestión



# 2020 - 2024

## ¡La IA en todos lados!

MODULO 01. NIVELATORIO

CLASE 03. INTRODUCCIÓN A LOS MODELOS IA (III)

2

.UBA económicas  
*Diplomaturas*

**IA** aplicada a  
Entornos Digitales  
de Gestión

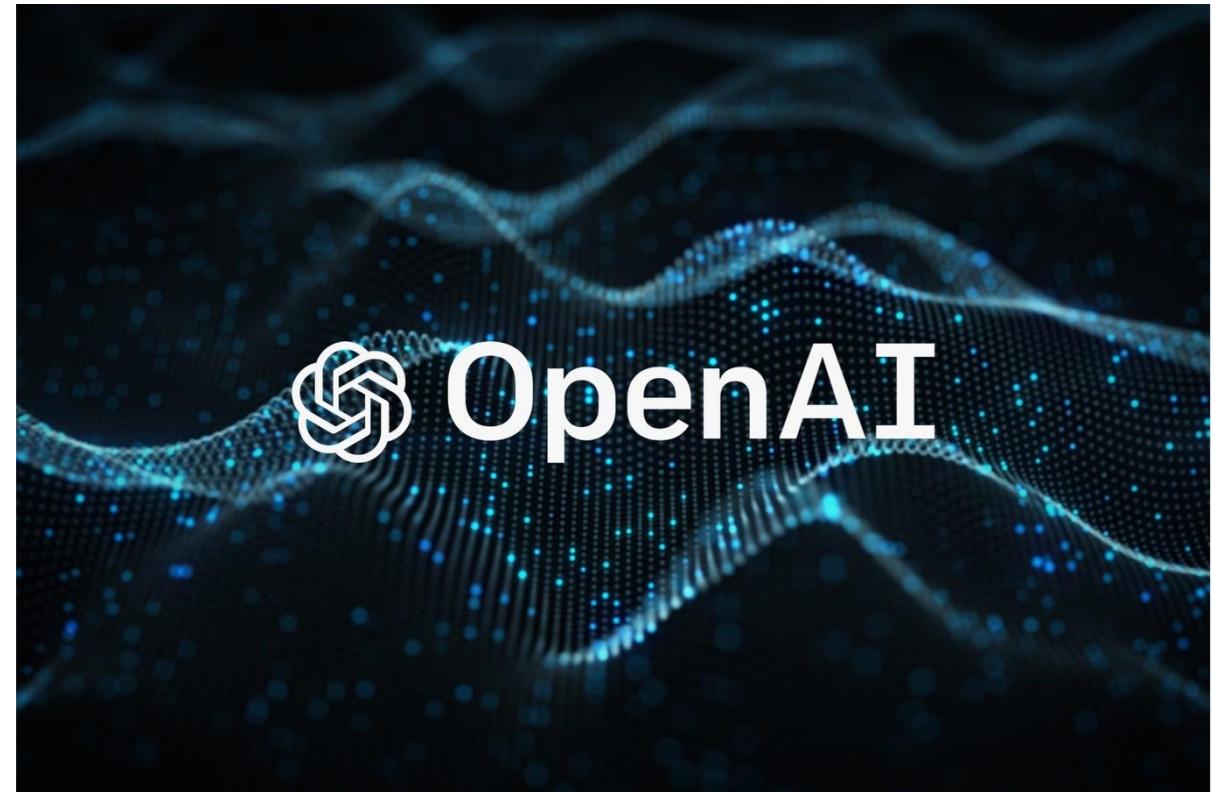


# 2020

## GPT-3 cobra vida

GPT-3, lanzado por OpenAI en junio de 2020, es un modelo de lenguaje avanzado con 175 mil millones de parámetros que puede generar texto coherente y contextual en respuesta a las entradas del usuario.

Este modelo ha sido utilizado en diversas aplicaciones, como asistentes virtuales, generación de contenido, y soporte al cliente, destacándose por su capacidad para realizar tareas complejas de procesamiento de lenguaje natural.



**GPT-3 El modelo LLM que generó la tendencia.**

**MODULO 01. NIVELATORIO**

**CLASE 03. INTRODUCCIÓN A LOS MODELOS IA (III)**

**3**

**.UBA** económicas  
*Diplomaturas*

**IA** aplicada a  
Entornos Digitales  
de Gestión



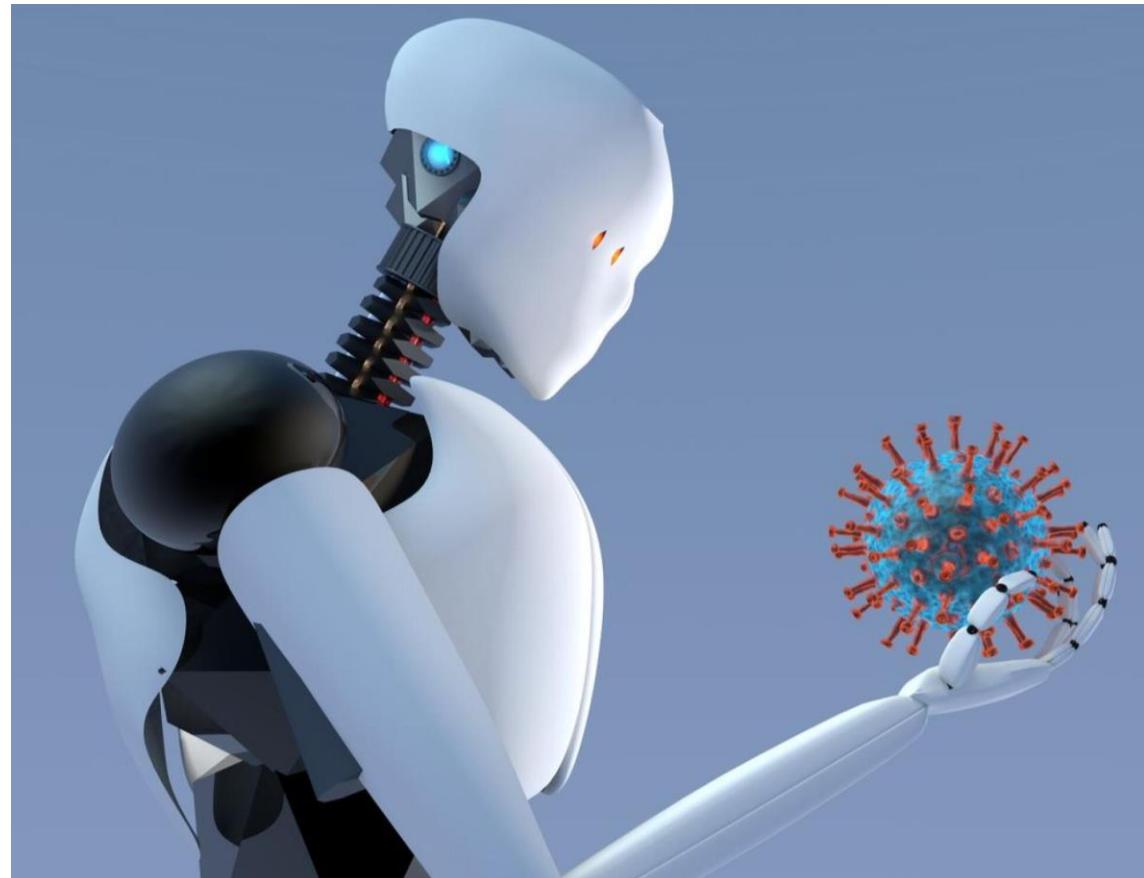
# 2020

## Estructuras Protéicas con IA

AlphaFold 2, presentado por DeepMind en diciembre de 2020, revolucionó la predicción de estructuras proteicas con una precisión comparable a métodos experimentales tradicionales, gracias a su avanzado modelo de redes neuronales profundas.

Este avance ha acelerado significativamente la investigación científica y el desarrollo de medicamentos, reduciendo costos y tiempos asociados a técnicas experimentales.

Particularmente, facilitó la investigación de vacunas contra el COVID-19.



La IA facilitó la lucha contra el COVID-19



# 2021

## La IA crea imágenes.

DALL-E 1 es un modelo de inteligencia artificial desarrollado por OpenAI que genera imágenes a partir de descripciones textuales. Utiliza una versión de GPT-3 para interpretar el texto y producir imágenes coherentes y detalladas basadas en la entrada proporcionada.

Fue lanzado en enero de 2021, marcando un avance significativo en la generación de imágenes mediante IA.



Prompt: "An elephant made of rubik's cube"



# 2021

## LoRA una futura revolución.

LoRA (Low-Rank Adaptation) es una técnica que permite ajustar grandes modelos de lenguaje de manera eficiente utilizando matrices de bajo rango, reduciendo la necesidad de reentrenar todo el modelo. Esto permite adaptar modelos preentrenados a nuevas tareas con menos datos y recursos computacionales, manteniendo un alto rendimiento.

Fue presentada en el documento original "LoRA: Low-Rank Adaptation of Large Language Models" por investigadores de Microsoft en 2021.

### LoRA: LOW-RANK ADAPTATION OF LARGE LANGUAGE MODELS

Edward Hu\* Yelong Shen\* Phillip Wallis Zeyuan Allen-Zhu  
Yuanzhi Li Shean Wang Lu Wang Weizhu Chen

Microsoft Corporation

{edwardhu, yeshe, phwallis, zeyuana, yuanzhil, swang, luw, wzchen}@microsoft.com  
yuanzhil@andrew.cmu.edu

(Version 2)

#### ABSTRACT

An important paradigm of natural language processing consists of large-scale pre-training on general domain data and adaptation to particular tasks or domains. As we pre-train larger models, full fine-tuning, which retrains all model parameters, becomes less feasible. Using GPT-3 175B as an example – deploying independent instances of fine-tuned models, each with 175B parameters, is prohibitively expensive. We propose **Low-Rank Adaptation**, or LoRA, which freezes the pre-trained model weights and injects trainable rank decomposition matrices into each layer of the Transformer architecture, greatly reducing the number of trainable parameters for downstream tasks. Compared to GPT-3 175B fine-tuned with Adam, LoRA can reduce the number of trainable parameters by 10,000 times and the GPU memory requirement by 3 times. LoRA performs on-par or better than fine-tuning in model quality on RoBERTa, DeBERTa, GPT-2, and GPT-3, despite having fewer trainable parameters, a higher training throughput, and, unlike adapters, *no additional inference latency*. We also provide an empirical investigation into rank-deficiency in language model adaptation, which sheds light on the efficacy of LoRA. We release a package that facilitates the integration of LoRA with PyTorch models and provide our implementations and model checkpoints for RoBERTa, DeBERTa, and GPT-2 at <https://github.com/microsoft/LoRA>.

LoRA ajusta LLMs usando matrices de bajo rango



# 2021

## ¿Poderoso? el chiquitín.

El modelo RETRO (Retrieval-Enhanced Transformer) de DeepMind combina aprendizaje profundo con técnicas de recuperación de información para mejorar el rendimiento en procesamiento de lenguaje natural, accediendo a grandes bases de datos de texto para generar respuestas precisas y contextualizadas. Presentado en diciembre de 2021, ha demostrado, en pruebas de laboratorio, ser más eficiente que los modelos tradicionales de gran escala. No fue liberado al público en general.



### Improving language models by retrieving from trillions of tokens

Sebastian Borgeaud<sup>†</sup>, Arthur Mensch<sup>†</sup>, Jordan Hoffmann<sup>†</sup>, Trevor Cai, Eliza Rutherford, Katie Millican, George van den Driessche, Jean-Baptiste Lespiau, Bogdan Damoc, Aidan Clark, Diego de Las Casas, Aurelia Guy, Jacob Menick, Roman Ring, Tom Hennigan, Saffron Huang, Loren Maggiore, Chris Jones, Albin Cassirer, Andy Brock, Michela Paganini, Geoffrey Irving, Oriol Vinyals, Simon Osindero, Karen Simonyan, Jack W. Rae<sup>‡</sup>, Erich Elsen<sup>‡</sup> and Laurent Sifre<sup>†,‡</sup>

All authors from DeepMind, <sup>†</sup>Equal contributions, <sup>‡</sup>Equal senior authorship

We enhance auto-regressive language models by conditioning on document chunks retrieved from a large corpus, based on local similarity with preceding tokens. With a 2 trillion token database, our Retrieval-Enhanced Transformer (RETRO) obtains comparable performance to GPT-3 and Jurassic-1 on the Pile, despite using  $25\times$  fewer parameters. After fine-tuning, RETRO performance translates to downstream knowledge-intensive tasks such as question answering. RETRO combines a frozen BERT retriever, a differentiable encoder and a chunked cross-attention mechanism to predict tokens based on an order of magnitude more data than what is typically consumed during training. We typically train RETRO from scratch, yet can also rapidly Retrofit pre-trained transformers with retrieval and still achieve good performance. Our work opens up new avenues for improving language models through explicit memory at unprecedented scale.

El paper fue presentado el 8/12/2021



# 2022

## DALL-E 2. El artista mejora.

DALL-E 2, anunciado por OpenAI en abril de 2022, tiene aproximadamente 3.5 mil millones de parámetros, lo que es significativamente menos que los 12 mil millones de DALL-E 1.

Esta versión mejorada del modelo original DALL-E 1 presenta capacidades avanzadas de generación de imágenes a partir de descripciones textuales, produciendo imágenes de mayor calidad y con más coherencia en relación a las entradas de texto proporcionadas.

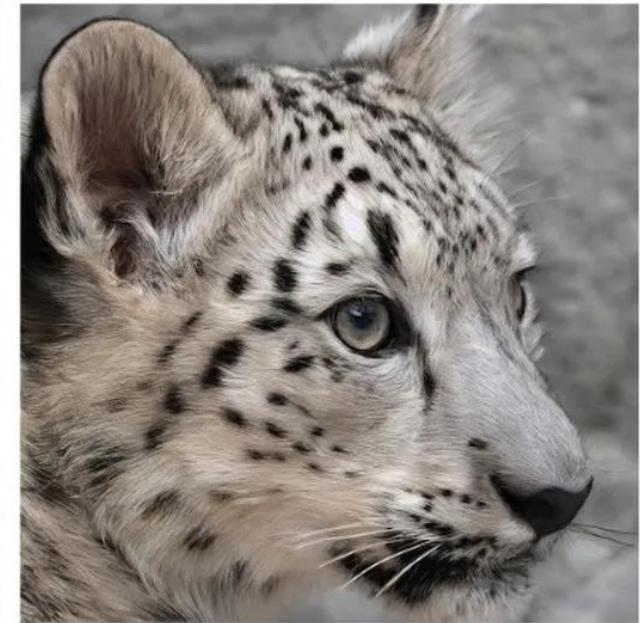


# 2022

## Stable Diffusion 2 y el Open Source.

Stable Diffusion 2, lanzado en noviembre de 2022, es una versión mejorada del modelo de difusión que genera imágenes de alta calidad a partir de descripciones textuales, manteniendo coherencia y detalle visual.

Es de código abierto y puede ejecutarse localmente, lo que ha generado una revolución al hacer accesible esta tecnología avanzada a una amplia audiencia.



Stable Diffusion 2 también ofrece reescalamiento de imágenes.

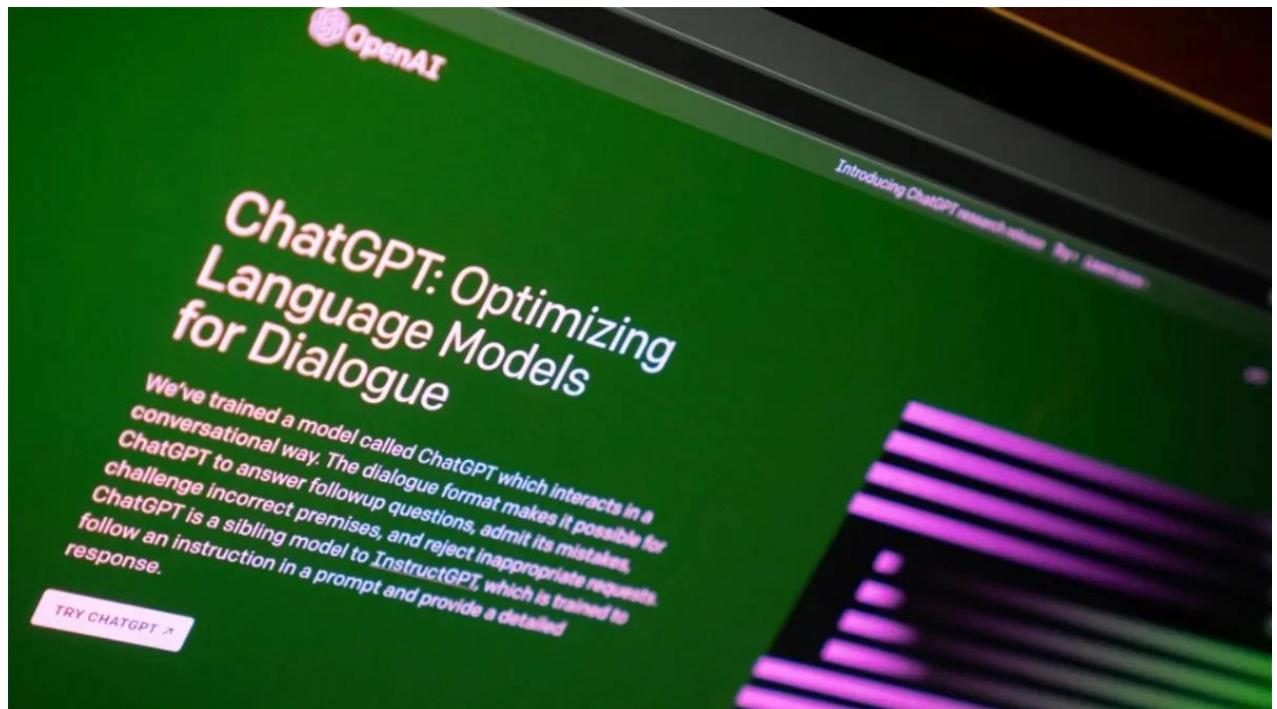


# 2022

## Con Chat GPT la IA “explotó”.

ChatGPT, lanzado por OpenAI el 30 de noviembre de 2022, es un modelo de lenguaje basado en la arquitectura GPT-3.5, diseñado para interactuar de manera conversacional.

Ofrece respuestas coherentes y detalladas a las consultas de los usuarios, mejorando la experiencia de interacción con inteligencia artificial. Su lanzamiento revolucionó la forma en que las personas interactúan con la tecnología.



ChatGPT la aplicación web de OpenAI optimizada para el diálogo.

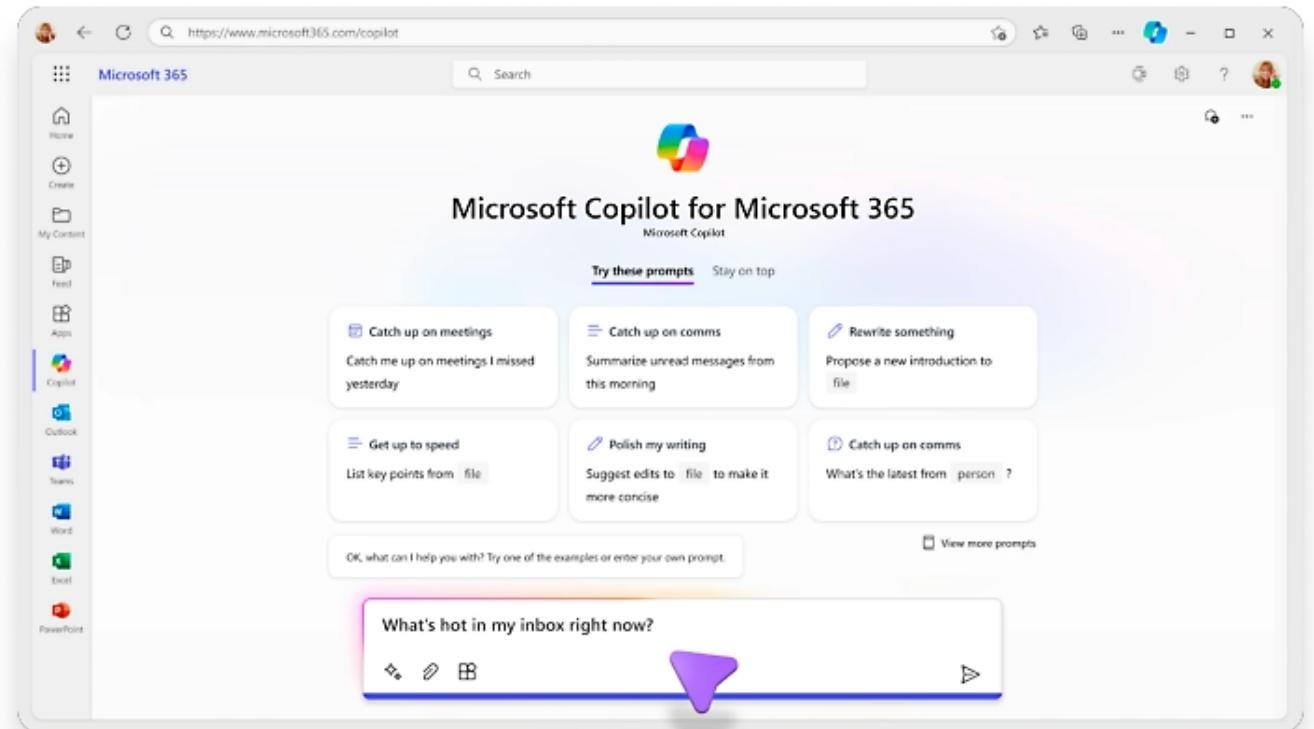


# 2023

## Microsoft se asocia a OpenAI.

A fines de enero del 2023, Microsoft anuncia su asociación con OpenAI con el objetivo de implementar su tecnología tanto en sus servicios online como en sus sistemas operativos.

A raíz de esto se facilita el acceso web de ChatGPT utilizando tecnología BING, además de mejorar los servicios de la plataforma, con una plataforma, primero denominada CHATBing y luego Microsoft Copilot.



Microsoft alimenta su servicio de IA generativa con tecnología OpenAI



# 2023

## Google lanza Bard... para ganar tiempo.

En Febrero 2023 Google anuncia su respuesta a ChatGPT, Bard. En su presentación brindó información incorrecta sobre el telescopio espacial James Webb. También enfrentó una fuerte competencia de ChatGPT de OpenAI, respaldado por Microsoft, que ya había ganado una sólida presencia en el mercado. Bard tuvo problemas de usabilidad y confiabilidad, como la incapacidad de mantener contextos en conversaciones largas y proporcionar respuestas incorrectas con excesiva confianza, lo que lo llevó al fracaso y a la búsqueda de un reemplazo.

Microsoft alimenta su servicio de IA generativa con tecnología OpenAI

MODULO 01. NIVELATORIO

CLASE 03. INTRODUCCIÓN A LOS MODELOS IA (III)

12

.UBA económicas  
Diplomaturas

IA  
aplicada a  
Entornos Digitales  
de Gestión



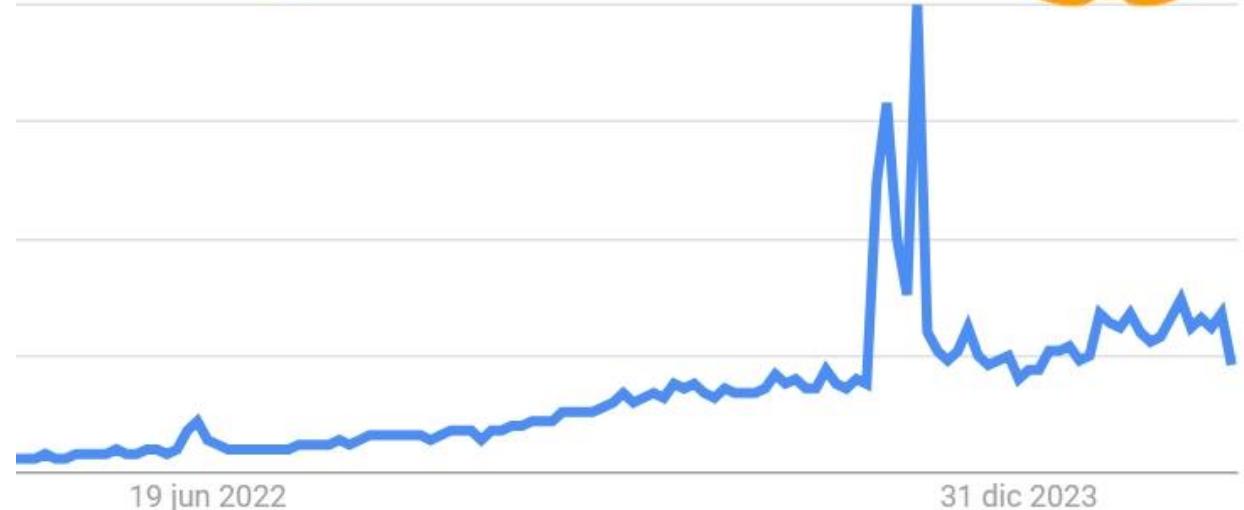
# 2023

Hugging Face se vuelve el lugar de encuentro.

El 21 de febrero de 2023, Hugging Face (HF) y Amazon Web Services (AWS) anuncian una colaboración para acelerar la adopción de aplicaciones de inteligencia artificial generativa, facilitando a los desarrolladores el acceso a modelos avanzados de HF a través de AWS.

Esta asociación ha permitido que una comunidad más amplia de desarrolladores pueda beneficiarse de las últimas innovaciones en inteligencia artificial a un costo accesible.

Google Trends



HF se vuelve tendencia cuando se lanzan modelos LLM Open Source



# 2023

## Modelos LLM Open Source Everywhere.

Durante el 2022 (y sobre todo el 2023 con la aparición de LLaMA), varias organizaciones y comunidades lanzaron modelos de lenguaje de gran escala (LLM) accesibles al público, lo que democratizó el acceso a tecnologías avanzadas de IA y fomentó la innovación y colaboración en la comunidad de desarrollo de inteligencia artificial. Esta tendencia ha permitido a más personas experimentar, mejorar y aplicar estos modelos en una variedad de aplicaciones prácticas.

### LLaMA: Open and Efficient Foundation Language Models

Hugo Touvron\*, Thibaut Lavril\*, Gautier Izacard\*, Xavier Martinet  
Marie-Anne Lachaux, Timothee Lacroix, Baptiste Rozière, Naman Goyal  
Eric Hambro, Faisal Azhar, Aurelien Rodriguez, Armand Joulin  
Edouard Grave\*, Guillaume Lample\*

Meta AI

#### Abstract

We introduce LLaMA, a collection of foundation language models ranging from 7B to 65B parameters. We train our models on trillions of tokens, and show that it is possible to train state-of-the-art models using publicly available datasets exclusively, without resorting to proprietary and inaccessible datasets. In particular, LLaMA-13B outperforms GPT-3 (175B) on most benchmarks, and LLaMA-65B is competitive with the best models, Chinchilla-70B and PaLM-540B. We release all our models to the research community<sup>1</sup>.

#### 1 Introduction

Large Languages Models (LLMs) trained on massive corpora of texts have shown their ability to per-

formance, a smaller one trained longer will ultimately be cheaper at inference. For instance, although Hoffmann et al. (2022) recommends training a 10B model on 200B tokens, we find that the performance of a 7B model continues to improve even after 1T tokens.

The focus of this work is to train a series of language models that achieve the best possible performance at various inference budgets, by training on more tokens than what is typically used. The resulting models, called *LLaMA*, ranges from 7B to 65B parameters with competitive performance compared to the best existing LLMs. For instance, LLaMA-13B outperforms GPT-3 on most benchmarks, despite being 10× smaller. We believe that this model will help democratize the access and

El paper técnico de LlaMA se publicó el 27/2/2023.



# 2023

## GPT 4 nace el rey.

Lanzado el 14 de marzo del 2023. GPT-4 es considerado el "rey" debido a su avanzada comprensión y generación de lenguaje, su mayor escala y eficiencia, y su capacidad para manejar tareas complejas con respuestas coherentes y contextuales. Además, sus mejoras en seguridad y ética, junto con su versatilidad en una amplia gama de aplicaciones, lo posicionan como una herramienta líder en inteligencia artificial.

## GPT-4 Technical Report

OpenAI\*

### Abstract

We report the development of GPT-4, a large-scale, multimodal model which can accept image and text inputs and produce text outputs. While less capable than humans in many real-world scenarios, GPT-4 exhibits human-level performance on various professional and academic benchmarks, including passing a simulated bar exam with a score around the top 10% of test takers. GPT-4 is a Transformer-based model pre-trained to predict the next token in a document. The post-training alignment process results in improved performance on measures of factuality and adherence to desired behavior. A core component of this project was developing infrastructure and optimization methods that behave predictably across a wide range of scales. This allowed us to accurately predict some aspects of GPT-4's performance based on models trained with no more than 1/1,000th the compute of GPT-4.

GPT 4 superó en todos los aspectos a su antecesor.



# 2023

## La niñera contra la IA.

La huelga de actores y guionistas de Hollywood, iniciada en julio de 2023, surgió debido a preocupaciones sobre el uso de la IA en la creación de contenido y la representación de actores. Los artistas temían que la IA pudiera reemplazarlos, amenazando su supervivencia en la industria. Tras el acuerdo alcanzado en noviembre de 2023 incluyó aumentos salariales, una mayor proporción de ingresos por transmisión y protecciones contra el uso ilimitado de la IA, permitiendo que la producción se reanude de manera responsable y ética.



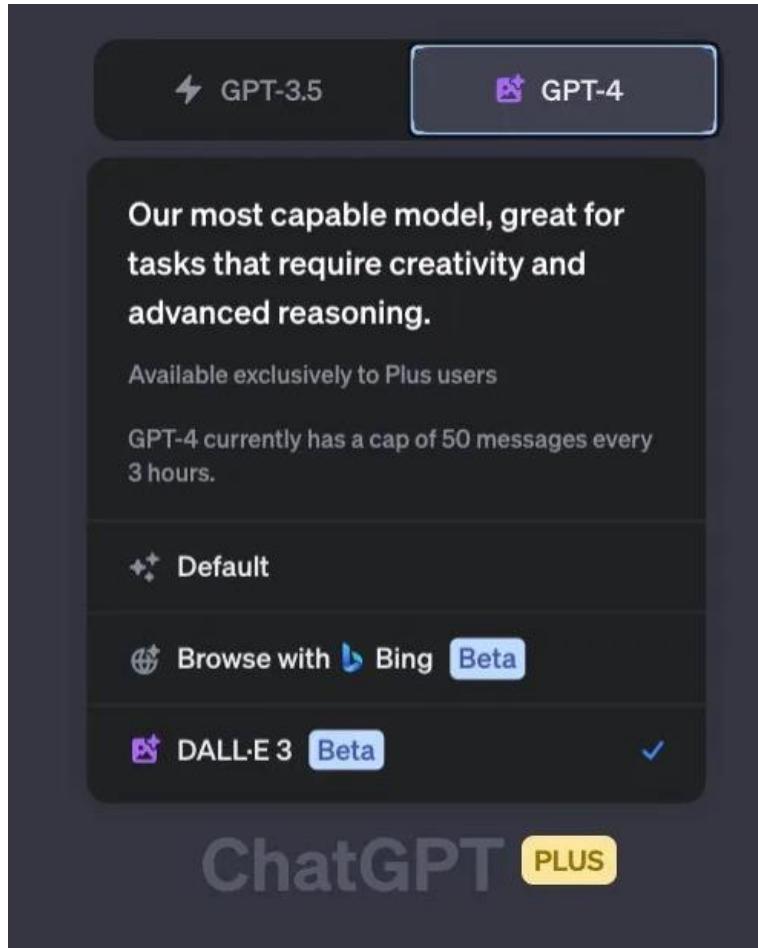
Fran Dreschler, se convirtió en la cara visible del reclamo.



# 2023

## ChatGPT se vuelve multimodal

La integración de DALL-E 3 con ChatGPT se anunció en septiembre de 2023 y se implementó en octubre del mismo año. Esta integración permite a los usuarios de ChatGPT Plus y Enterprise crear imágenes a partir de descripciones textuales directamente en la interfaz de ChatGPT, sin necesidad de utilizar complementos adicionales. Esta mejora facilita la creación de arte digital, haciendo que sea accesible incluso para aquellos que no tienen experiencia en ingeniería de prompts.



La integración de ChatGPT con DALL-E 3 inició la era de la multimodalidad.

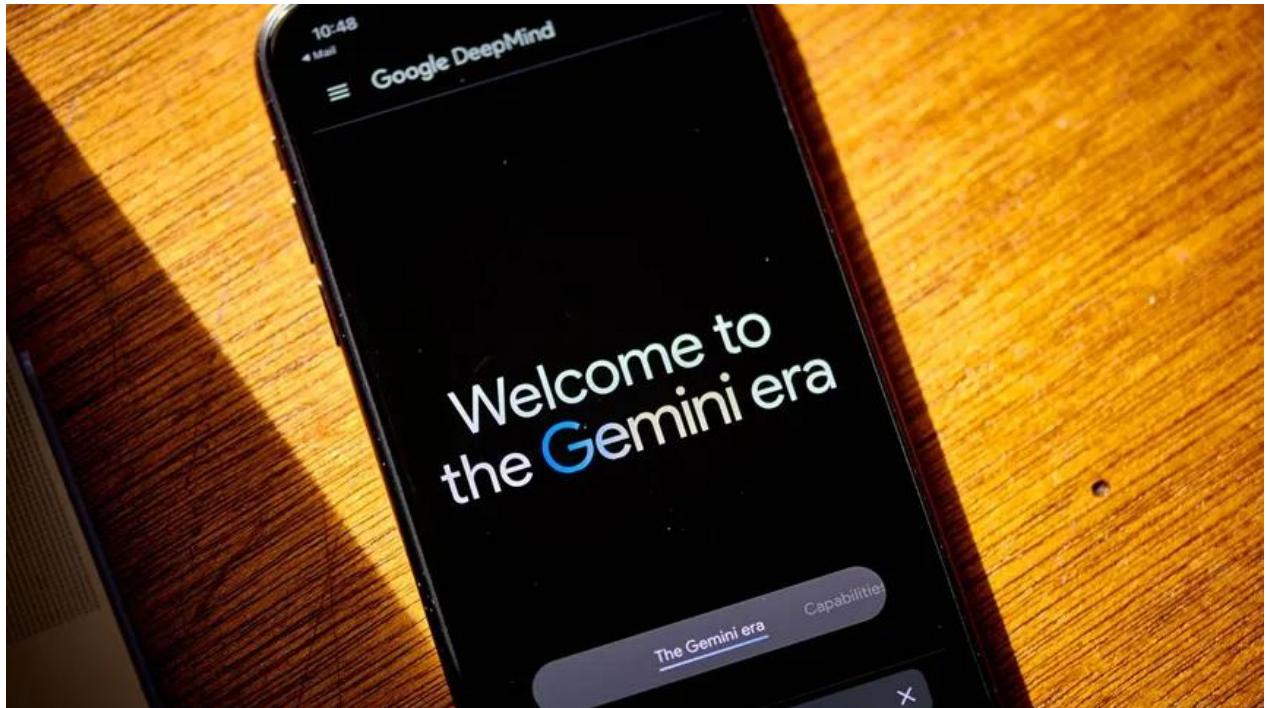


# 2023

## Google lanza Gemini

A mediados de diciembre del 2023, luego del fracaso de su modelo BARD, Google se apresura y disponibiliza Gemini (en estado Beta).

Según Google, Gemini cuenta con habilidades de programación avanzadas, incluida la generación de código de alta calidad mediante AlphaCode 2, un sistema avanzado de generación de código. También puede resolver problemas de programación complejos y colaborar con desarrolladores.



Google cerró el 2023 con el lanzamiento de Gemini, su modelo IA.



# Seguimos la próxima



# 2024

## ¡La IA en todos lados!

MODULO 01. NIVELATORIO

CLASE 03. INTRODUCCIÓN A LOS MODELOS IA (III)

20

.UBA económicas  
*Diplomaturas*

**IA** aplicada a  
Entornos Digitales  
de Gestión



# MÓDULO 01 - CLASE 03

## Biblioteca - Datos - Listas - Tuplas



# Google colab

---

## Google Colab Clase 2

[acortar.link/clase2python](https://acortar.link/clase2python)



MODULO 01. NIVELATORIO

CLASE 03. Python (III)

01

.UBA económicas  
*Diplomaturas*

**IA** aplicada a  
Entornos Digitales  
de Gestión



# Biblioteca/Libreria/Modulo en Python





# Biblioteca

En Python, una "biblioteca" (también conocida como "librería" o "módulo") se refiere a un conjunto de funciones y métodos predefinidos que se pueden utilizar para realizar tareas específicas. Estas bibliotecas están diseñadas para ser reutilizables y pueden incluir una variedad de funcionalidades, desde operaciones matemáticas hasta manipulación de archivos, acceso a bases de datos, creación de interfaces gráficas de usuario y mucho más.





# Biblioteca

## Las Bibliotecas más populares

- NumPy: Para cálculos numéricos y manipulación de matrices.
- Pandas: Para el análisis y manipulación de datos estructurados.
- Matplotlib y Seaborn: Para la creación de gráficos y visualización de datos.
- Requests: Para realizar solicitudes HTTP a través de la web.
- Django y Flask: Para el desarrollo web.
- TensorFlow y PyTorch: Para el aprendizaje profundo y la inteligencia artificial.
- Tkinter y PyQt: Para crear interfaces gráficas de usuario (GUI).
- Os: Para funciones relacionadas con el sistema operativo y archivos.
- DateTime: Para manipulación de fechas y horas.
- Random: Para generación de números aleatorios.

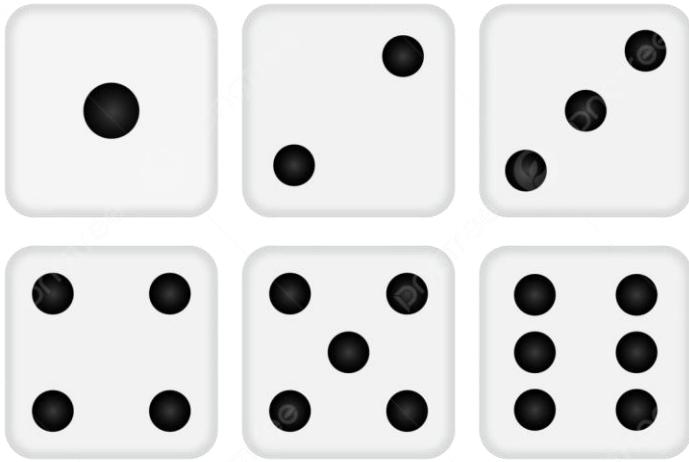




# RANDOM

La biblioteca random de Python proporciona herramientas para trabajar con números aleatorios y realizar operaciones relacionadas con la aleatoriedad, como la generación de números aleatorios, la selección de elementos al azar y la mezcla de secuencias.





# Visualización de datos estáticos





# Lanzar 1 dado

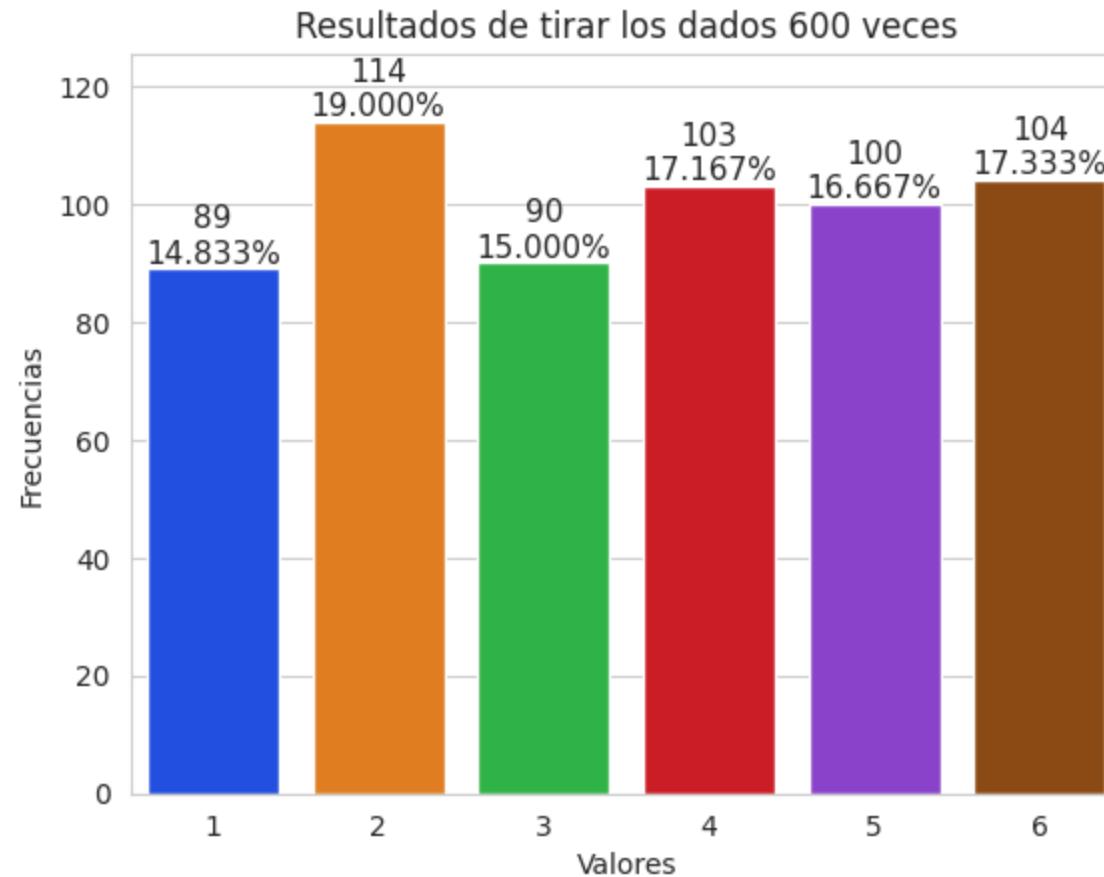
$$\text{Probabilidad (Evento)} = \frac{\text{Sucesos Favorables}}{\text{Sucesos Posibles}}$$

Evento (Resultado al Lanzar el Dado)	Probabilidad
1	$\frac{1}{6} = 0,1\bar{6} \rightarrow \text{Aprox } 16,6\%$
2	$\frac{1}{6} = 0,1\bar{6} \rightarrow \text{Aprox } 16,6\%$
3	$\frac{1}{6} = 0,1\bar{6} \rightarrow \text{Aprox } 16,6\%$
4	$\frac{1}{6} = 0,1\bar{6} \rightarrow \text{Aprox } 16,6\%$
5	$\frac{1}{6} = 0,1\bar{6} \rightarrow \text{Aprox } 16,6\%$
6	$\frac{1}{6} = 0,1\bar{6} \rightarrow \text{Aprox } 16,6\%$





# Lanzar 1 dado 600 veces



# Google colab

---

## Google Colab Clase 3

[acortar.link/clase3python](https://acortar.link/clase3python)



# Listas y Tuplas





# Listas y Tuplas

En términos matemáticos una lista es una matriz o un vector. En términos de programación es un conjunto de elementos que contienen datos.



**Una lista en Python es una colección ordenada y mutable de elementos que pueden ser de diferentes tipos. Las listas se definen utilizando corchetes “[]” y los elementos están separados por comas. Las listas permiten modificar, agregar y eliminar elementos después de su creación.**





# Lista

a[0]	a[1]	a[2]	a[3]	a[4]	a[5]
25	13	15	58	16	87
a[-6]	a[-5]	a[-4]	a[-3]	a[-2]	a[-1]



Una tupla en Python es una colección ordenada pero inmutable de elementos (durante su ejecución). Una vez creada, no se pueden modificar, agregar ni eliminar elementos de una tupla (durante su ejecución). Las tuplas se definen utilizando paréntesis “( )” y los elementos están separados por comas.





# Listas VS Tuplas

## Mutabilidad

- **Lista:** Mutable. Sus elementos se pueden modificar, agregar o eliminar después de la creación.
- **Tupla:** Inmutable. Una vez creada, no se pueden cambiar, agregar ni eliminar sus elementos.





# Listas VS Tuplas

## Sintaxis

- **Lista:** Se define con corchetes [].
- **Tupla:** Se define con paréntesis () .





# Listas VS Tuplas

## Uso

- **Lista:** Se usa cuando se necesita una colección de elementos que pueda cambiar a lo largo del tiempo.
- **Tupla:** Se usa cuando se necesita una colección de elementos que no debe cambiar. Ideal para elementos que deben permanecer constantes, como coordenadas (x, y) o claves de diccionarios.





# Listas Multidimensionales

Una lista de varias dimensiones, también conocida como lista multidimensional o matriz, es una estructura de datos que contiene otras listas como sus elementos. Esto permite organizar datos en forma de tablas, cubos o cualquier otra estructura de múltiples niveles. En Python, una lista multidimensional se implementa como una lista de listas.



**PARTE I**

**CLASE**

**4**

# MÓDULO 01 - CLASE 04

# INTRODUCCIÓN A LOS MODELOS IA (IV)

.UBA económicas  
*Diplomaturas*

**IA** aplicada a  
Entornos Digitales  
de Gestión



# 2024

## ¡La IA en todos lados!

MODULO 01. NIVELATORIO

CLASE 04. INTRODUCCIÓN A LOS MODELOS IA (IV)

2

.UBA económicas  
*Diplomaturas*

**IA** aplicada a  
Entornos Digitales  
de Gestión



# 01-24

## Samsung anuncia Galaxy AI.

Galaxy AI es una suite integral de inteligencia artificial (IA) desarrollada por Samsung para sus dispositivos móviles Galaxy. Combina IA en el dispositivo y en la nube para ofrecer una amplia gama de funciones que mejoran la experiencia del usuario. Fue presentada en el lanzamiento del Samsung Galaxy S24 Ultra, y luego incorporandose a distintos dispositivos de gama alta de la firma.



Imagen de las VISION PRO.



# 02-24

## Apple comercializa las Vision Pro

Presentadas en Junio del 2023, Apple Vision Pro es un dispositivo de computación espacial que permite a los usuarios interactuar con el mundo digital de una manera completamente nueva. Combina la realidad aumentada (AR) y la realidad virtual (VR) para crear una experiencia inmersiva que te permite superponer contenido digital sobre tu entorno físico.



Imagen de las VISION PRO.



# 02-24

## Google presenta Gemini Pro

El gigante tecnológico Google lanzó Gemini 1.5 Pro, un modelo de inteligencia artificial (IA) que puede procesar grandes cantidades de información de una sola vez, incluida una hora de video, once horas de audio, 30.000 líneas de código de programación o más de 700.000 palabras.

El modelo Gemini se desarrolló íntegramente pensando en la multimodalidad con la intención de posicionarse como el LLM nº 1.



# Welcome to the Gemini era

Gemini 1.5 llega hasta el millón de tokens de contexto.



# 02-24

## OpenAI Anuncia Sora.

Sora es un sistema de IA generativa que es capaz de crear videos de hasta un minuto con escenas complejas que contengan múltiples personajes, tipos específicos de movimientos y detalles precisos de los sujetos y el fondo. Por el momento, solo está disponible para expertos y algunos artistas que están participando de un proceso de desarrollo y retroalimentación.



Sora es capaz de crear videos de alta calidad.



# 03-24

## Se aprueba la ley de IA en la UE.

El Parlamento Europeo aprobó el 14 de marzo de 2024, la Ley que regulará la Inteligencia Artificial en la UE. Su objetivo es proteger los derechos fundamentales, la democracia, el Estado de derecho y la sostenibilidad medioambiental frente a la IA que entraña un alto riesgo, impulsando al mismo tiempo la innovación y erigiendo a Europa en líder del sector.



El parlamento europeo aprueba la ley IA



# 03-24

## La ONU se expide sobre la IA.

La Asamblea General de la ONU adoptó por aclamación el jueves una resolución histórica sobre la promoción de sistemas de inteligencia artificial (IA) "seguros y fiables" que beneficien también al desarrollo sostenible para todos.

Al adoptar sin votación un proyecto de resolución liderado por Estados Unidos, la Asamblea también destacó el respeto, la protección y la promoción de los derechos humanos en el diseño, el desarrollo, el despliegue y el uso de la IA. El texto fue "copatrocinado" o respaldado por más de 120 Estados miembros.



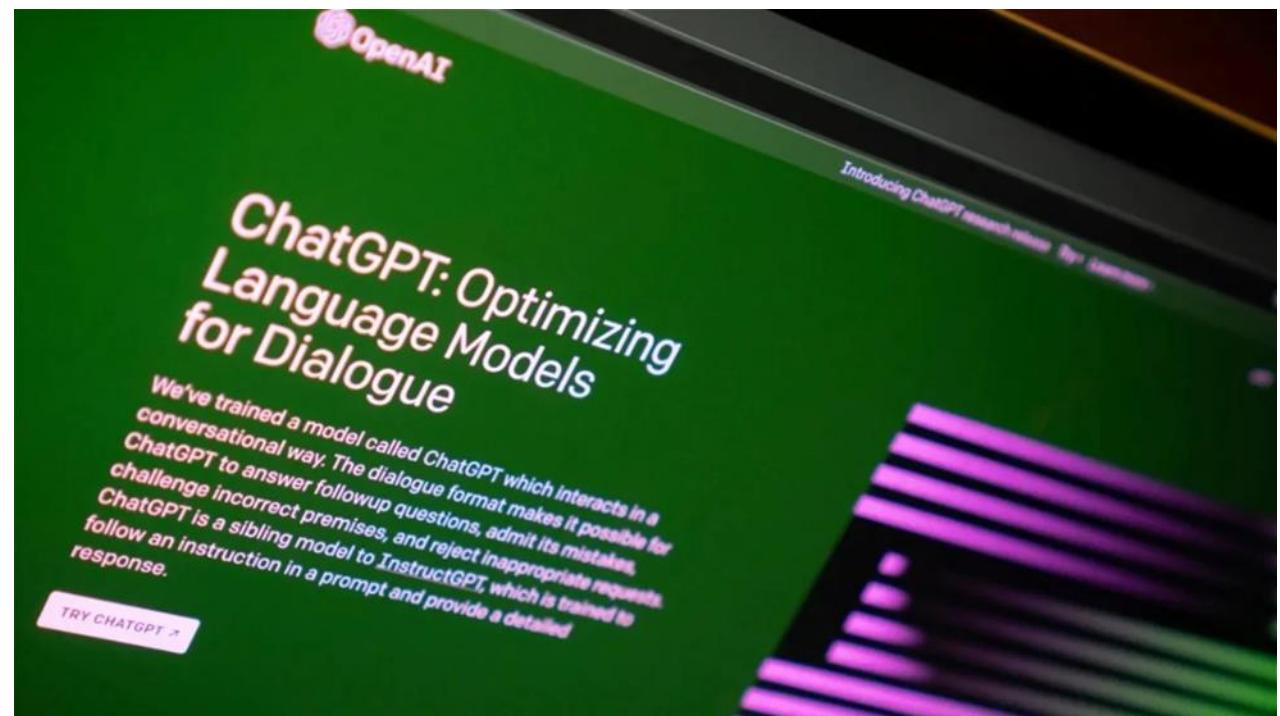
Asamblea general de la ONU.



# 04-24

## Chat GPT sin logueo.

A fin de atraer una mayor cantidad de usuarios. OpenAI dispuso la no obligatoriedad de generar una cuenta de usuario para utilizar ChatGPT. De esta manera, se hizo posible la utilización del sistema en modo "incognito".



El parlamento europeo aprueba la ley IA



# 04-24

## Se presenta Meta Llama 3.

La tercera iteración del modelo LLM Open Source Llama fue lanzada en abril del 2024. Tiene tres versiones de tamaño, 8B, 70B y 400B.

Diversas pruebas indican que su versión más poderosa iguala a ChatGPT 4. La empresa planea incorporarlo como complemento en sus servicios de WhatsApp, Instagram, Facebook y Messenger.



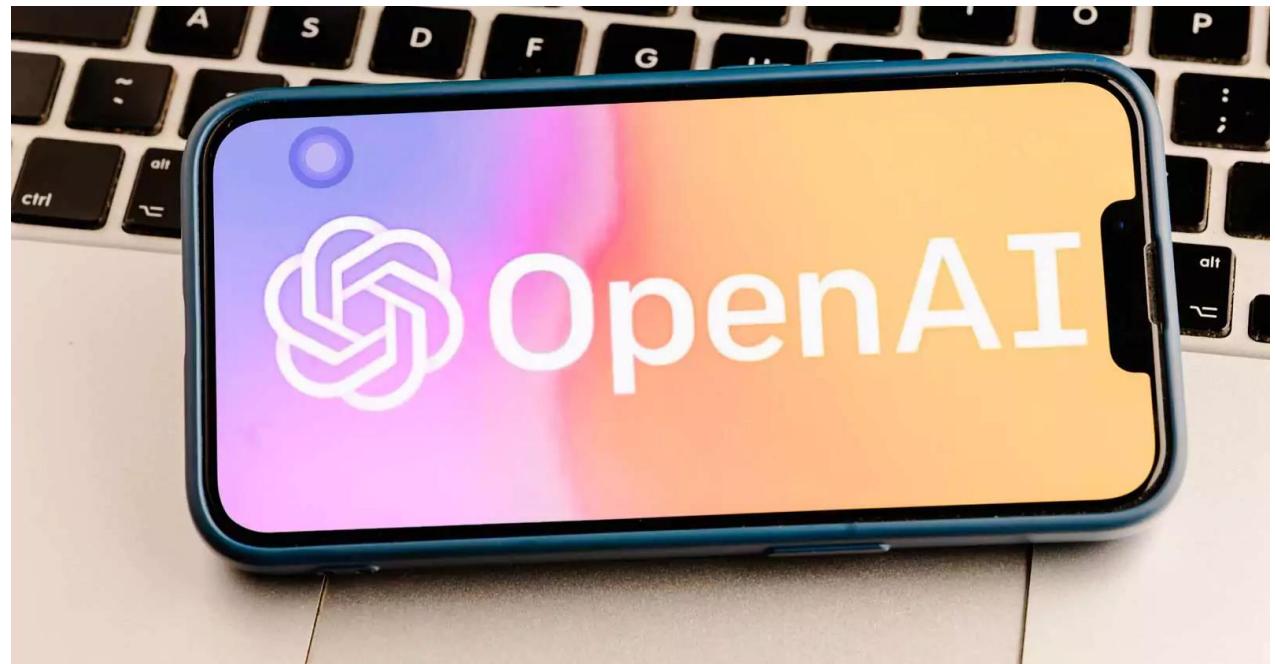
Llama 3 se erige como el LLM Open Source más poderoso.



# 05-24

## Se presenta GPT4o

GPT-4o (omni) es una evolución notable de GPT-4 que destaca por su capacidad multimodal nativa, lo que permite interacciones en tiempo real con baja latencia. Este avance implica mejoras significativas en sus habilidades de texto, audio y visión. Una característica destacada es su disponibilidad gratuita para todos los usuarios de ChatGPT. La reducción de la latencia es una de las principales mejoras de GPT-4o, con tiempos de respuesta casi instantáneos en comparación con los 5 segundos de GPT-4, con un promedio de 320 milisegundos.



GPT4o



# 05-24

## Google Presenta Project Astra

Project Astra es un proyecto innovador de Google Al que tiene como objetivo crear el asistente virtual del futuro. Se basa en la tecnología de Gemini, un modelo de lenguaje a gran escala (LLM) desarrollado por Google, y promete ofrecer una experiencia de usuario mucho más natural e intuitiva que los asistentes virtuales actuales.



Presentación de Project Astra



# 05-24

## Microsoft presenta Copilot + PC

Los nuevos Copilot<sup>+</sup> PC son los equipos con Windows más rápidos e inteligentes lanzados hasta la fecha. Sus procesadores son capaces de alcanzar la increíble cifra de 40+ TOPS (billones de operaciones por segundo), cuentan con una batería con autonomía para toda la jornada y acceso a los modelos de IA más avanzados.



Copilot+PC es la nueva estrategia de Microsoft



# Seguimos la próxima



# 2024

## ¡La IA en todos lados!

MODULO 01. NIVELATORIO

CLASE 04. INTRODUCCIÓN A LOS MODELOS IA (IV)

15

.UBA económicas  
*Diplomaturas*

**IA** aplicada a  
Entornos Digitales  
de Gestión



# MÓDULO 01 - CLASE 04

## Listas - Diccionarios - NUMPY - APIs



# Google colab

---

## Google Colab Clase 4

[acortar.link/clase4python](https://acortar.link/clase4python)



MODULO 01. NIVELATORIO

CLASE 04. Python (IV)

01

.UBA económicas  
*Diplomaturas*

**IA** aplicada a  
Entornos Digitales  
de Gestión



**Una lista en Python es una colección ordenada y mutable de elementos que pueden ser de diferentes tipos. Las listas se definen utilizando corchetes “[]” y los elementos están separados por comas. Las listas permiten modificar, agregar y eliminar elementos después de su creación.**



Una tupla en Python es una colección ordenada pero inmutable de elementos. Una vez creada, no se pueden modificar, agregar ni eliminar elementos de una tupla. Las tuplas se definen utilizando paréntesis "()" y los elementos están separados por comas.





# Ordenar Listas

MODULO 01. NIVELATORIO

CLASE 04. Python (IV)

04

.UBA económicas  
*Diplomaturas*

IA  
aplicada a  
Entornos Digitales  
de Gestión





# Listas Multidimensionales

Una lista de varias dimensiones, también conocida como lista multidimensional o matriz, es una estructura de datos que contiene otras listas como sus elementos. Esto permite organizar datos en forma de tablas, cubos o cualquier otra estructura de múltiples niveles. En Python, una lista multidimensional se implementa como una lista de listas.





# Diccionarios

Un diccionario tiene es una colección no ordenada de parejas, estas parejas contienen llaves y datos. Los datos pueden ser números (enteros, flotantes), cadenas (strings, booleanos) o arreglos (listas,tuplas) e incluso otro diccionario.

Es posible modificar los datos de un diccionario, o su longitud. Es decir, los diccionarios son objetos mutables, sin embargo, las llaves asociadas a los datos, no lo son.





# Diccionarios

Para crear un diccionario puedes utilizar llaves e ingresar sus elementos separados con una coma.

Cada elemento se escribe con la siguiente notación:

Llave : valor





# Arreglos y NUMPY

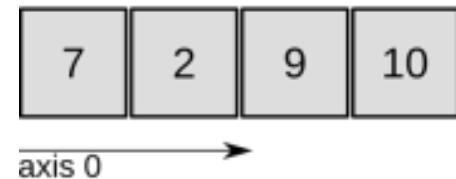
---

NumPy es una de las librerías más populares de Python, ya que puede procesar listas de múltiples dimensiones que incluyen lazos anidados, o comprensión de listas con diversas sentencias for.



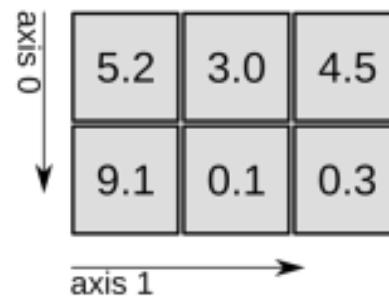
# Python Arreglos y NUMPY

1D array



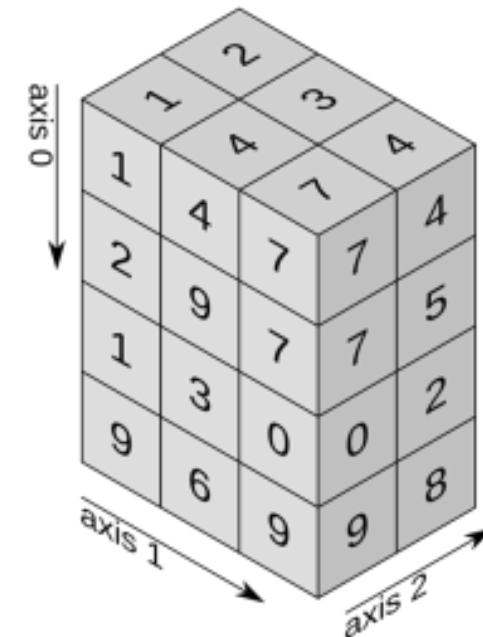
shape: (4,)

2D array



shape: (2, 3)

3D array



shape: (4, 3, 2)





# Arreglos y NUMPY

---

NumPy está diseñado para trabajar principalmente con datos homogéneos, lo que significa que todos los elementos dentro de un array deben ser del mismo tipo de datos. Esto contrasta con las listas de Python regulares, que pueden contener elementos de diferentes tipos.





# Arreglos y NUMPY

---

[numpy.org/doc/stable](https://numpy.org/doc/stable)



Una API (Interfaz de Programación de Aplicaciones) es un conjunto de reglas, protocolos y herramientas que permiten la comunicación entre diferentes aplicaciones de software. Proporciona un conjunto de métodos y funciones que otras aplicaciones pueden utilizar para acceder y manipular los datos o funcionalidades de un servicio o sistema de software específico de manera segura y controlada.



## Endpoints

- Los endpoints son URLs específicas dentro de una API que corresponden a recursos o funciones específicas. Por ejemplo, un endpoint puede ser /usuarios para acceder a información sobre usuarios en un sistema.



## Métodos HTTP

- Los métodos HTTP, como **GET**, **POST**, **PUT** y **DELETE**, se utilizan para indicar la acción que se realizará en un recurso específico a través de un endpoint. Por ejemplo, **GET** se utiliza para recuperar información, **POST** para crear nuevos recursos, **PUT** para actualizar recursos existentes, y **DELETE** para eliminar recursos.



## Parámetros de Consulta (Query Parameters)

- Los parámetros de consulta se utilizan para filtrar, ordenar o paginar los resultados devueltos por un endpoint. Por ejemplo, en la URL `/usuarios?estado=activo&limite=10`, "estado" y "limite" son parámetros de consulta.



## Formatos de Datos (Payload):

- Los formatos de datos se utilizan para representar la información que se envía o se recibe a través de una API. Los formatos comunes incluyen JSON (JavaScript Object Notation) y XML (Extensible Markup Language).



## Autenticación y Autorización

- La autenticación y autorización se utilizan para garantizar que solo los usuarios autorizados puedan acceder a los recursos y funcionalidades de una API. Esto puede implicar el uso de tokens de acceso, claves de API, o sistemas de autenticación basados en OAuth.



## Respuestas de la API

- Las respuestas de la API son los datos devueltos por un endpoint después de que se ha realizado una solicitud. Esto puede incluir información solicitada, mensajes de estado, códigos de error, y más.

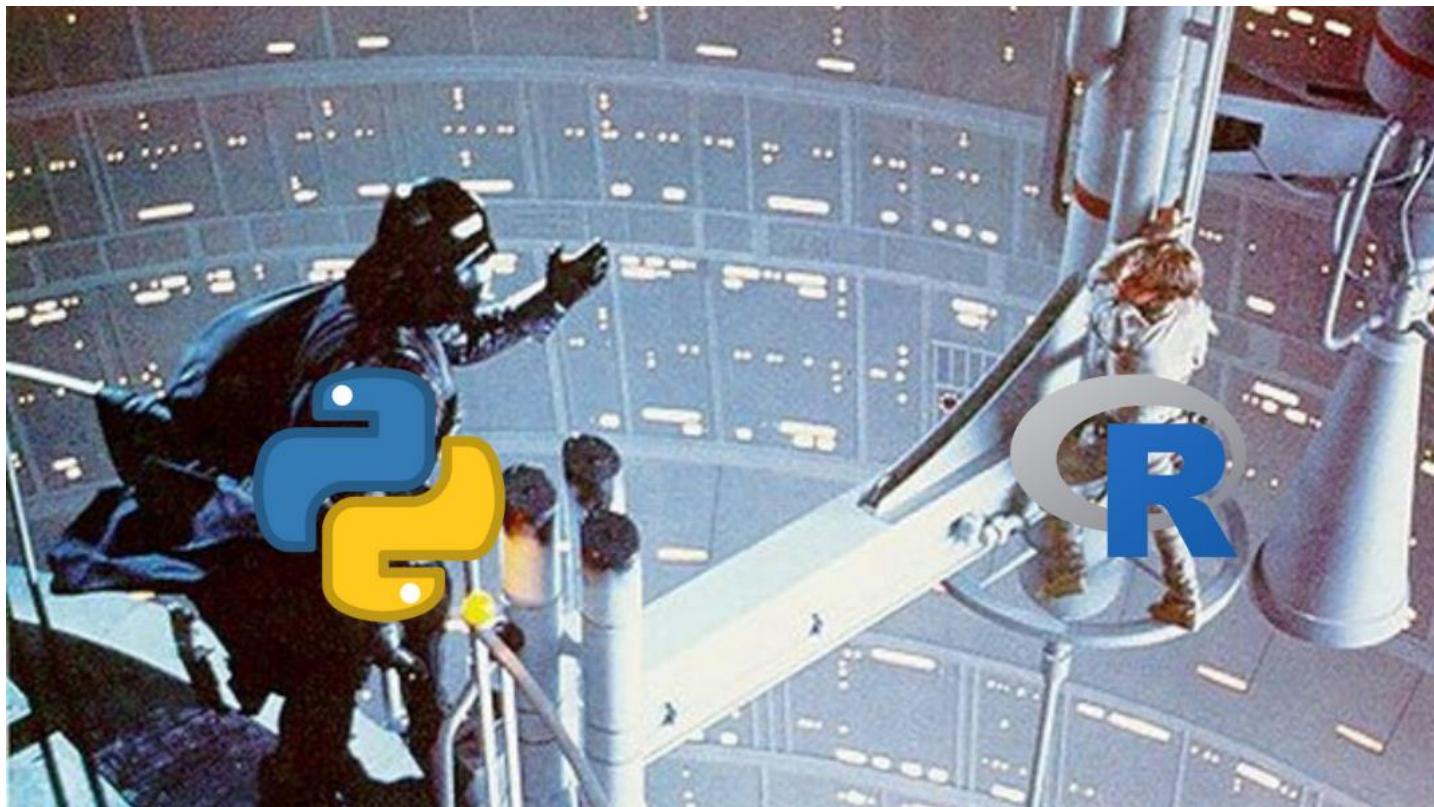




[openweathermap.org/](https://openweathermap.org/)

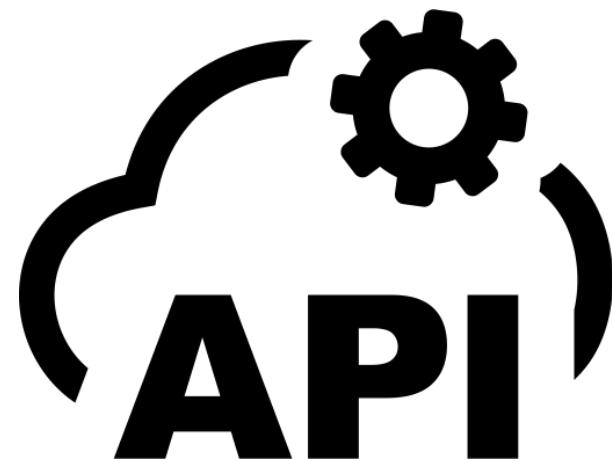


# Python APIs



swapi.dev/





publicapis.dev/



**PARTE II**

**CLASE**

**5**

# MÓDULO 02 - CLASE 01

# INTRODUCCIÓN

# AL MACHINE LEARNING

(y un poco más de Python)

.UBA económicas  
*Diplomaturas*

IA aplicada a  
Entornos Digitales  
de Gestión



# Definición de los sistemas de ML

MODULO 02. MACHINE LEARNING.

CLASE 01. INTRODUCCIÓN AL MACHINE LEARNING.

2

.UBA económicas  
*Diplomaturas*

IA  
aplicada a  
Entornos Digitales  
de Gestión



# ¿Qué es?

## Concepto de Aprendizaje Automático.

"El aprendizaje automático es un subdominio de la IA que proporciona a los sistemas la capacidad de aprender y mejorar automáticamente a partir de la experiencia sin ser programados explícitamente a tales fines. Se basa en la hipótesis subyacente de crear el modelo y tratar de mejorarlo a través de la adición de más datos en éste a lo largo del tiempo".

El término fue popularizado en 1959 en el paper de Arthur Samuel "Some Studies in Machine Learning. Using the Game of Checkers".

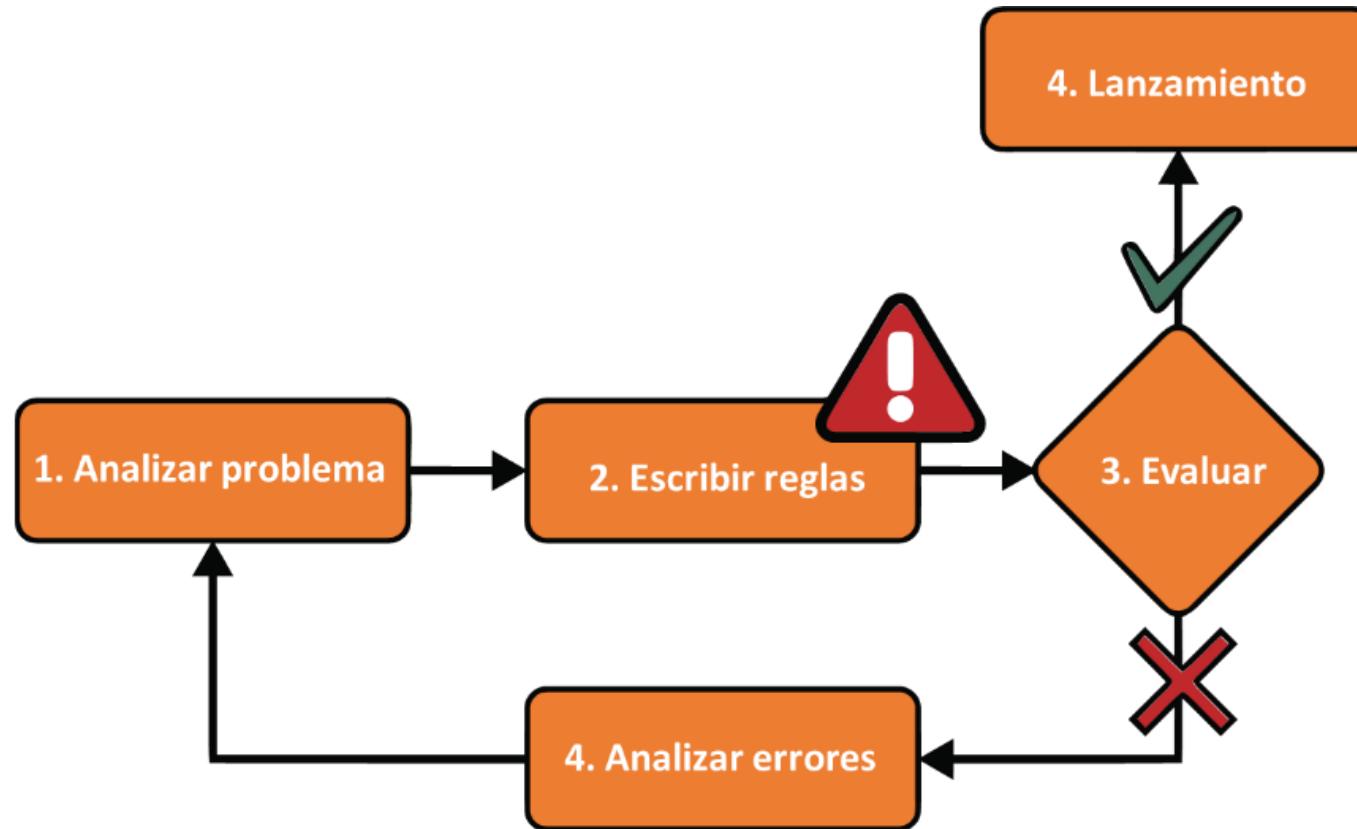
A. L. Samuel

### **Some Studies in Machine Learning Using the Game of Checkers**

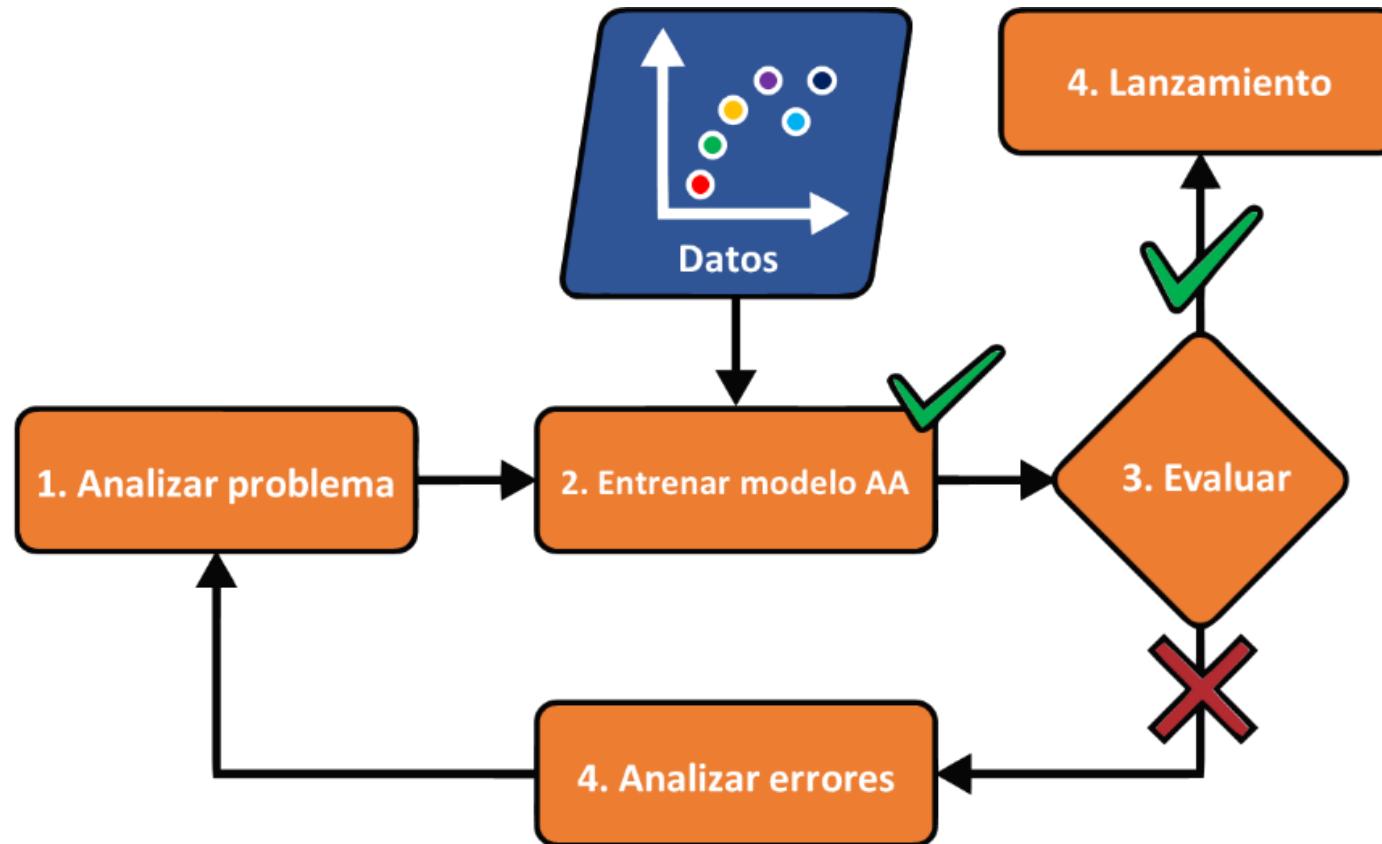
**Abstract:** Two machine-learning procedures have been investigated in some detail using the game of checkers. Enough work has been done to verify the fact that a computer can be programmed so that it will learn to play a better game of checkers than can be played by the person who wrote the program. Furthermore, it can learn to do this in a remarkably short period of time (8 or 10 hours of machine-playing time) when given only the rules of the game, a sense of direction, and a redundant and incomplete list of parameters which are thought to have something to do with the game, but whose correct signs and relative weights are unknown and unspecified. The principles of machine learning verified by these experiments are, of course, applicable to many other situations.



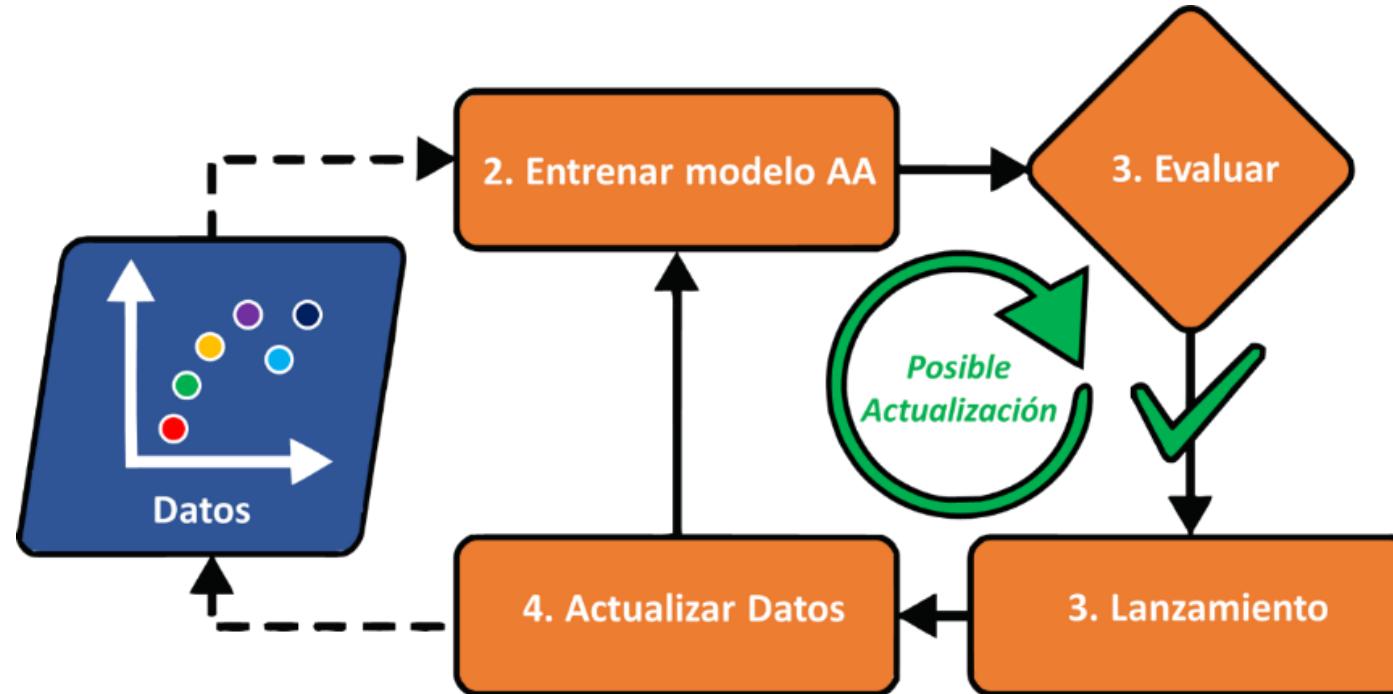
**Se plantea la construcción de un filtro capaz de identificar y bloquear correos de SPAM.  
(paradigma antiguo)**



**Se plantea la construcción de un filtro capaz de identificar y bloquear correos de SPAM.  
(enfoque del aprendizaje automático)**



**Se plantea la construcción de un filtro capaz de identificar y bloquear correos de SPAM.  
(enfoque del aprendizaje automático con datos cambiantes)**



# ¿Cuándo utilizar ML?

- En soluciones que funcionan mediante la aplicación de un conjunto extenso de reglas o heurísticas. (**Filtro de SPAM**).
- En problemas complejos en los que un analista no es capaz de determinar una solución a partir de la información existente. (**Clasificación de IMÁGENES**).
- En entornos que fluctúan o varían con frecuencia. (**Tráfico de REDES**).
- Apoyo de la fase de análisis en enfoques tradicionales en los que se dispone de conjuntos de datos muy grandes y difíciles de interpretar.



# Usos de Machine Learning

Detección de Rostros



# Usos de Machine Learning

Detección de SPAM  
en sistemas de correo  
electrónico.



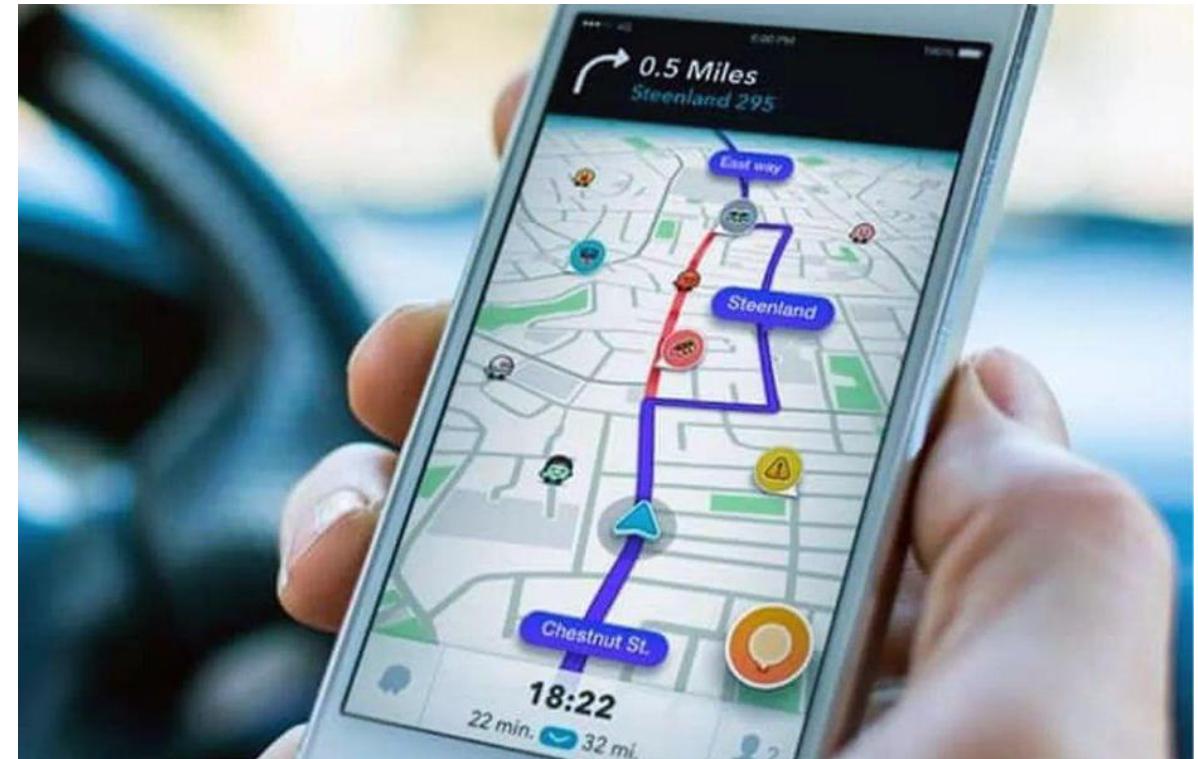
# Usos de Machine Learning

Sistemas de recomendaciones.



# Usos de Machine Learning

Predicción de tráfico.



# Clasificación de los sistemas de ML

MODULO 02. MACHINE LEARNING.

CLASE 01. INTRODUCCIÓN AL MACHINE LEARNING.

12

.UBA económicas  
*Diplomaturas*

IA  
aplicada a  
Entornos Digitales  
de Gestión



# Clasificación de sistemas de ML

1. En función de la manera en que se entrenan los modelos.
2. En función de la manera en la que los modelos aprenden en el tiempo.
3. En función de la forma en la que los modelos realizan las predicciones.



# En función de la manera en que se entrenan los modelos

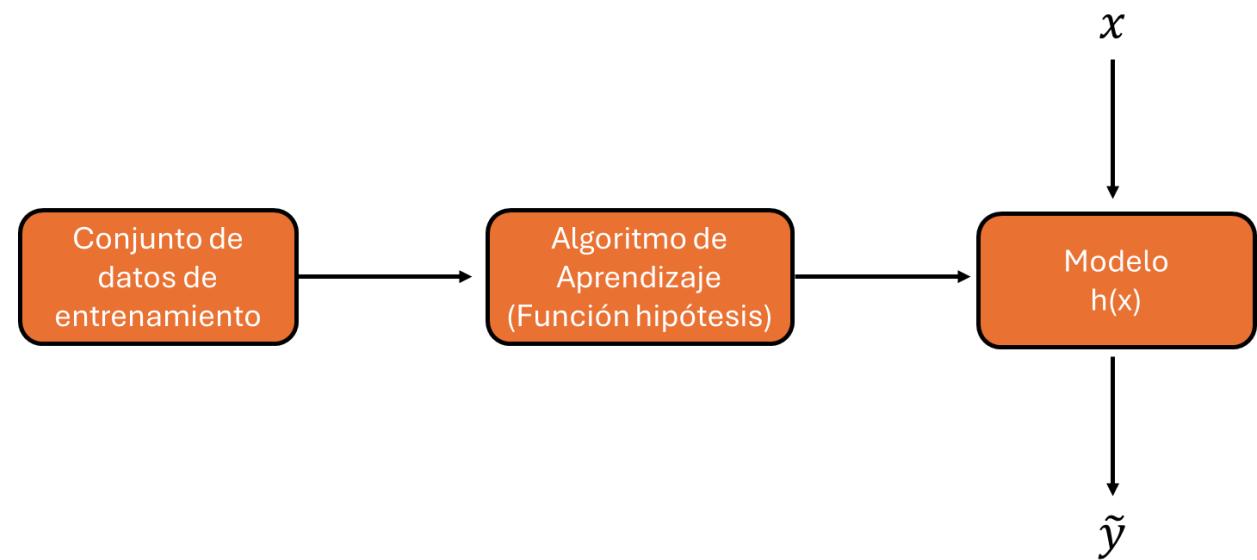
1. Aprendizaje supervisado.
2. Aprendizaje no supervisado.
3. Aprendizaje semi-supervisado.
4. Aprendizaje reforzado.



# Aprendizaje supervisado

"El aprendizaje supervisado es la tarea de aprendizaje automático que consiste en aprender una función que mapea una entrada a una salida basada en pares de entrada-salida de ejemplo".

La función resultante es utilizada posteriormente para predecir valores a partir de nuevos ejemplos.

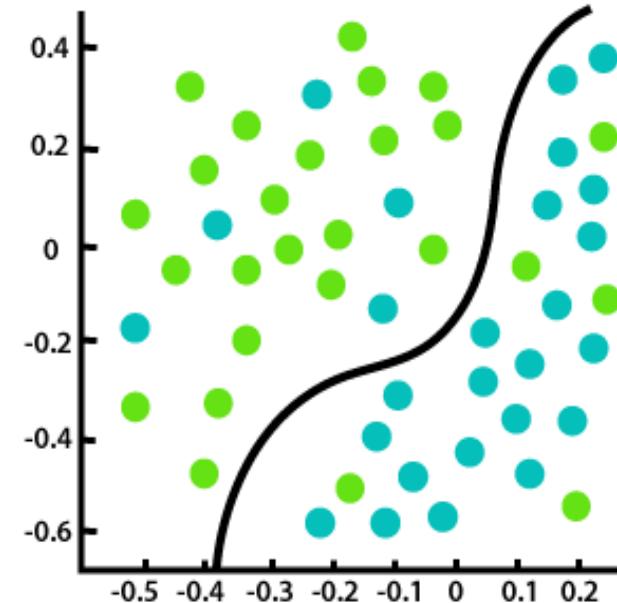


# Aprendizaje supervisado

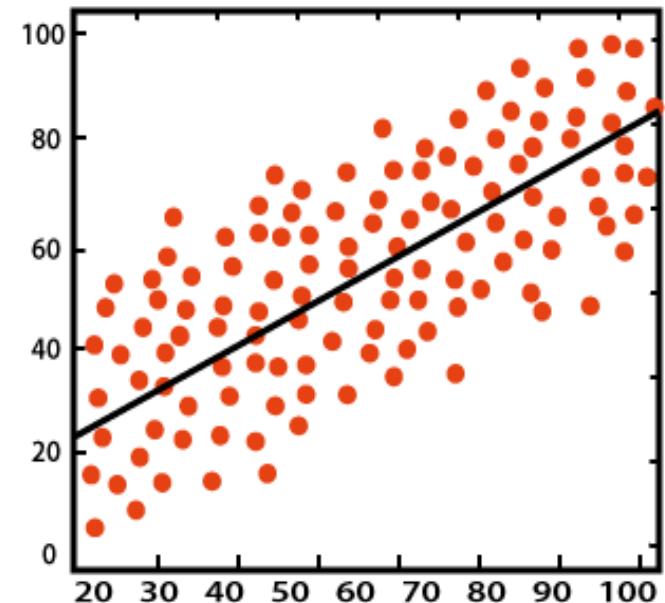
Existen dos tipos principales de aprendizaje supervisado:

**Regresión.** (Intenta predecir valores)

**Clasificación.** (Intenta predecir categorías o clases)



CLASIFICACIÓN

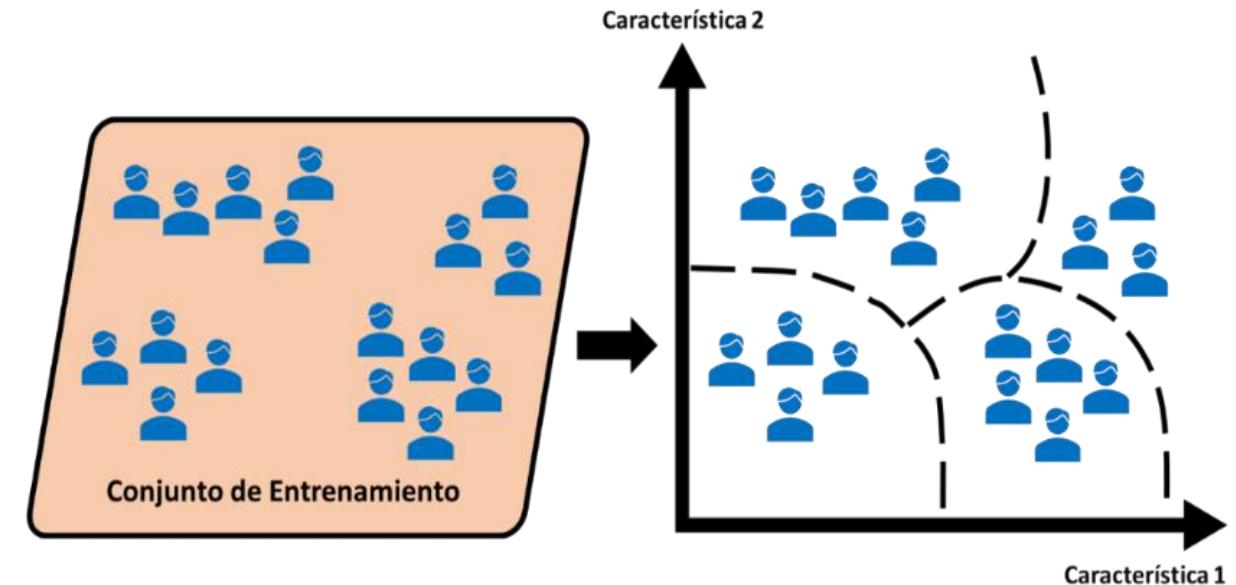


REGRESIÓN



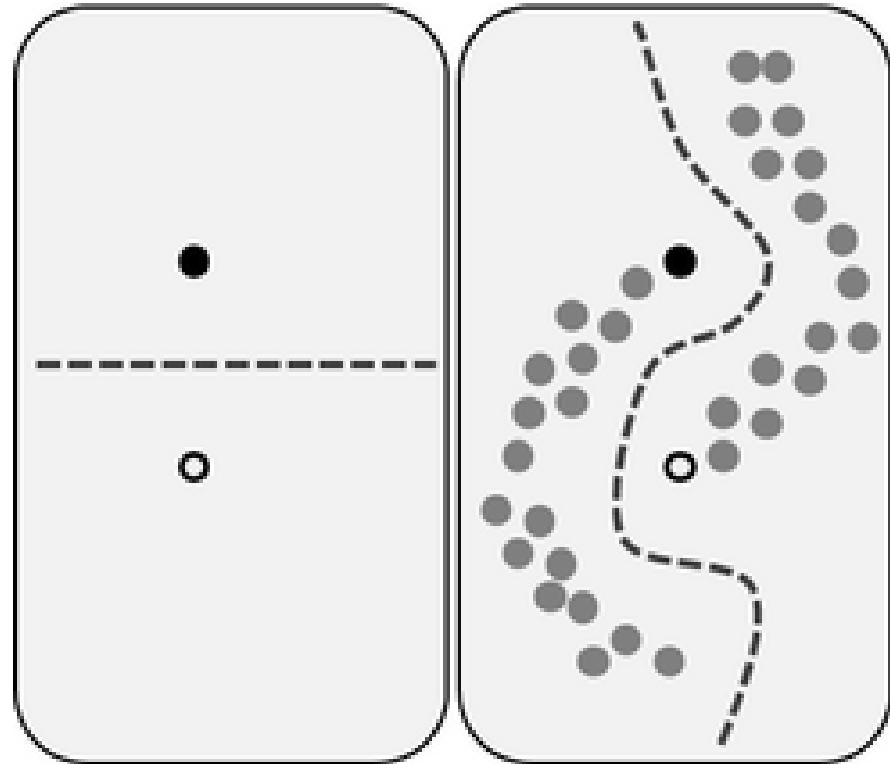
# Aprendizaje no supervisado

El aprendizaje no supervisado es la tarea de aprendizaje automático que consiste en **inferir una función que describe la estructura de un conjunto de datos sin etiquetar, clasificar o categorizar.**



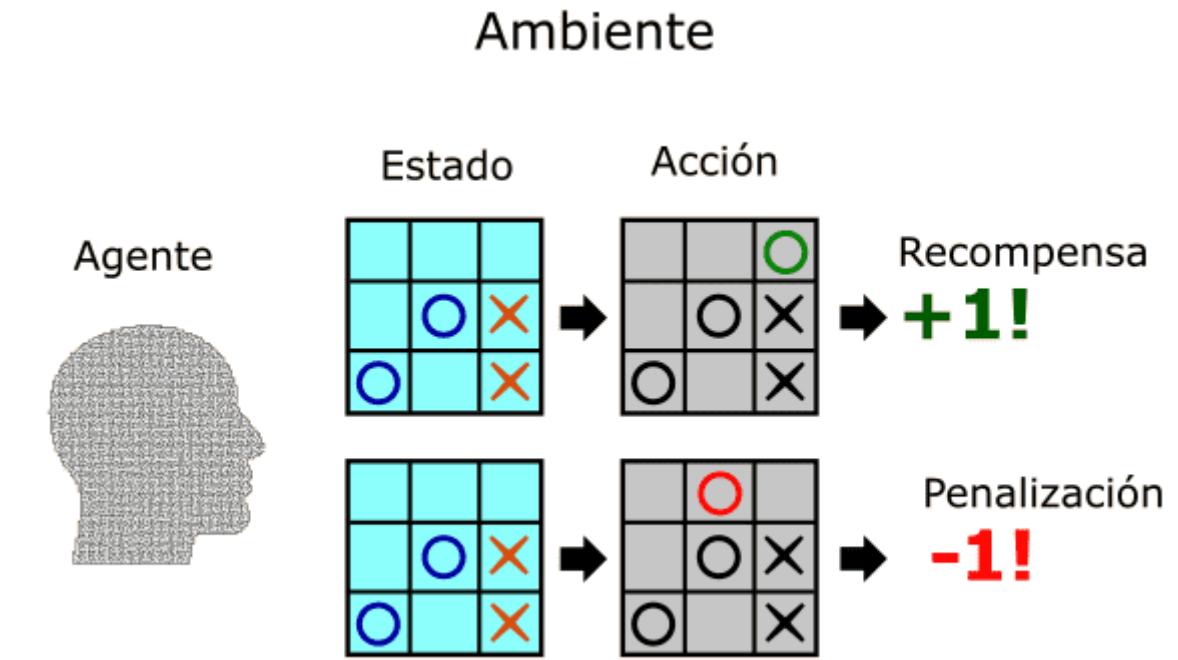
# Aprendizaje Semi - supervisado

El aprendizaje semi - supervisado es un tipo de aprendizaje automático que combina el aprendizaje supervisado y no supervisado para entrenar modelos de IA. Se utiliza cuando se tiene un conjunto de datos parcialmente etiquetado.



# Aprendizaje por refuerzo (RL)

El RL imita el proceso de aprendizaje por ensayo y error que los humanos utilizan para lograr sus objetivos. En el aprendizaje por refuerzo, una máquina aprende a partir de su propia experiencia, interaccionando con el entorno hasta dar con el comportamiento ideal. La máquina emprenderá acciones que repetirá y “reforzará” según las recompensas que obtenga, que pueden ser positivas o negativas.



# En función de la manera en la que los modelos aprenden en el tiempo

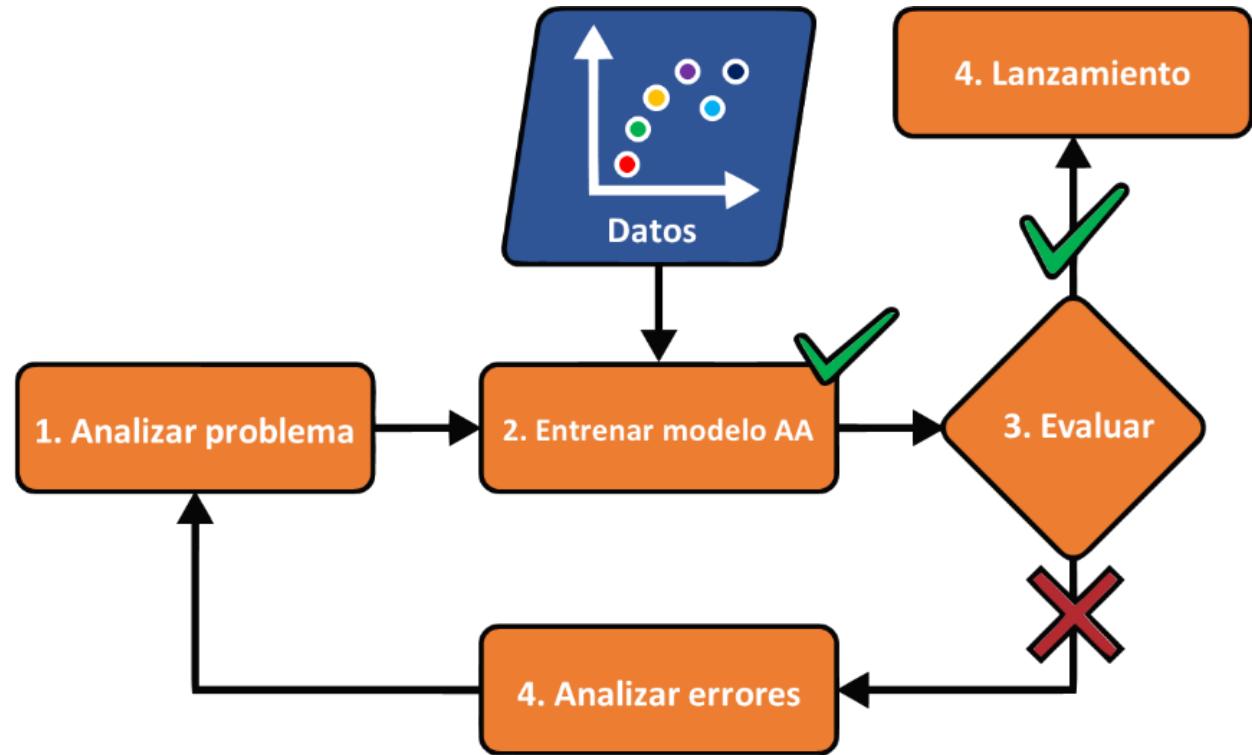
1. Aprendizaje online.
2. Aprendizaje batch.



# Aprendizaje batch

Los sistemas basados en aprendizaje batch no aprenden de manera incremental, se entrenan utilizando todos los datos disponibles.

Ejemplo, se utiliza una base de datos con distintas características de las viviendas y en función a dichas características, se intenta generar un modelo para predecir su precio.



# Aprendizaje batch

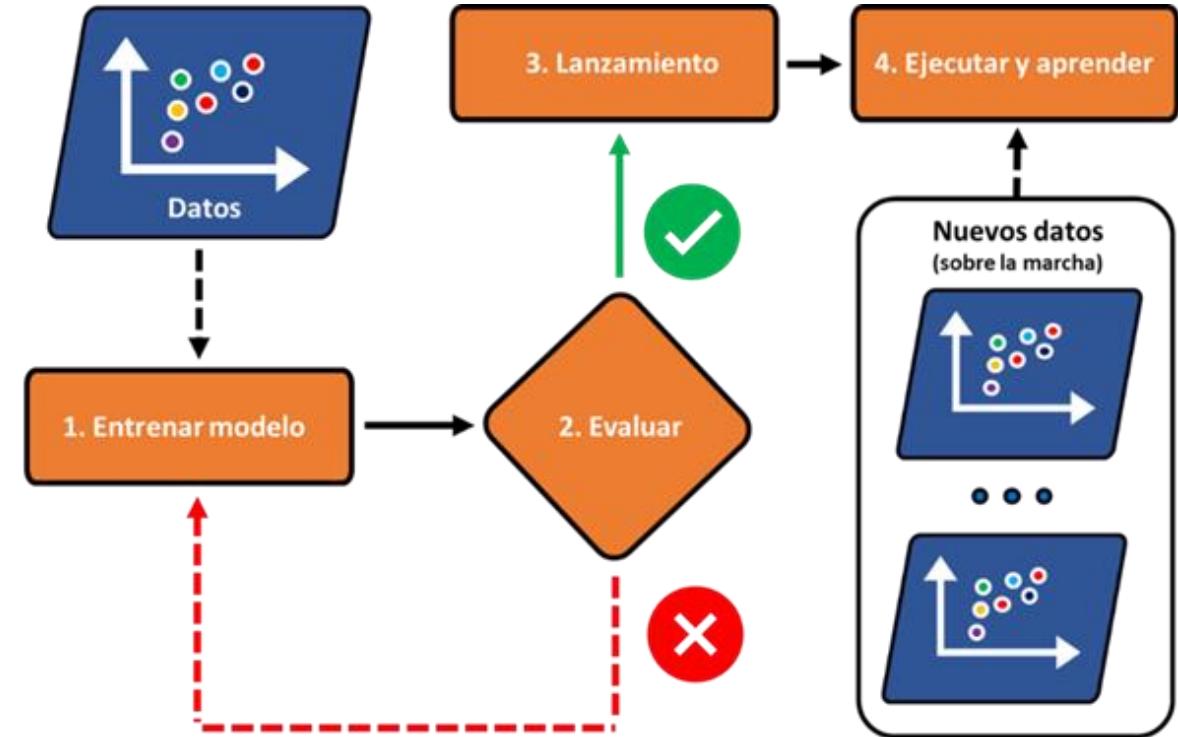
- Si se desea que el sistema se adapte a un nuevo tipo de dato, se debe entrenar de nuevo con todos los datos disponibles.
- Solución sencilla.
- Funciona bien para sistemas que no requieren un conjunto de datos muy grande ni adaptarse a nuevos datos de manera muy rápida.
- Muy restringido para dispositivos con una capacidad limitada de recursos, como un smartphone.



# Aprendizaje online

Los sistemas basados en aprendizaje online se entrena incrementalmente, mediante el consumo incremental de datos, ya sea individuales o en pequeños grupos (mini-batches).

Ejemplo (Detección de fraudes en transacciones bancarias).



# Aprendizaje online

- Los sistemas basados en aprendizaje online se entrenan incrementalmente, mediante el consumo incremental de datos, ya sea individuales o en pequeños grupos (mini-batches).
- Solución ideal para sistemas que reciben datos continuamente y requieren adaptarse a ellos de manera rápida.
- Es capaz de lidiar con grandes conjuntos de datos que puede que no entren en una sola máquina.
- Aparecen algunas variables importantes que hay que determinar, como la ratio de aprendizaje.
- Pueden ser muy inestables si por alguna razón se consumen datos de baja calidad.



# En función de la forma en la que los modelos realizan las predicciones

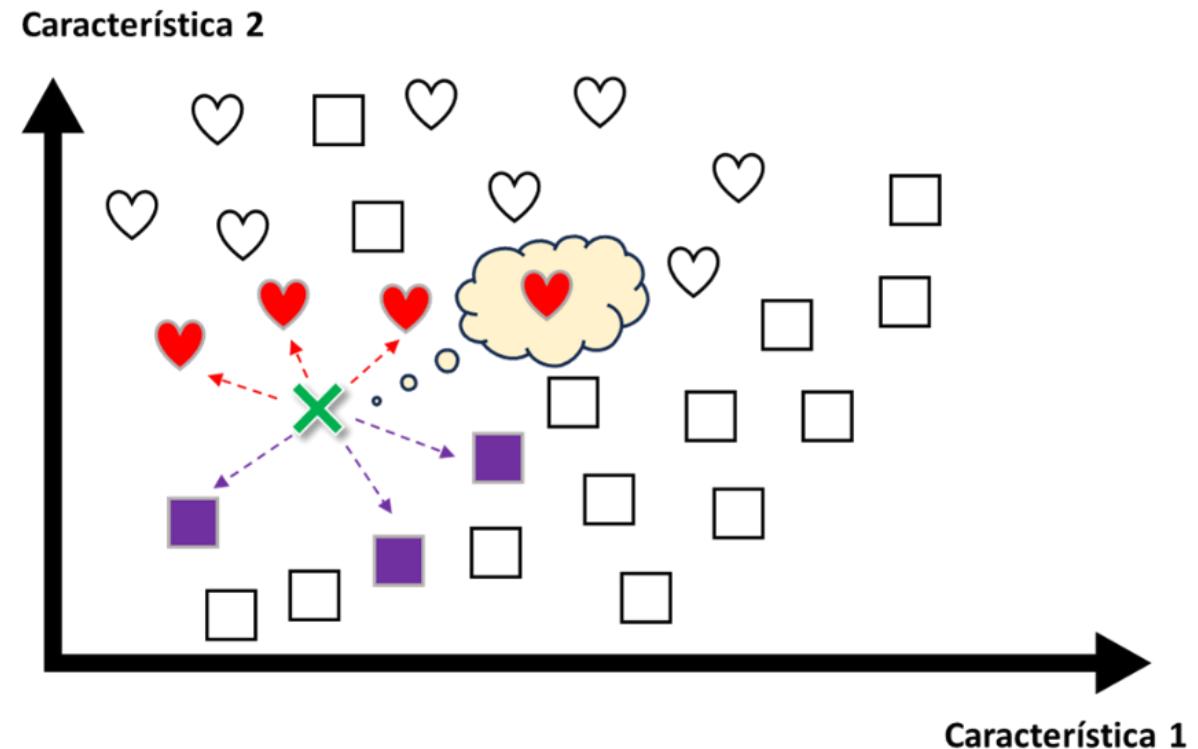
1. Aprendizaje basado en instancias.
2. Aprendizaje basado en modelos.



# Aprendizaje basado en instancias

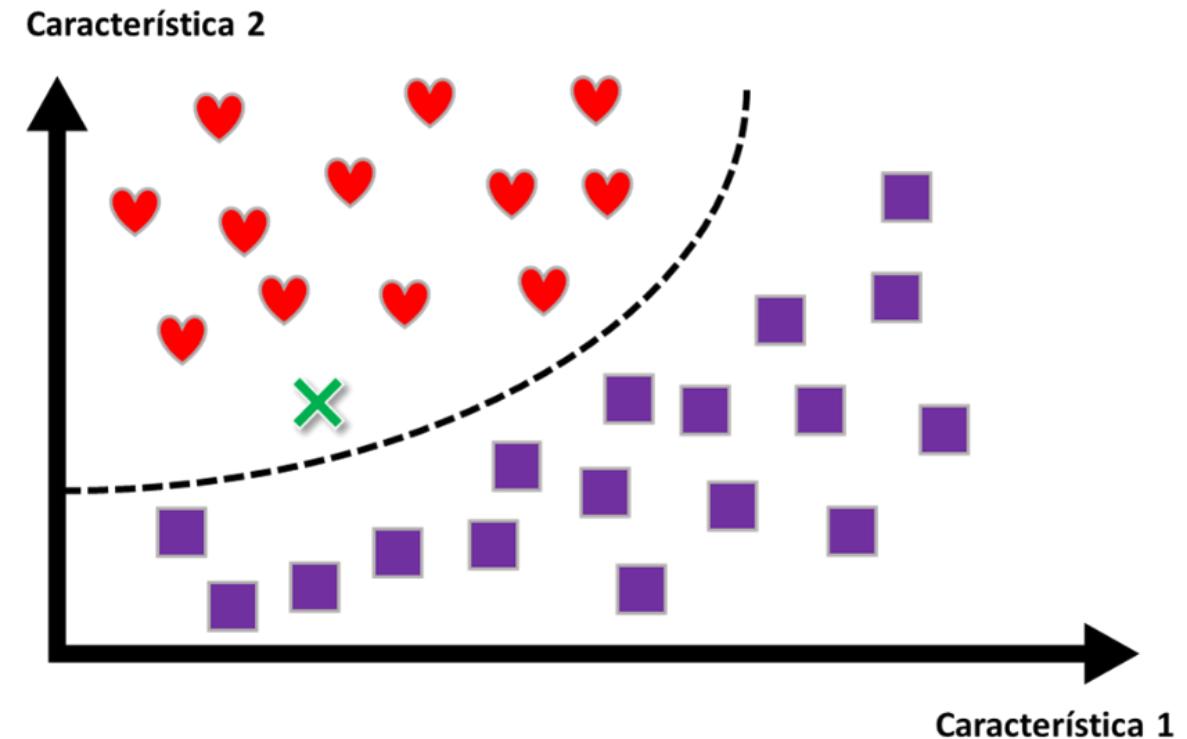
El sistema aprende los ejemplos del conjunto de datos de entrenamiento y luego intenta generalizar para nuevos ejemplares, requiriendo una medida de similitud.

Este tipo de algoritmos "memoriza" las instancias o muestras del conjunto de entrenamiento, utilizando esta información a la hora de realizar una predicción. El mejor ejemplo de esto es el algoritmo *k-Nearest Neighbors* que clasifica una nueva observación a partir de la clasificación que tengan las observaciones de entrenamiento más próximas.

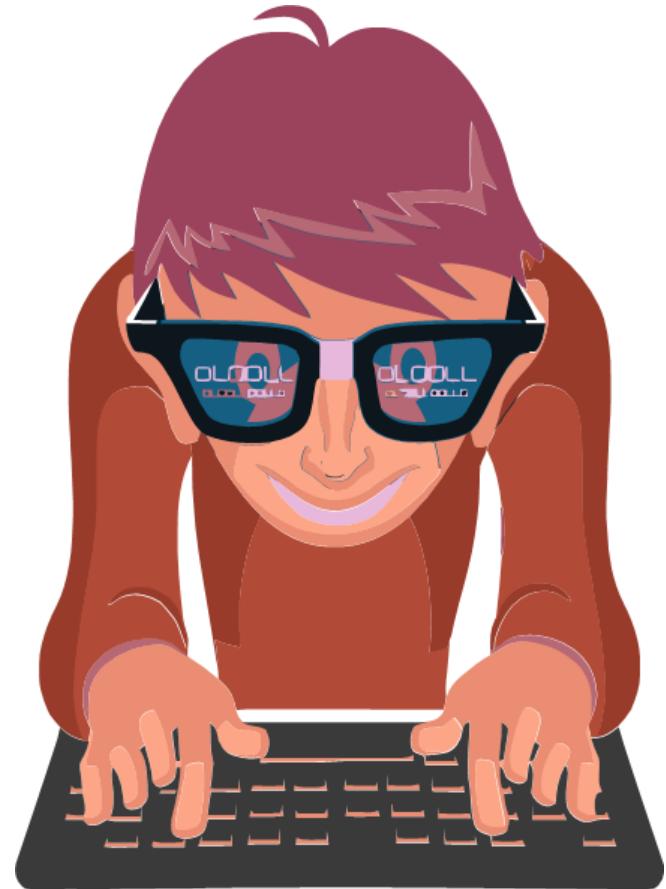


# Aprendizaje basado en modelos

En el aprendizaje basado en modelos, se crea un modelo a partir de datos de entrenamiento para realizar predicciones futuras. Una vez creado, los datos de entrenamiento se descartan. Por ejemplo, en una regresión lineal, se genera una recta de la forma  $ax + b$  a partir de los datos de entrenamiento, y esta recta se utiliza para predecir las etiquetas de nuevas muestras.



# Empecemos ¡con Machine Learning!



MODULO 02. MACHINE LEARNING.

CLASE 01. INTRODUCCIÓN AL MACHINE LEARNING.

28

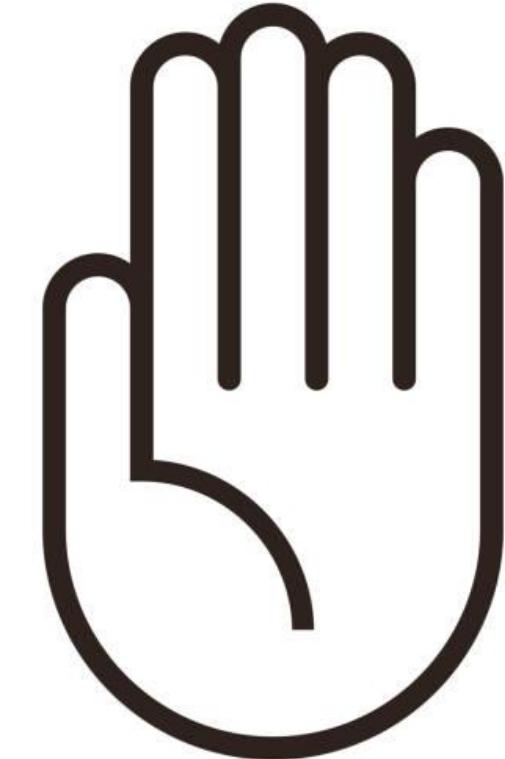
.UBA económicas  
*Diplomaturas*

**IA** aplicada a  
Entornos Digitales  
de Gestión



# Antes de seguir

## ¡Necesitamos más Python!



# NUMPY

¡La Biblioteca Esencial  
de Cálculo Numérico en Python!



# PANDAS

## ¡La Biblioteca Esencial de Análisis de datos en Python!

MODULO 02. MACHINE LEARNING.

CLASE 01. INTRODUCCIÓN AL MACHINE LEARNING.

31

.UBA económicas  
*Diplomaturas*

IA  
aplicada a  
Entornos Digitales  
de Gestión



# MATPLOTLIB

¡La Biblioteca Esencial  
de Visualización de datos en Python!

MODULO 02. MACHINE LEARNING.

CLASE 01. INTRODUCCIÓN AL MACHINE LEARNING.

32

.UBA económicas  
*Diplomaturas*

IA aplicada a  
Entornos Digitales  
de Gestión



**PARTE II**

**CLASE**

**6**

# MÓDULO 02 - CLASE 02

# APRENDIZAJE SUPERVISADO (y aún más de Python)

.UBA económicas  
*Diplomaturas*

IA aplicada a  
Entornos Digitales  
de Gestión



# NOTACIÓN VECTORIAL

MODULO 02. MACHINE LEARNING.

CLASE 02. APRENDIZAJE SUPERVISADO (I).

2

.UBA económicas  
*Diplomaturas*

**IA** aplicada a  
Entornos Digitales  
de Gestión



# Notación Vectorial

Un vector es un elemento matemático que se representa como un n-tupla ordenada de números reales (o complejos), usualmente escrito como  $(v_1, v_2, \dots, v_n)$  donde cada  $v_i$  es una componente del vector.

$$\vec{v} = (v_1, v_2, \dots, v_{n-1}, v_n)$$



# MACHINE LEARNING

## (modelo general)

MODULO 02. MACHINE LEARNING.

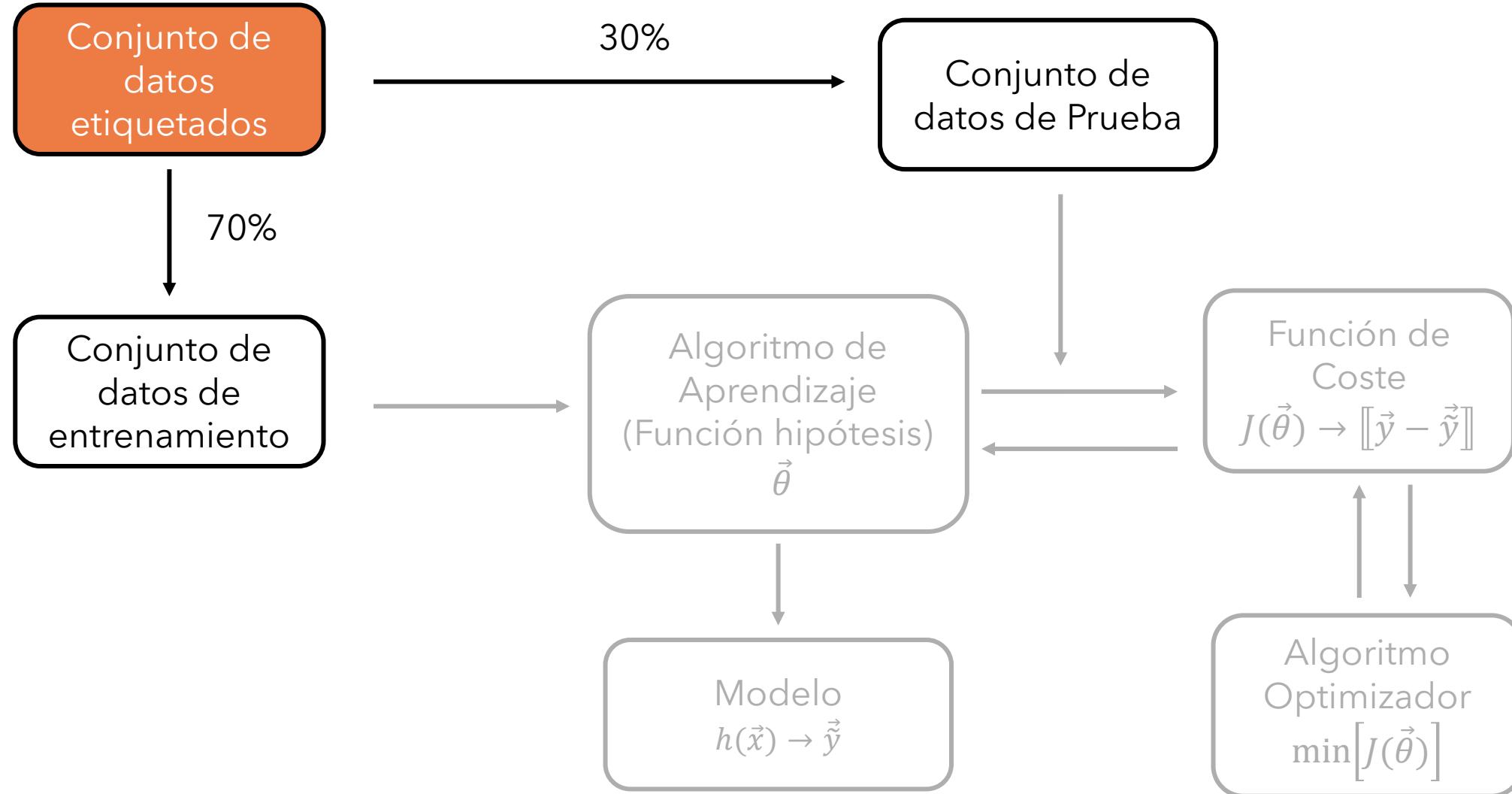
CLASE 02. APRENDIZAJE SUPERVISADO (I).

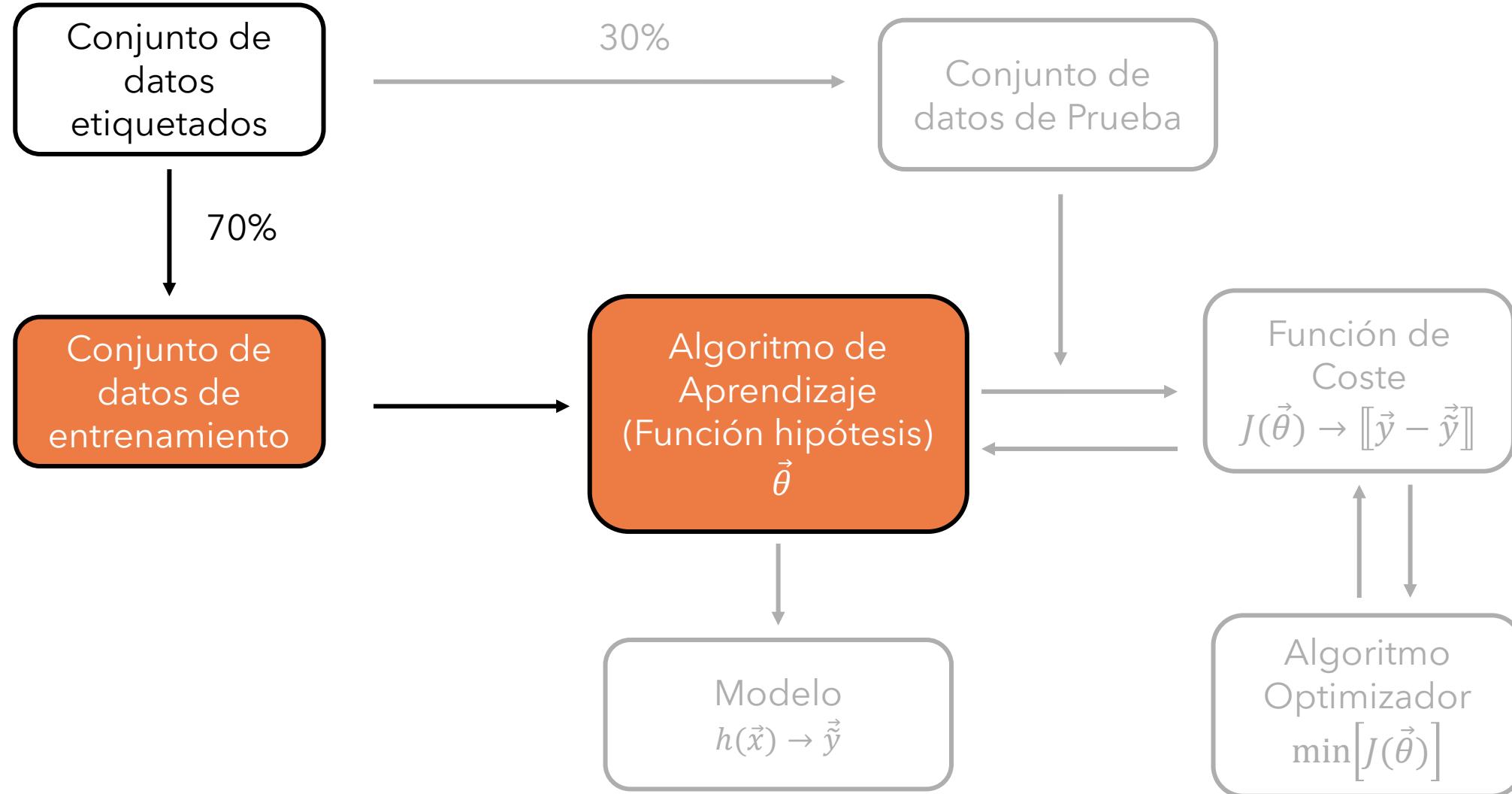
4

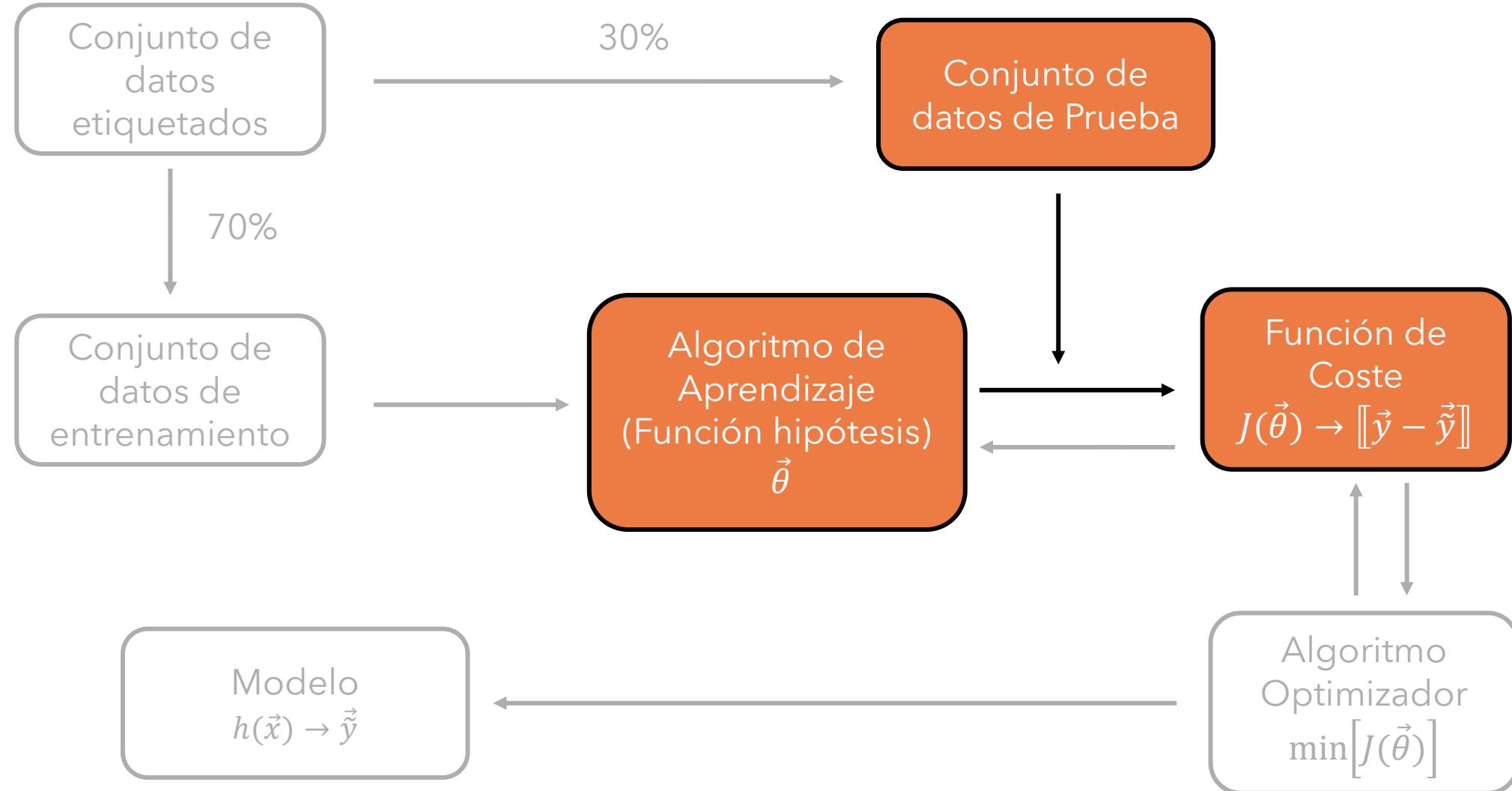
.UBA económicas  
*Diplomaturas*

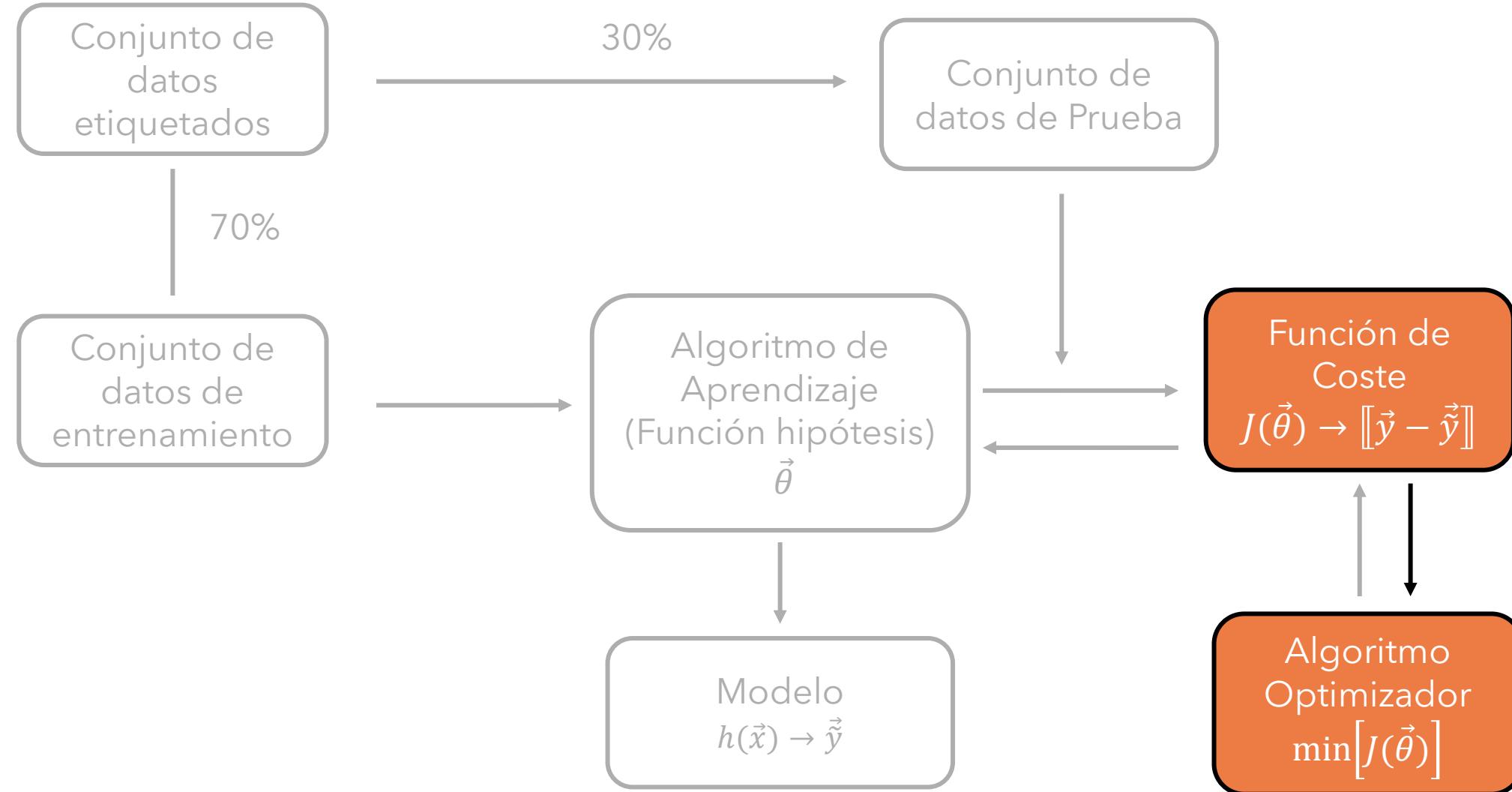
**IA** aplicada a  
Entornos Digitales  
de Gestión

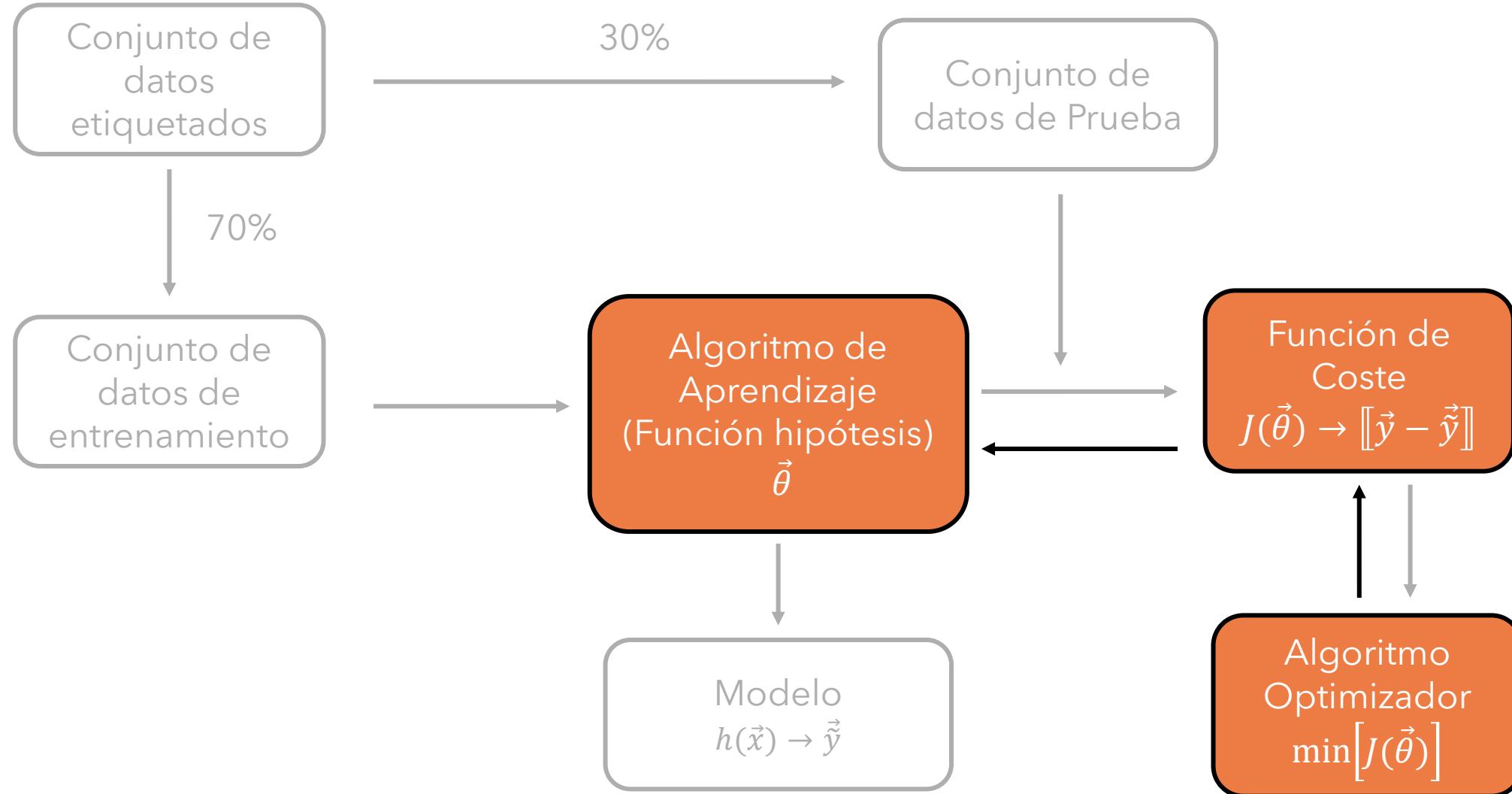


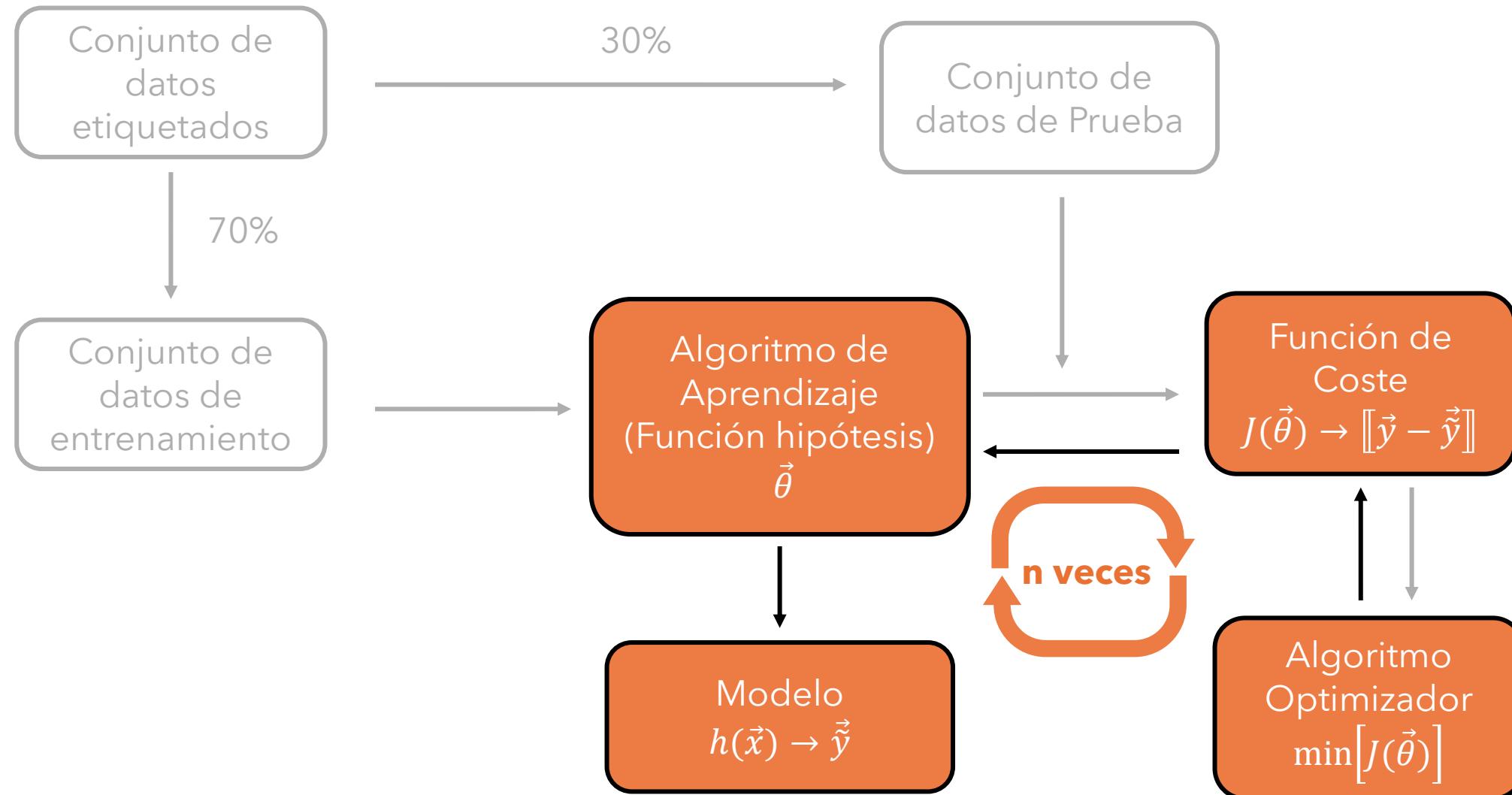












# TIPOS DE APRENDIZAJE SUPERVISADO

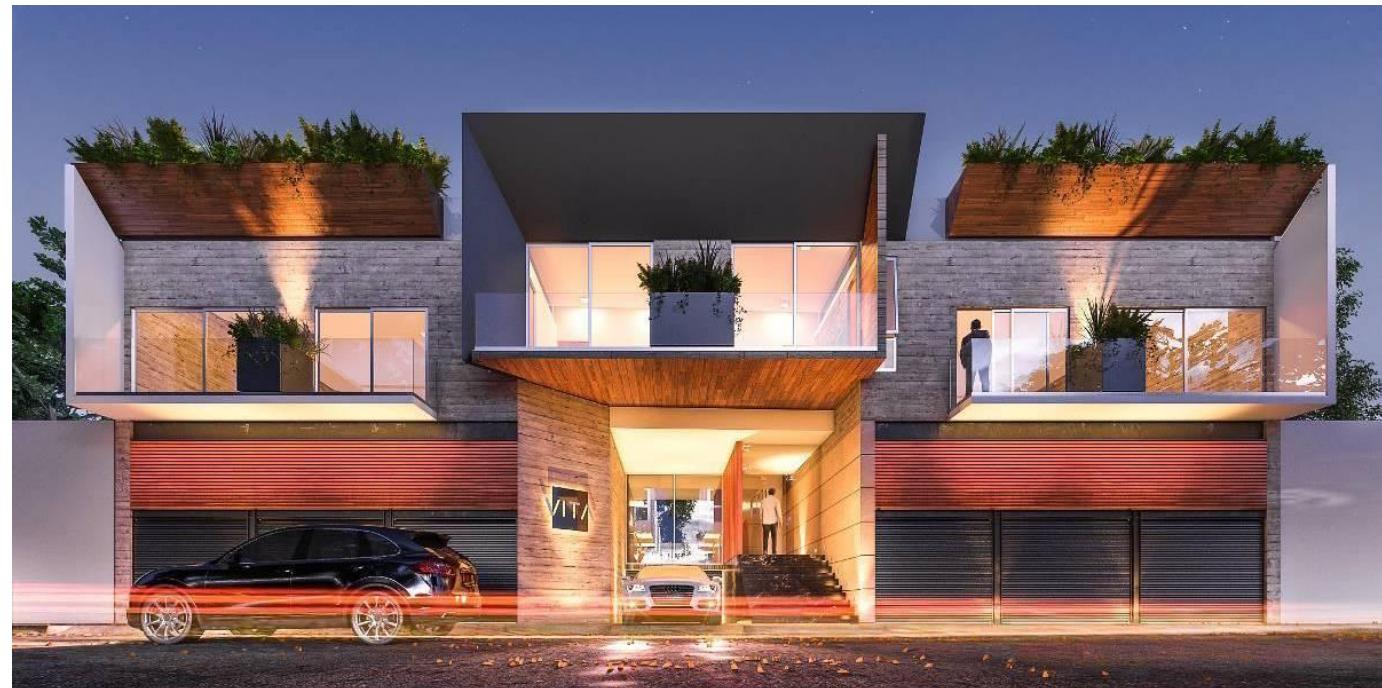


# Tipos de Aprendizaje Supervisado

## Regresión:

En los problemas de regresión, los valores a predecir son continuos. Es necesario identificar las variables independientes para establecer una fórmula matemática (generalmente una función lineal) que relacione el valor de la variable dependiente con las variables independientes.

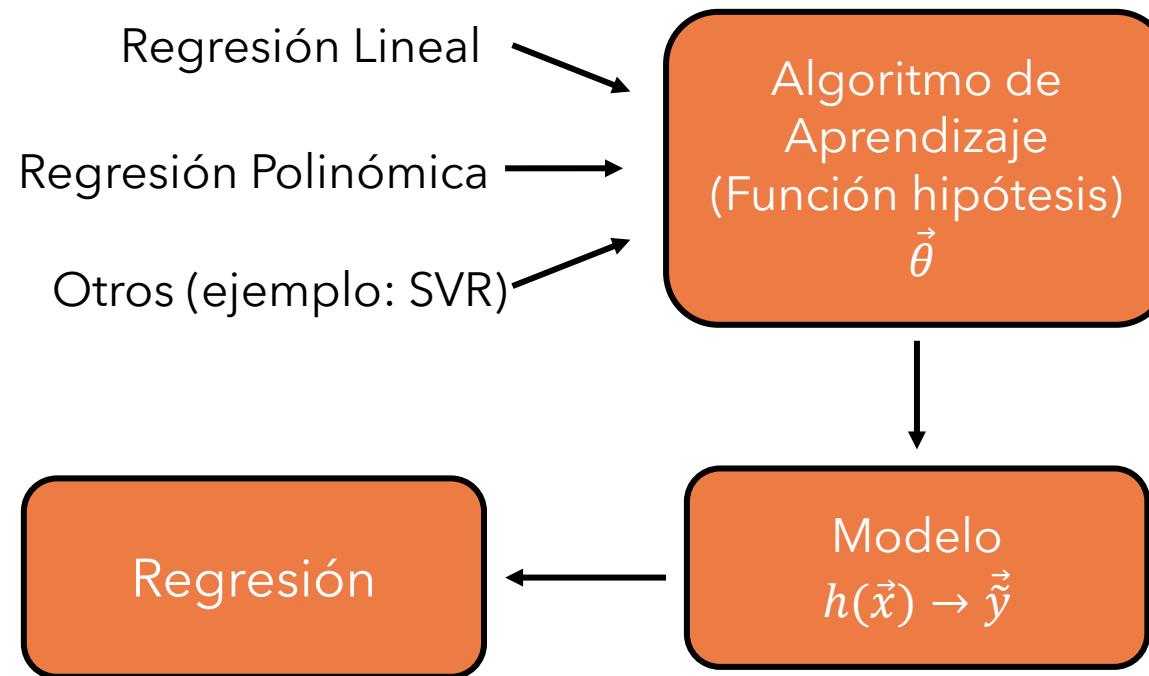
Esta variable dependiente es el valor que se desea predecir.



Ejemplo de Regresión. Predicción del precio de las casas.



# Tipos de Aprendizaje Supervisado



# Tipos de Aprendizaje Supervisado

## Clasificación:

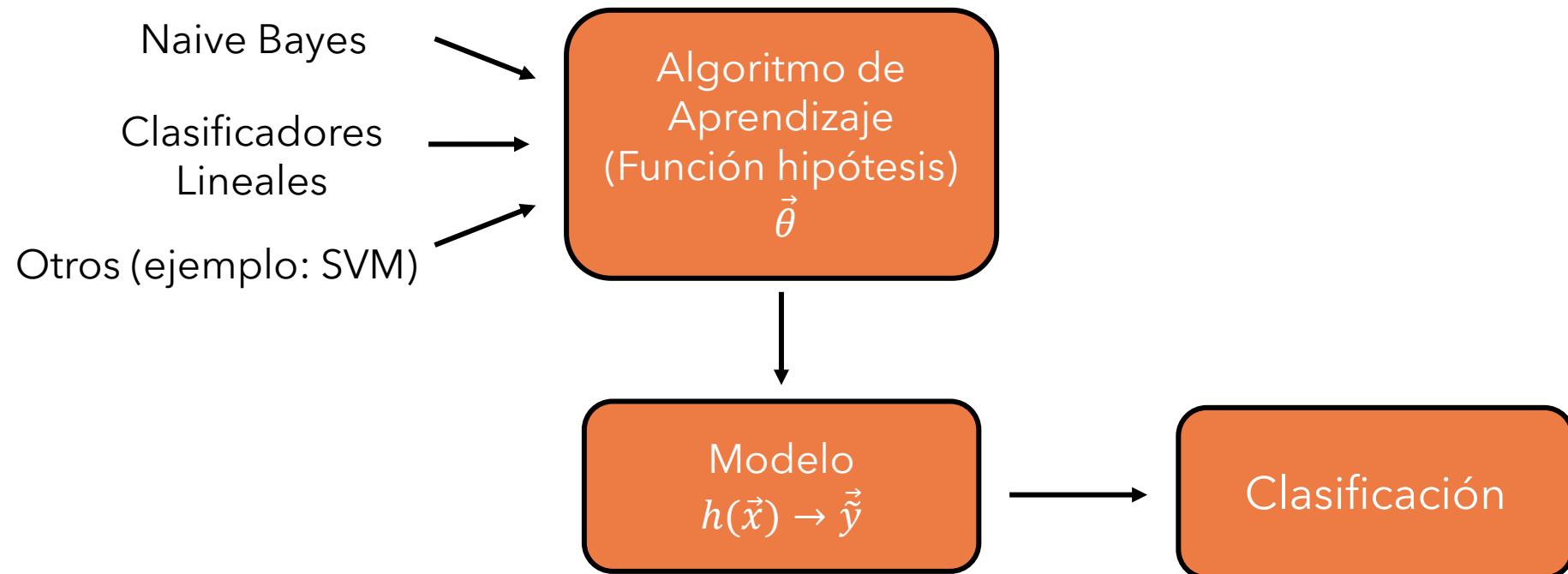
La clasificación es un método de aprendizaje supervisado en Machine Learning en el cual el modelo se entrena para predecir la etiqueta correcta de un conjunto de datos de entrada. Durante el proceso, el modelo se entrena exhaustivamente utilizando datos de entrenamiento y posteriormente se evalúa con datos de prueba. Una vez completado este proceso, el modelo está listo para realizar predicciones sobre nuevos datos no vistos.



Ejemplo de clasificación. Fraude en transacciones bancarias.



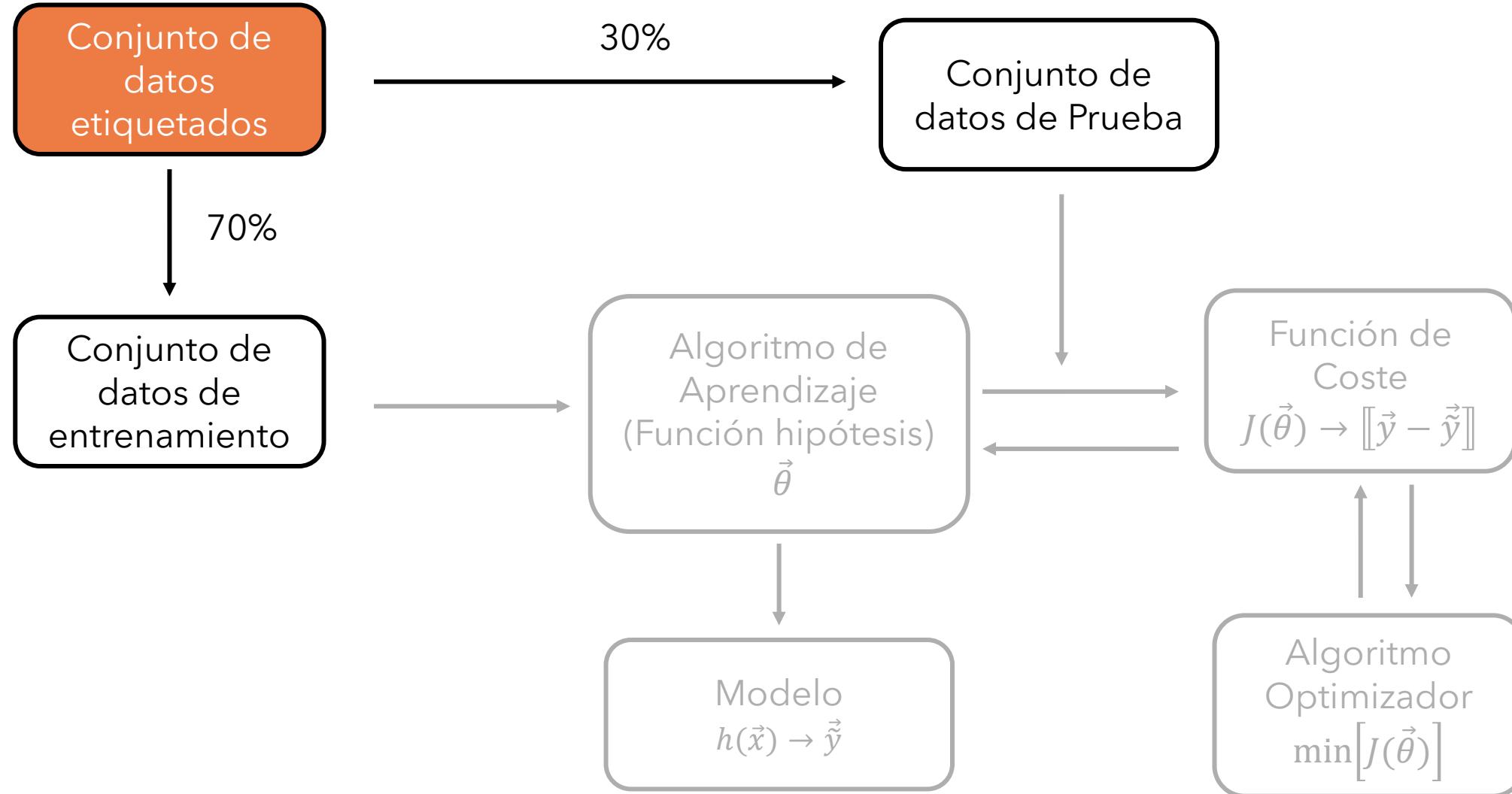
# Tipos de Aprendizaje Supervisado



# REGRESIÓN LINEAL

## (Conceptos Generales)





# REGRESIÓN LINEAL (CONCEPTOS GENERALES)

Datos de Entrenamiento (70%)  
(seleccionados aleatoriamente)

$X = Predictor$

$Y = Respuesta$

$(x_i; y_i) = observación i_{ésima}$

Observacion (xi)	Features (X)	Target Values (Y)
$x_1$	1000	10000
$x_2$	1500	20000
$x_3$	500	5500
...	...	...
$x_{n-2}$	...	...
$x_{n-1}$	...	...
$x_n$	200	2500

Aprendizaje Automático (Machine Learning)



# REGRESIÓN LINEAL (CONCEPTOS GENERALES)

## Datos de Prueba (30%)

(seleccionados aleatoriamente)

$X = Predictor$

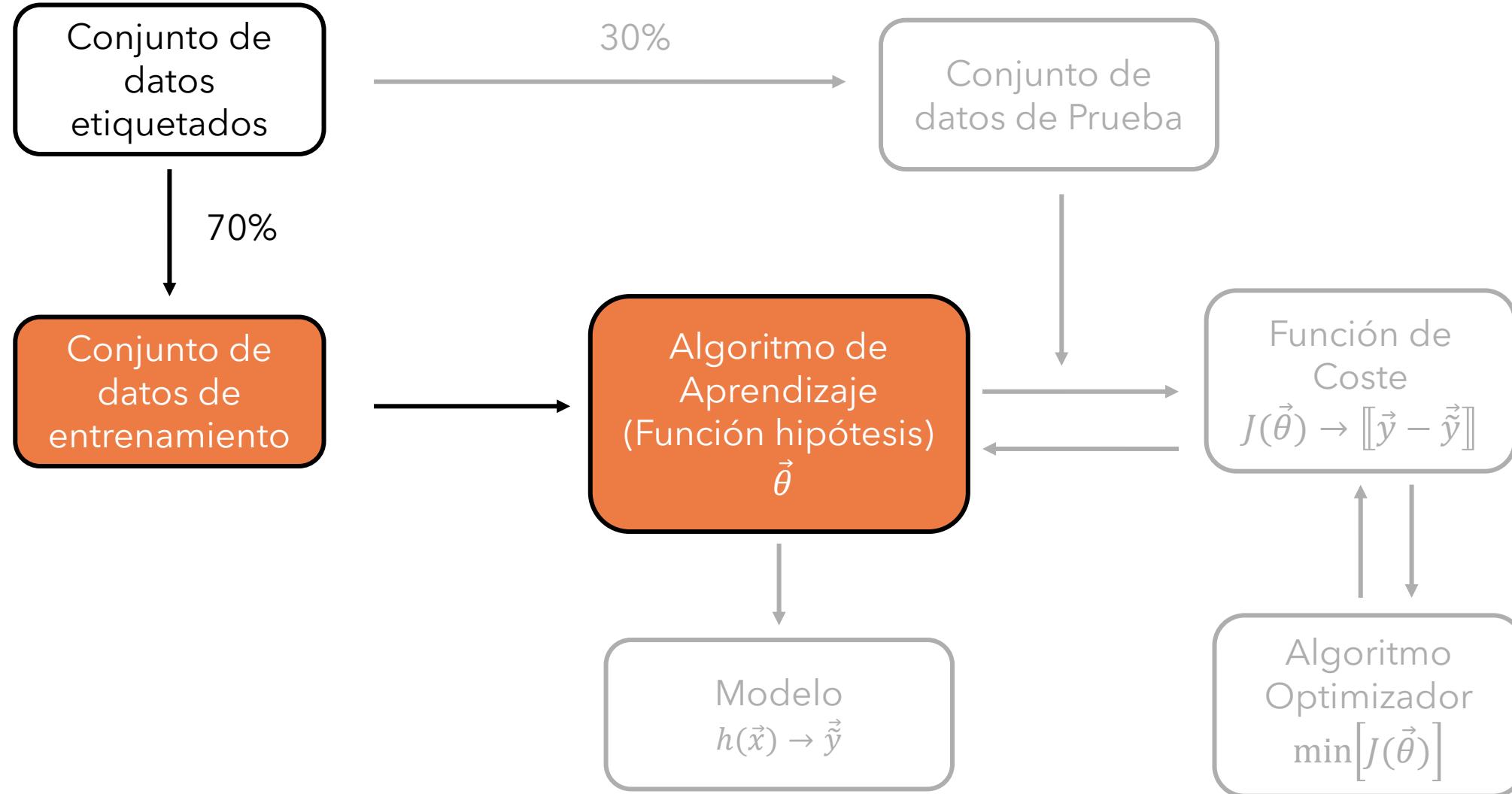
$Y = Respuesta$

$(x_i; y_i) = observación i_{ésima}$

Observacion (xi)	Features (X)	Target Values (Y)
$x_{11}$	1000	10000
$x_{22}$	1500	20000
$x_{53}$	500	5500
...	...	...
$x_{n-20}$	...	...
$x_{n-15}$	...	...
$x_{n-10}$	200	2500

Aprendizaje Automático (Machine Learning)

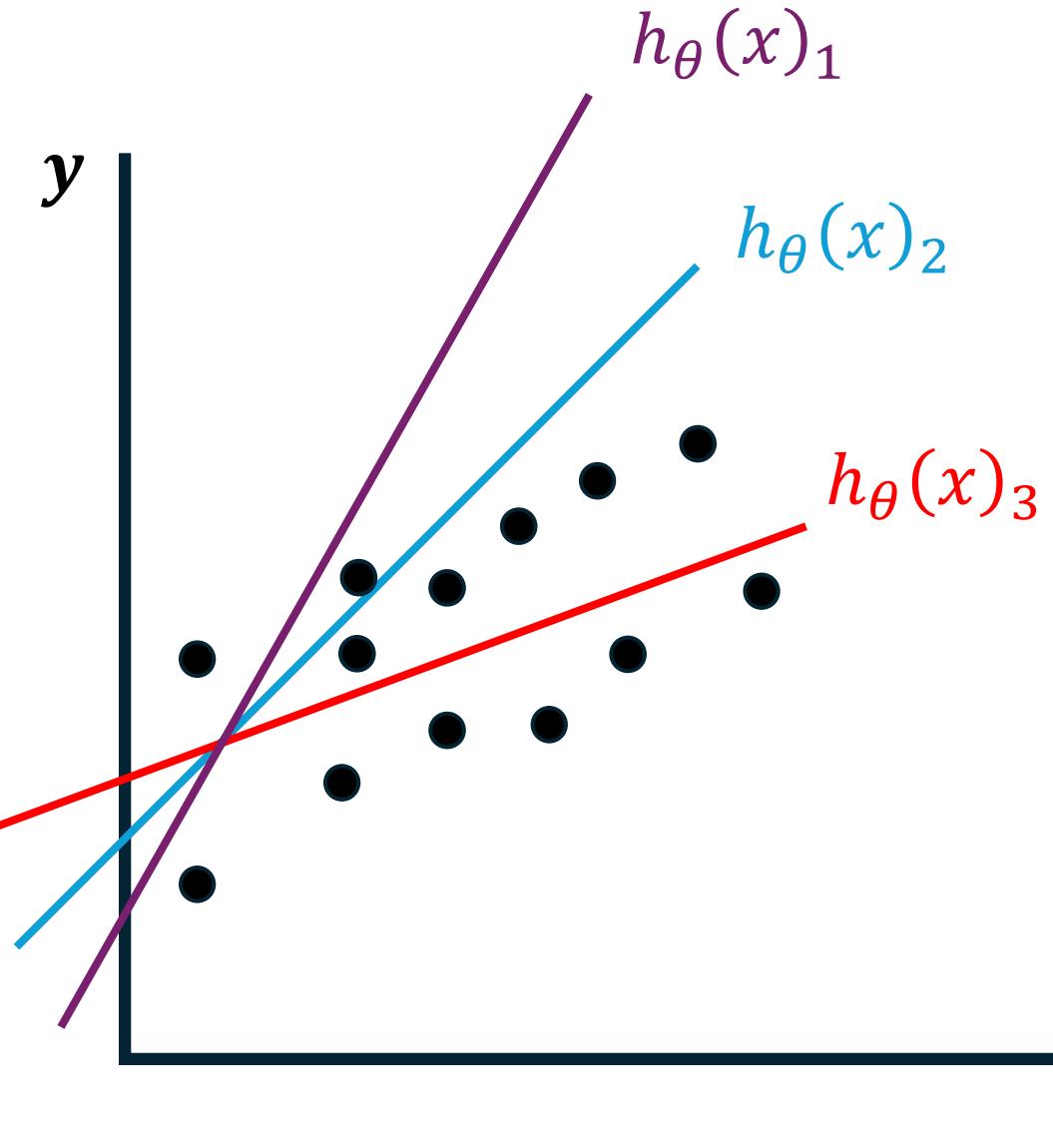




# Algoritmo de aprendizaje

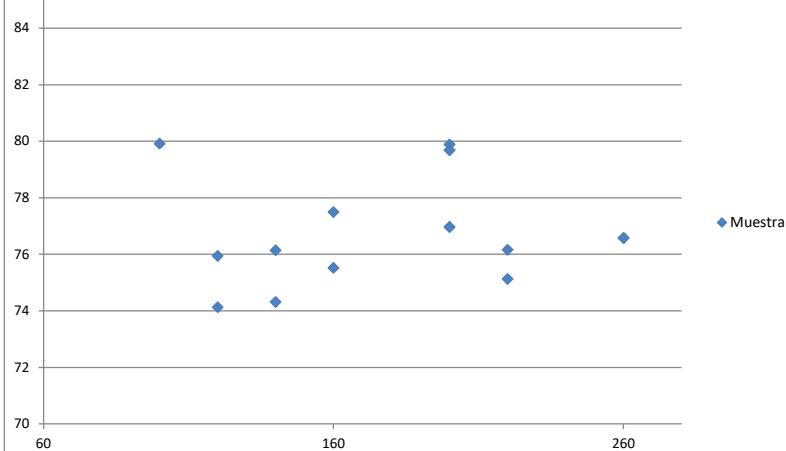
$$h_{\theta}(x) = \theta_0 + \theta_1 x$$

$$h_{\theta}(\vec{X}) = \theta_0 + \theta_1 x_1 + \theta_2 x_2 + \dots + \theta_n x_n$$

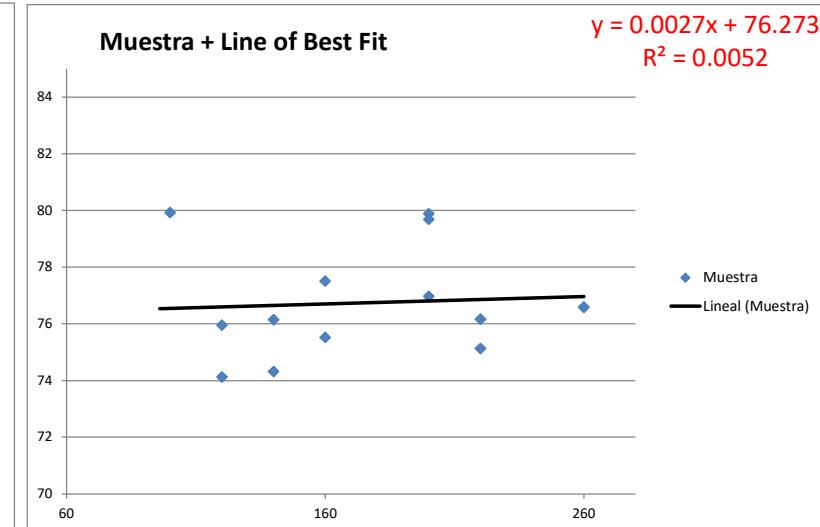


Elemento	Valor X	Valor Y
23	140	74.31975518
14	100	79.92308852
31	160	75.52259959
17	120	75.95312768
65	260	76.58259265
44	200	79.88675267
48	200	76.96845031
16	120	74.12923951
33	160	77.50324772
49	200	79.68899676
55	220	75.13063869
53	220	76.16740298
27	140	76.14701011
65	260	76.58259265

Muestra (Análisis de Correlación)



Muestra + Line of Best Fit



Modelo Poblacional

$$y = 0.0029x + 76.565$$

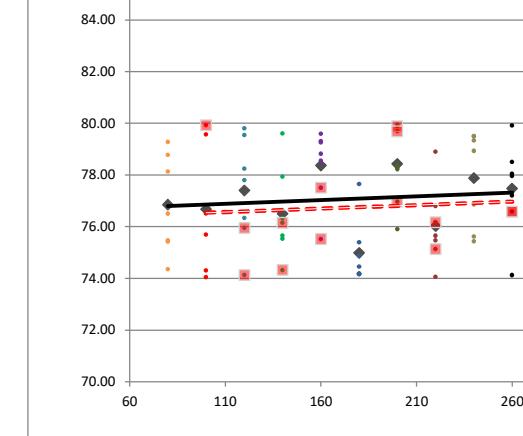
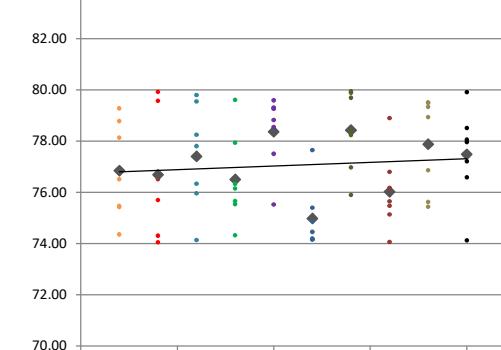
$$R^2 = 0.0267$$

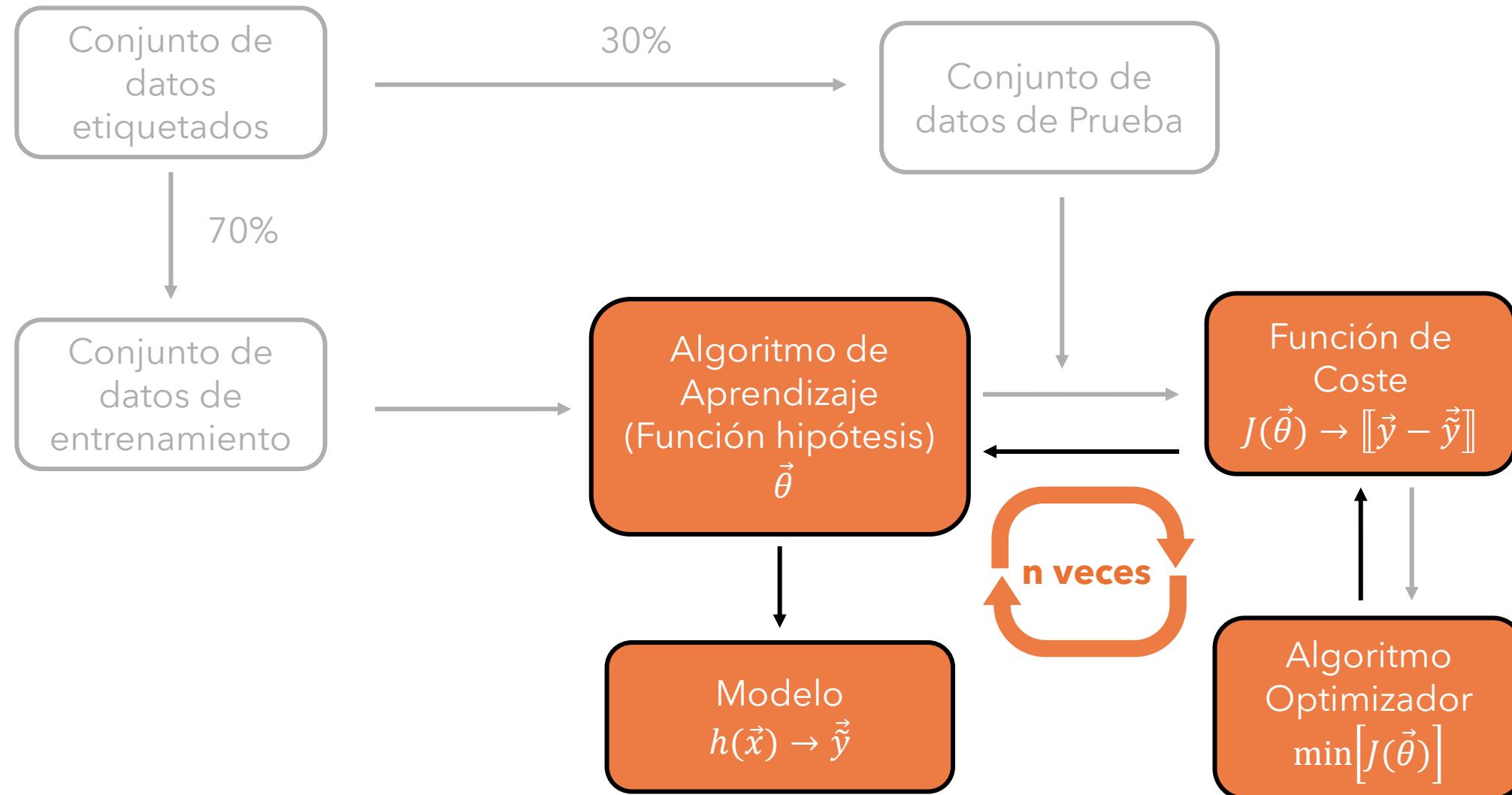
- $X=100$
- ◆  $E(Y_i|X)$
- $X=80$
- $X=120$
- $X=140$
- $X=160$
- $X=180$
- $X=200$
- $X=220$
- $X=240$
- $X=260$
- Lineal ( $E(Y_i|X)$ )

Modelo Poblacional + Line of Best Fit

$y = 0.0029x + 76.565$   
 $R^2 = 0.0267$

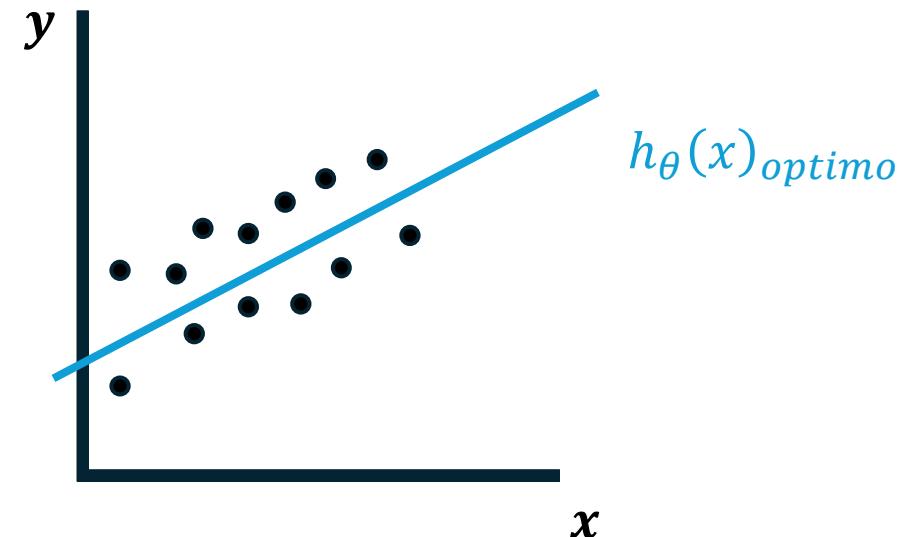
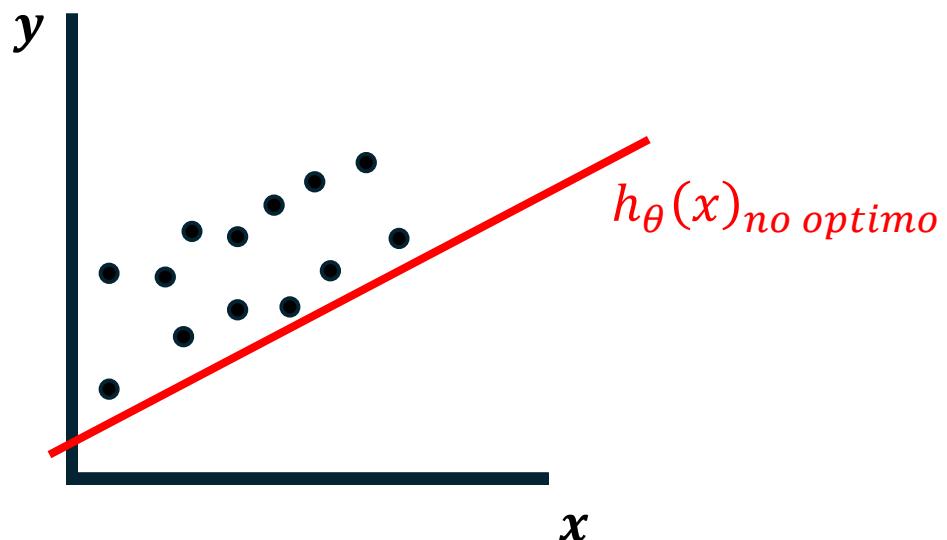
$y = 0.0027x + 76.273$   
 $R^2 = 0.0052$





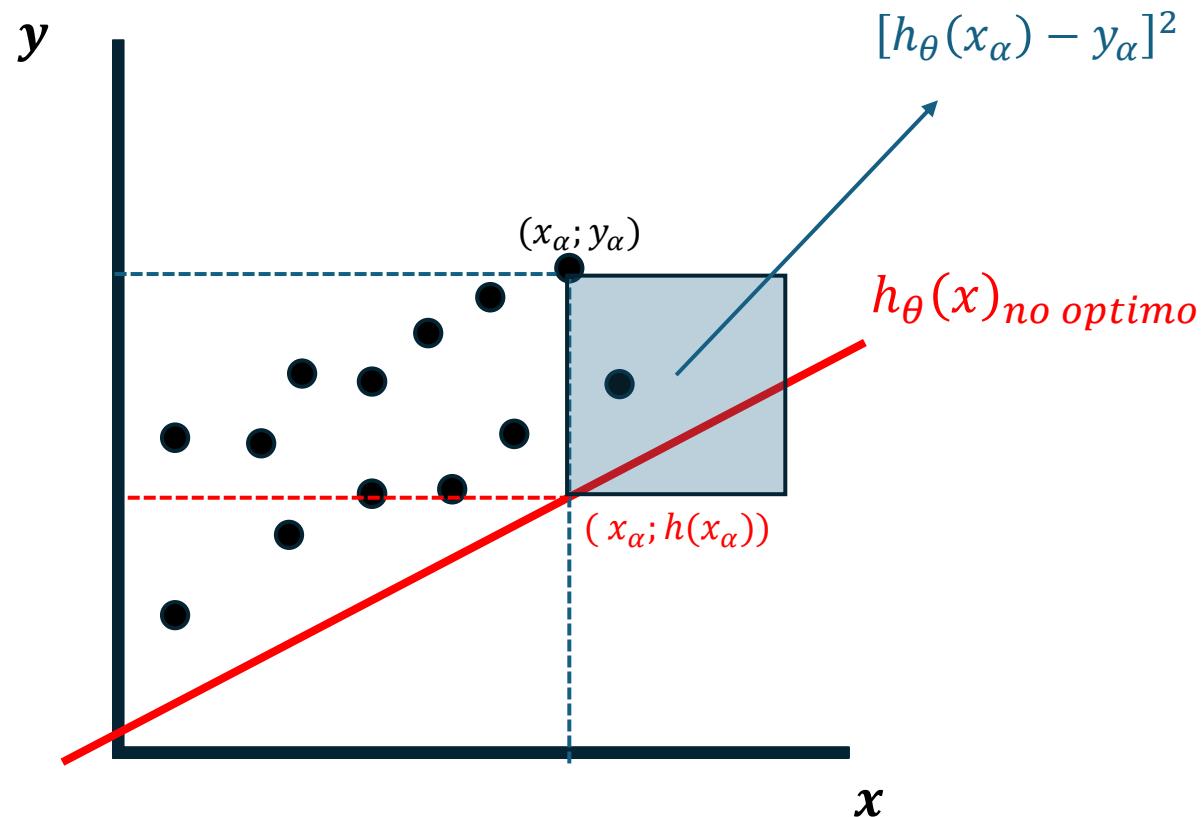
# Construcción del modelo

Buscar los parámetros  $\theta_0$  y  $\theta_1$  que generen la función hipótesis  $h_{\theta}(x)$  que mejor se adapte al conjunto de datos de entrenamiento  $(X; Y)$ . Se debe minimizar una función de coste  $J(\theta)$



# Función de coste

En Regresión Lineal se utiliza una función de coste denominada Mean Square Error (MSE).

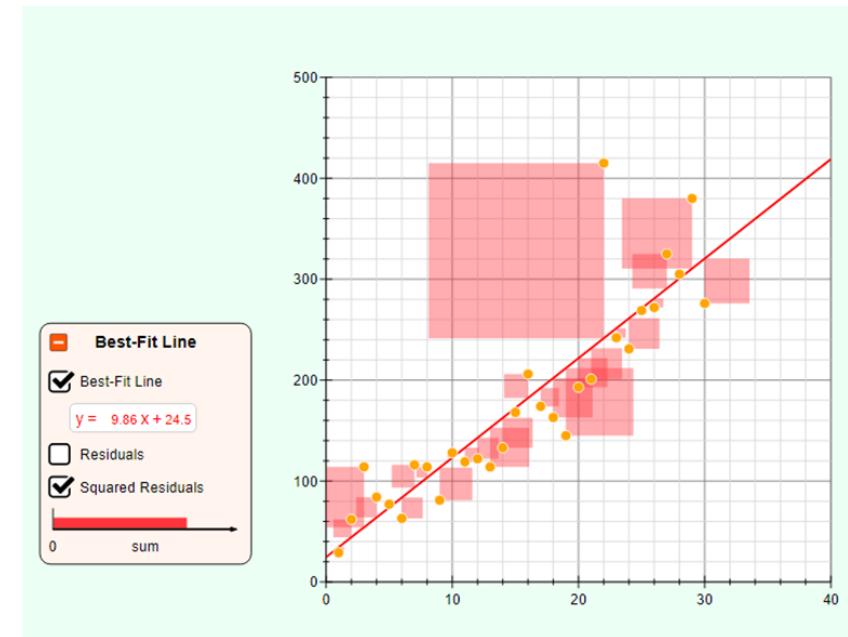
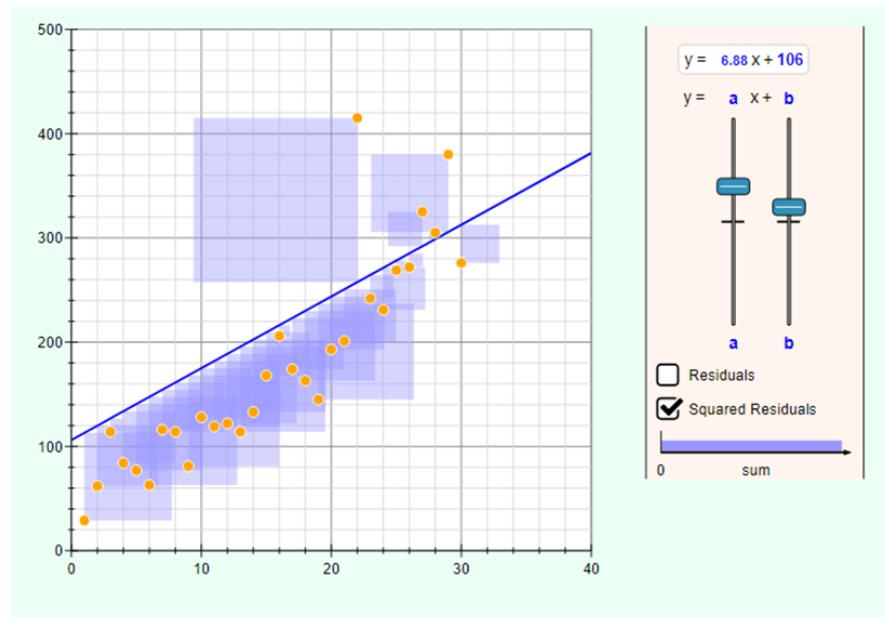


$$J(\theta_0, \theta_1) = \frac{1}{m} \sum_{i=1}^m [h_\theta(x_i) - y_i]^2$$



# Función de optimización (Gradient Descent)

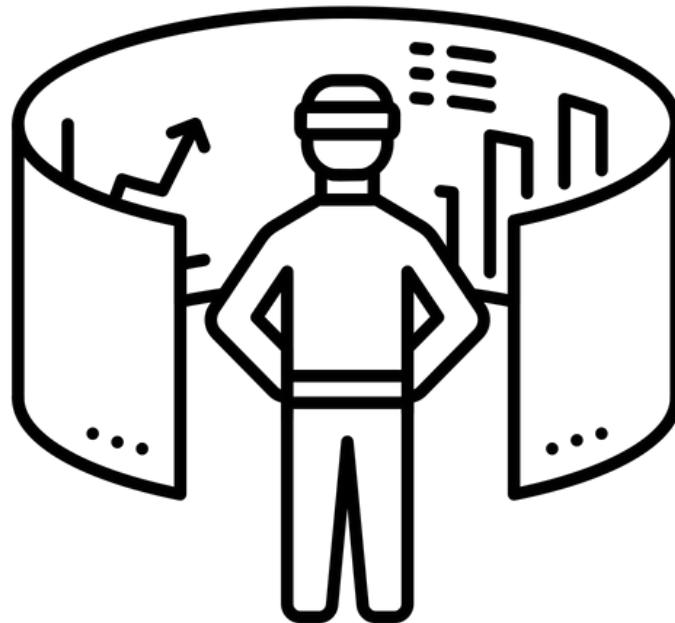
Para llegar a la función de costo óptima (que minimiza MSE) se utiliza un algoritmo denominado Descenso del Gradiente (que veremos en próximas clases). (Clic en la imagen para simulación)



# ¡Alerta de dolor de estómago!

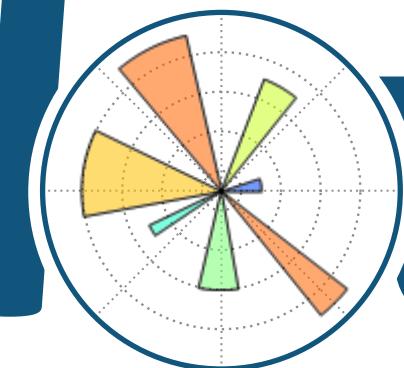


# ¡Vamos a una simulación!



Y después vamos a ver...

matplotlib



**PARTE II**

**CLASE**

**7**

# MÓDULO 02 - CLASE 03

# APRENDIZAJE SUPERVISADO (II) (y obviamente Python)

.UBA económicas  
*Diplomaturas*

IA aplicada a  
Entornos Digitales  
de Gestión



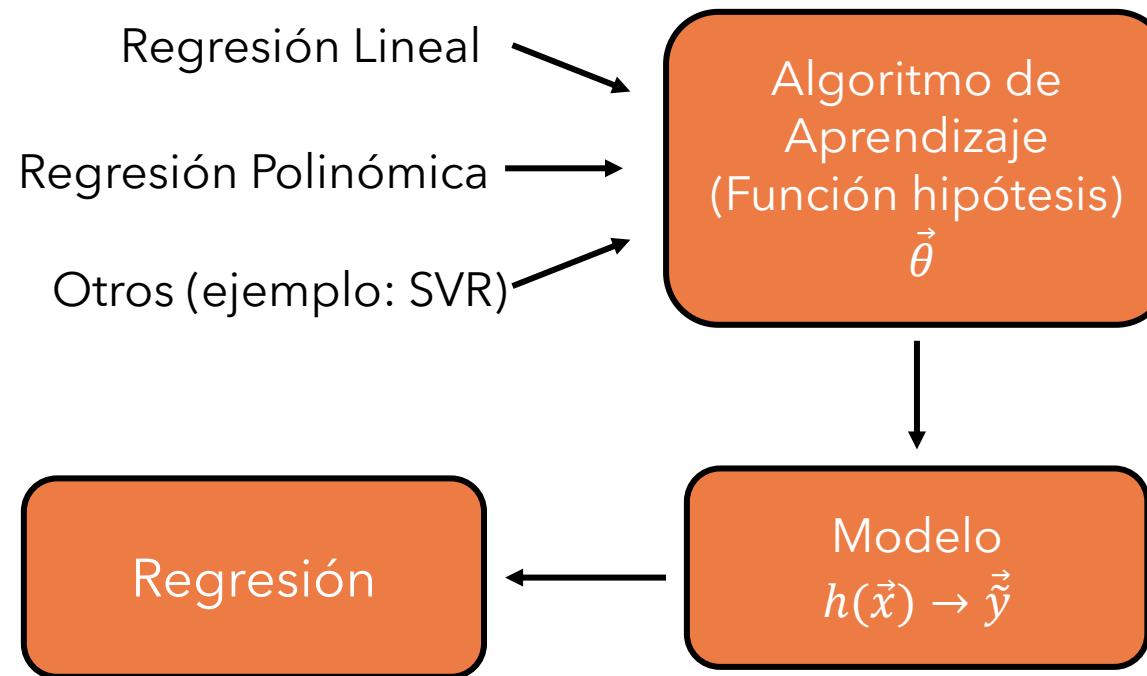
# Debemos recordar este esquema de la clase pasada

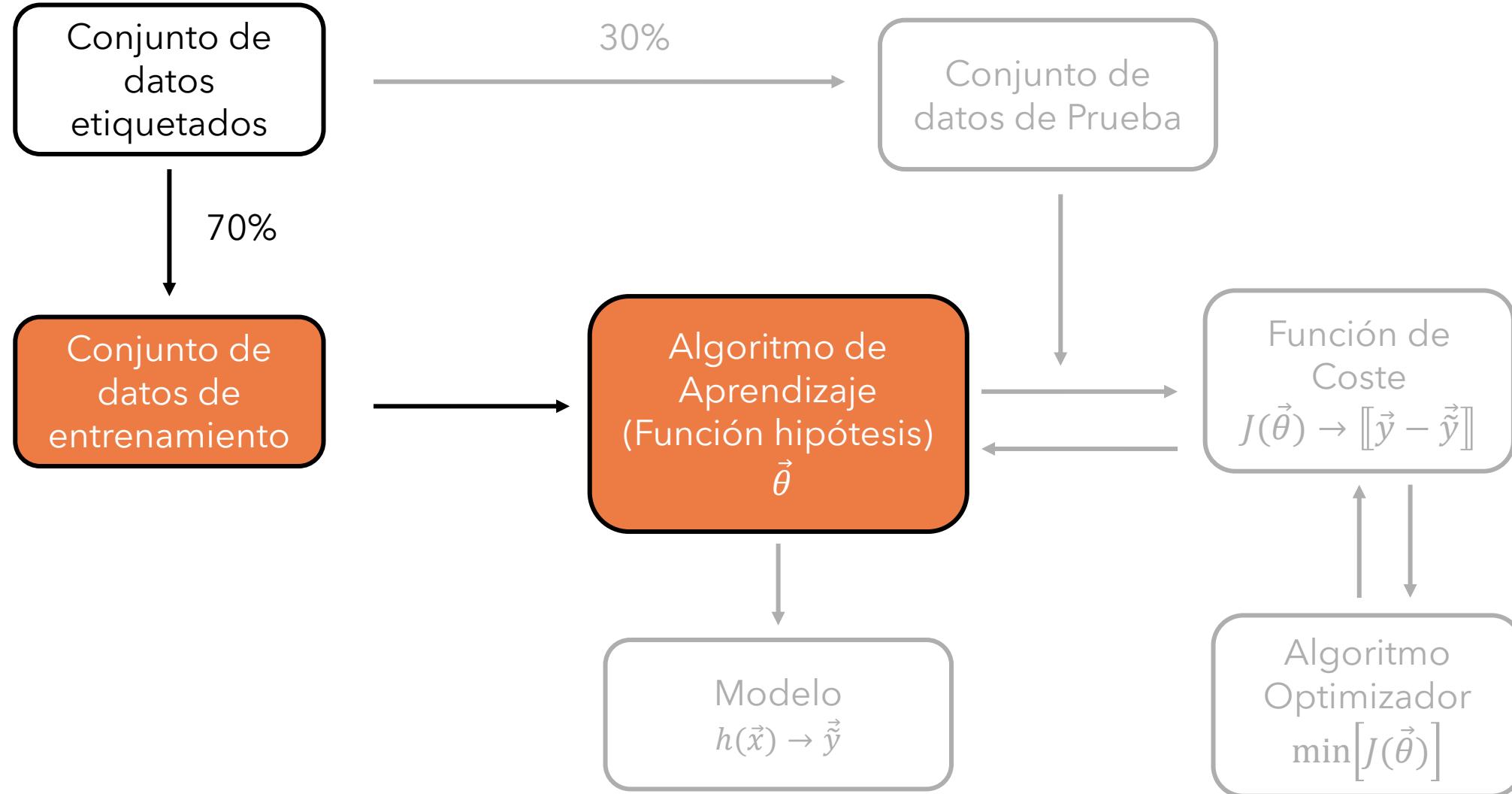


# Que refiere al modelo de Regresión Lineal aplicado a ML



# Tipos de Aprendizaje Supervisado

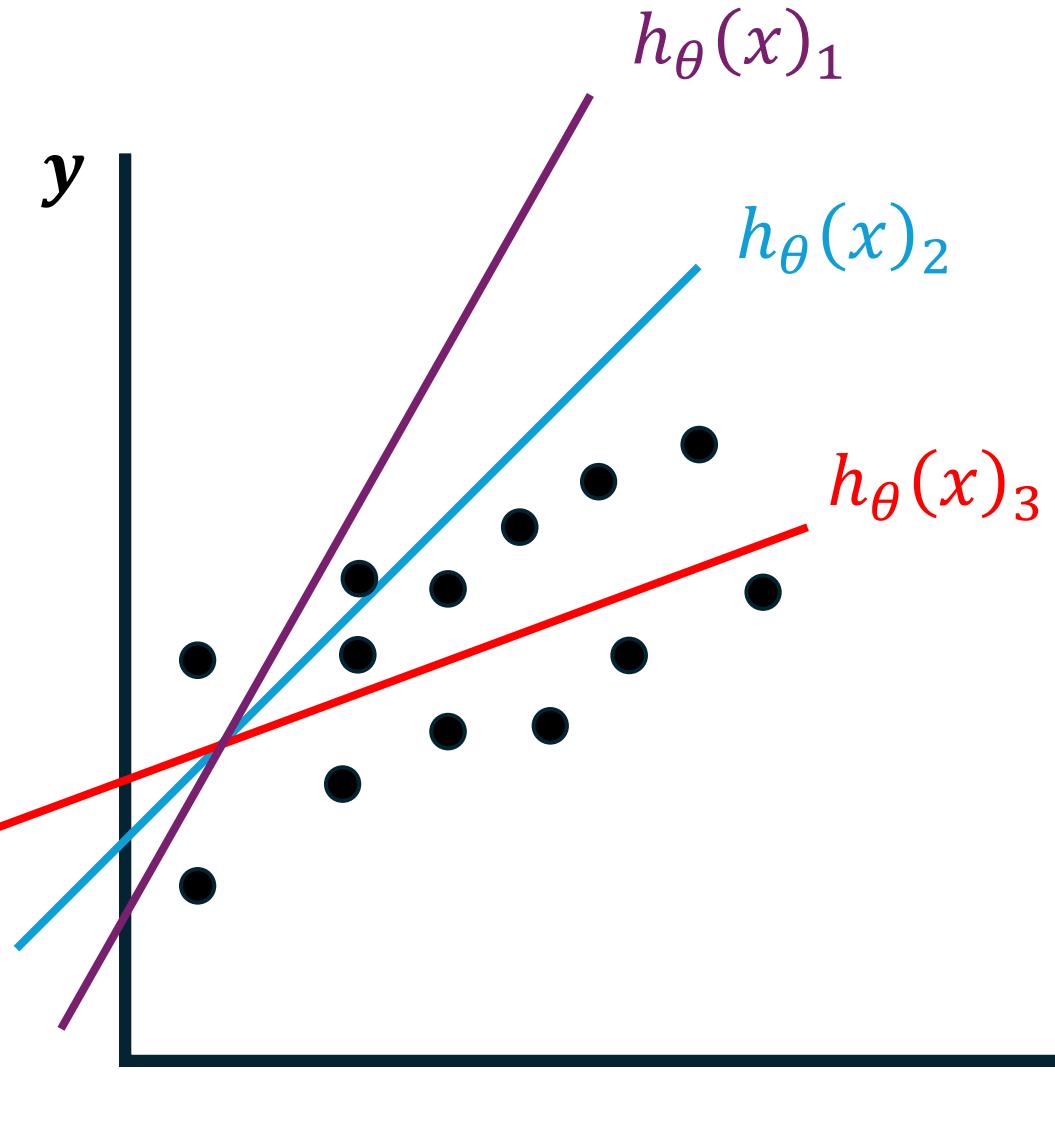


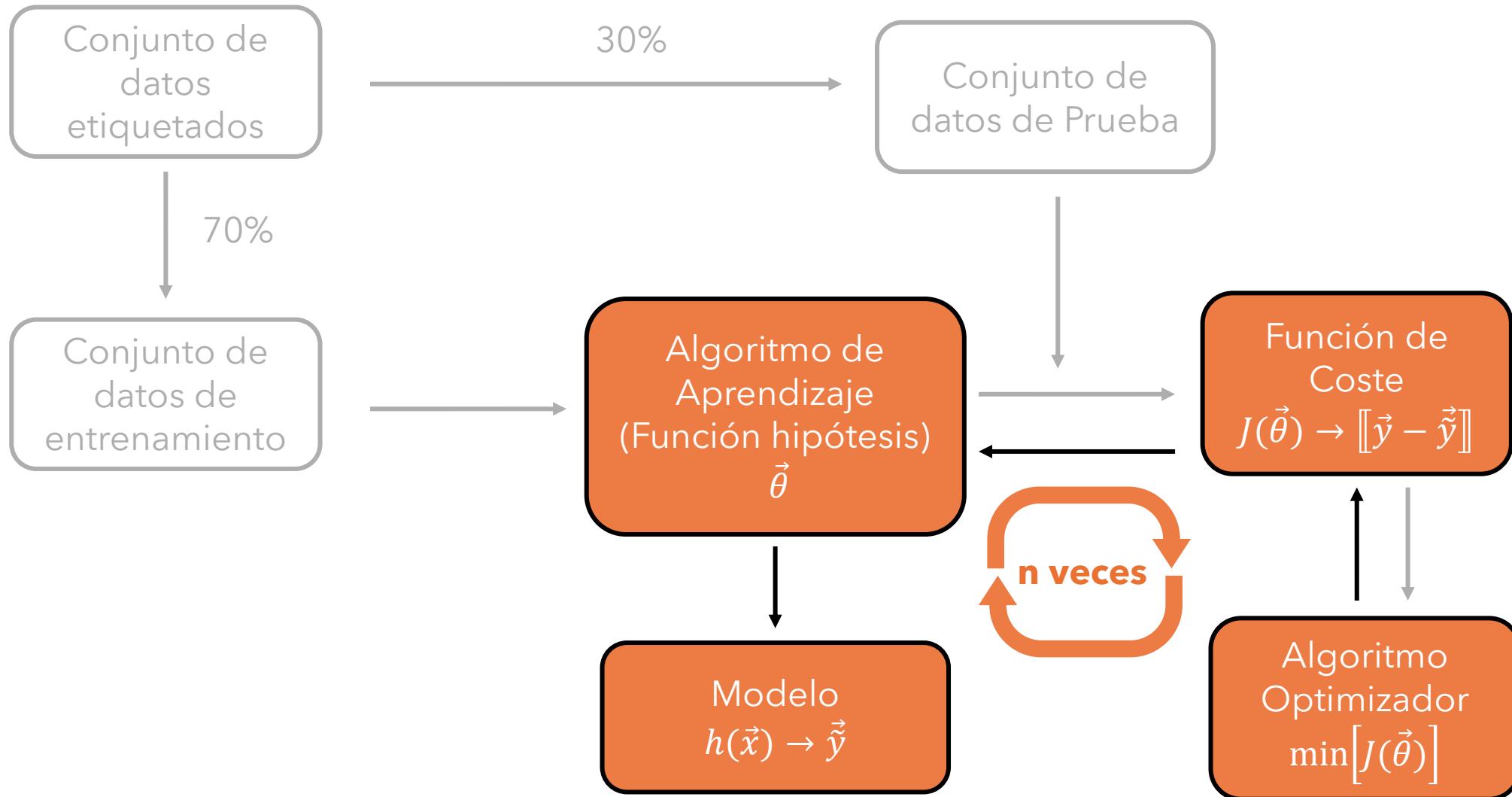


# Algoritmo de aprendizaje

$$h_{\theta}(x) = \theta_0 + \theta_1 x$$

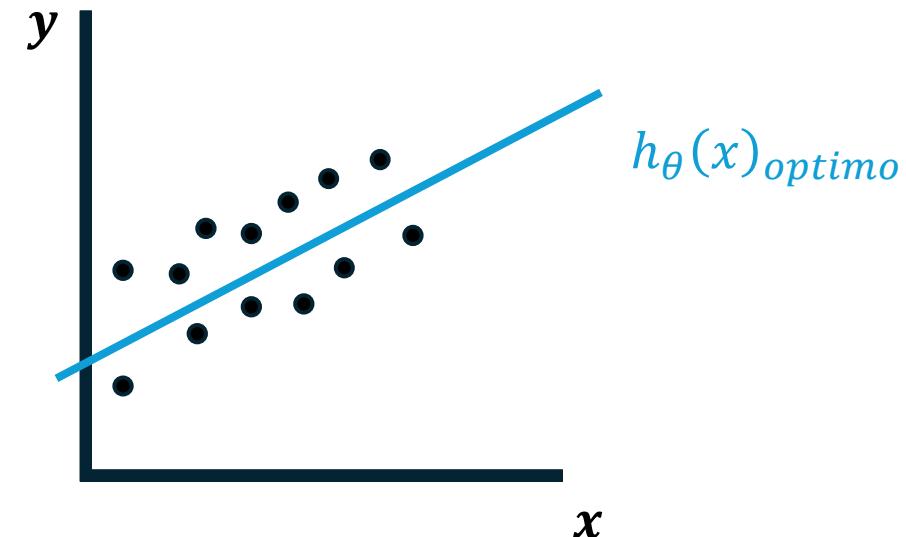
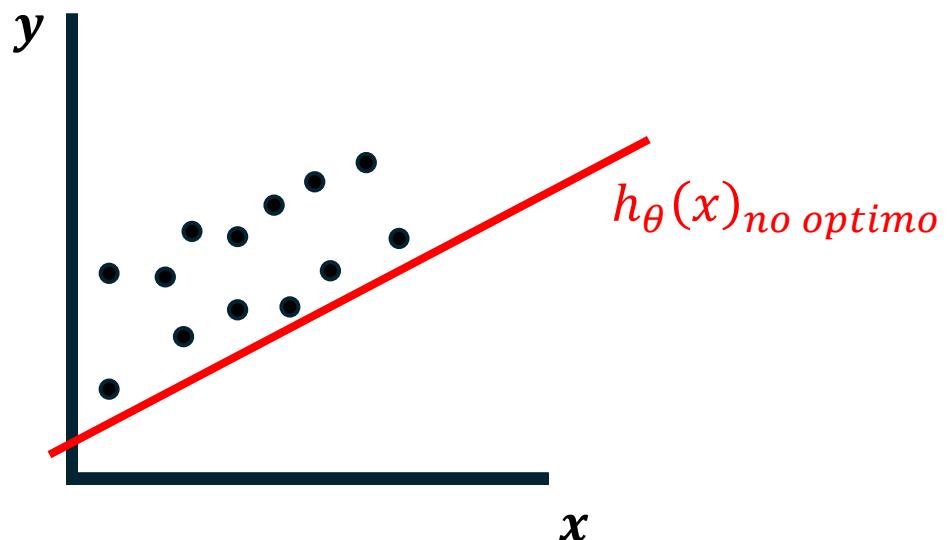
$$h_{\theta}(\vec{X}) = \theta_0 + \theta_1 x_1 + \theta_2 x_2 + \dots + \theta_n x_n$$





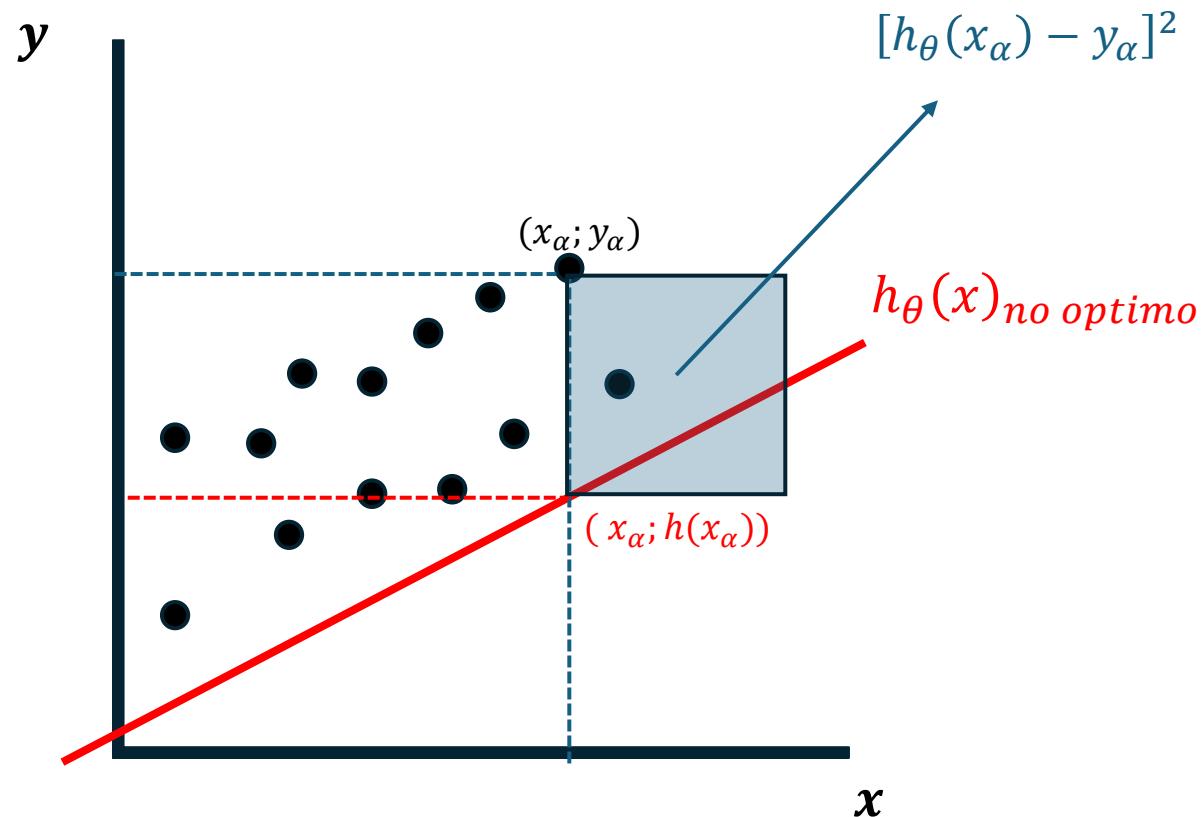
# Construcción del modelo

Buscar los parámetros  $\theta_0$  y  $\theta_1$  que generen la función hipótesis  $h_{\theta}(x)$  que mejor se adapte al conjunto de datos de entrenamiento  $(X; Y)$ . Se debe minimizar una función de coste  $J(\theta)$



# Función de coste

En Regresión Lineal se utiliza una función de coste denominada Mean Square Error (MSE).

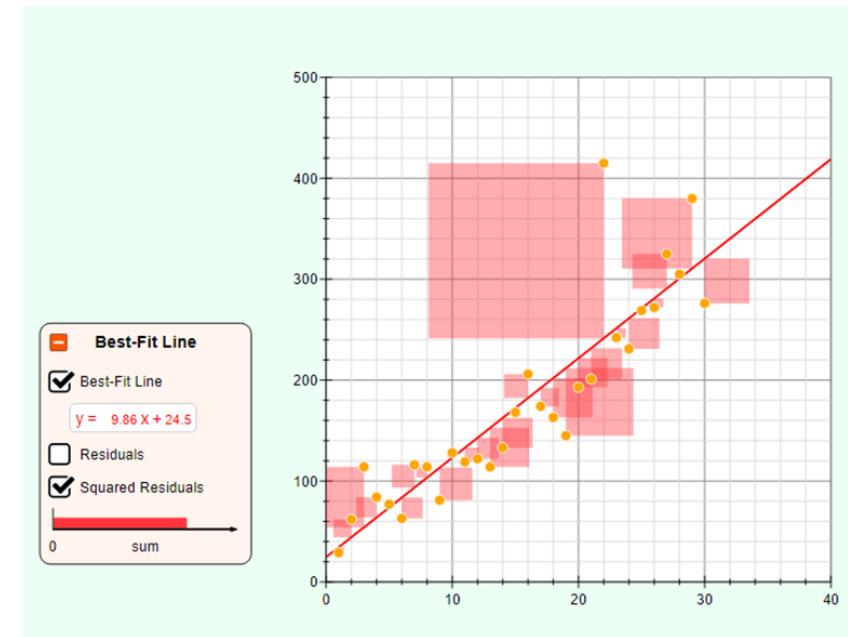
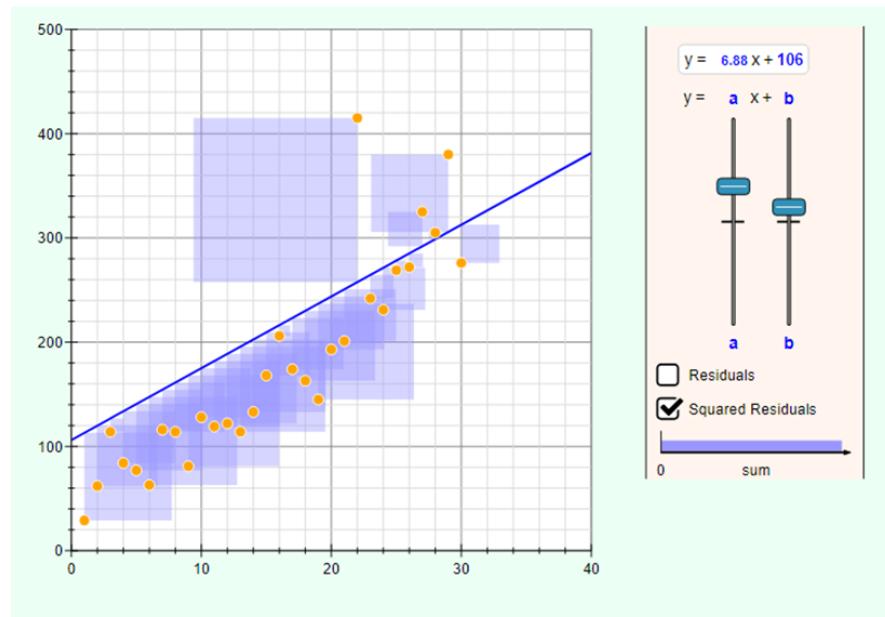


$$J(\theta_0, \theta_1) = \frac{1}{m} \sum_{i=1}^m [h_\theta(x_i) - y_i]^2$$



# Función de optimización (Gradient Descent)

Para llegar a la función de costo óptima (que minimiza MSE) se utiliza un algoritmo denominado Descenso del Gradiente (que veremos en próximas clases). (Clic en la imagen para simulación)



# Hoy trabajaremos con un ejercicio completo de MRL



# Para esto utilizaremos la librería ML de Python:



# MÓDULO 02 - CLASE 03

## Actualidad IA

.UBA económicas  
*Diplomaturas*

**IA** aplicada a  
Entornos Digitales  
de Gestión



# GPT - 4o

14 de mayo 2024

## El Futuro de la Inteligencia Artificial Multimodal

MODULO 02. MACHINE LEARNING

CLASE 03. ACTUALIDAD IA

01

.UBA económicas  
*Diplomaturas*

IA  
aplicada a  
Entornos Digitales  
de Gestión



# Multimodalidad Nativa

## Anteriores

Modelos como GPT-3 y GPT-4 eran principalmente modelos de procesamiento de texto.

## GPT - 4o

Es un modelo nativamente multimodal, lo que significa que puede procesar texto, audio e imágenes de manera integrada y sin necesidad de conversiones intermedias.



# Procesamiento Directo de Audio e Imágenes

## Anteriores

Necesitaban etapas de preprocessamiento y herramientas externas para manejar audio e imágenes.

## GPT - 4o

Puede interactuar directamente con audio e imágenes, permitiendo aplicaciones en tiempo real



# Aplicaciones Ampliadas

## Anteriores

Sus aplicaciones estaban más centradas en el procesamiento del lenguaje natural (NLP)

## GPT - 4o

Amplía su ámbito de aplicación a áreas como la transcripción y traducción en vivo, asistencia virtual personalizada y análisis de imágenes, entre otros, gracias a su capacidad multimodal.



# ¿Nos FALTA algo ?

Aplicación para integrar con PC o Smartphone



# Copilot+

19 de mayo 2024

## El Futuro de la Inteligencia Artificial Multimodal



# Copilot+

## Integración de Windows y GPT - 4o



# Copilot+

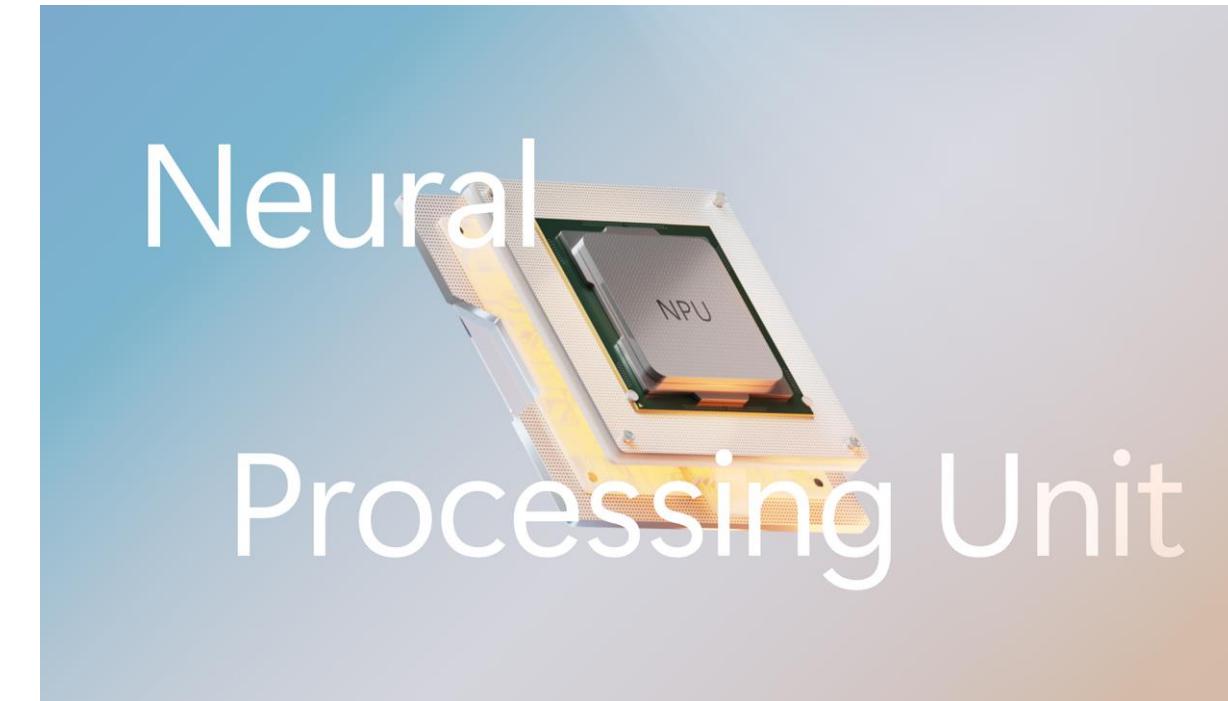
¿Qué podemos hacer ?





# Copilot+Pc

**NPU (Unidad de  
Procesamiento Neuronal**





# Claude 3.5 Sonnet

Supera a los modelos de la competencia y a Claude 3  
Opus



# Claude 3.5 Sonnet

	Claude 3.5 Sonnet	Claude 3 Opus	GPT-4o	Gemini 1.5 Pro	Llama-400b (early snapshot)
Graduate level reasoning <i>GPQA, Diamond</i>	<b>59.4%*</b> 0-shot CoT	<b>50.4%</b> 0-shot CoT	<b>53.6%</b> 0-shot CoT	—	—
Undergraduate level knowledge <i>MMLU</i>	<b>88.7%**</b> 5-shot	<b>86.8%</b> 5-shot	—	<b>85.9%</b> 5-shot	<b>86.1%</b> 5-shot
	<b>88.3%</b> 0-shot CoT	<b>85.7%</b> 0-shot CoT	<b>88.7%</b> 0-shot CoT	—	—
Code <i>HumanEval</i>	<b>92.0%</b> 0-shot	<b>84.9%</b> 0-shot	<b>90.2%</b> 0-shot	<b>84.1%</b> 0-shot	<b>84.1%</b> 0-shot
Multilingual math <i>MGSM</i>	<b>91.6%</b> 0-shot CoT	<b>90.7%</b> 0-shot CoT	<b>90.5%</b> 0-shot CoT	<b>87.5%</b> 8-shot	—
Reasoning over text <i>DROP, F1 score</i>	<b>87.1</b> 3-shot	<b>83.1</b> 3-shot	<b>83.4</b> 3-shot	<b>74.9</b> Variable shots	<b>83.5</b> 3-shot Pre-trained model
Mixed evaluations <i>BIG-Bench-Hard</i>	<b>93.1%</b> 3-shot CoT	<b>86.8%</b> 3-shot CoT	—	<b>89.2%</b> 3-shot CoT	<b>85.3%</b> 3-shot CoT Pre-trained model
Math problem-solving <i>MATH</i>	<b>71.1%</b> 0-shot CoT	<b>60.1%</b> 0-shot CoT	<b>76.6%</b> 0-shot CoT	<b>67.7%</b> 4-shot	<b>57.8%</b> 4-shot CoT
Grade school math <i>GSM8K</i>	<b>96.4%</b> 0-shot CoT	<b>95.0%</b> 0-shot CoT	—	<b>90.8%</b> 11-shot	<b>94.1%</b> 8-shot CoT

\* Claude 3.5 Sonnet scores 67.2% on 5-shot CoT GPQA with maj@32

\*\* Claude 3.5 Sonnet scores 90.4% on MMLU with 5-shot CoT prompting



# Claude 3.5 Sonnet

	Claude 3.5 Sonnet	Claude 3 Opus	GPT-4o	Gemini 1.5 Pro
Visual math reasoning <i>MathVista (testmini)</i>	<b>67.7%</b> 0-shot CoT	<b>50.5%</b> 0-shot CoT	<b>63.8%</b> 0-shot CoT	<b>63.9%</b> 0-shot CoT
Science diagrams <i>AI2D, test</i>	<b>94.7%</b> 0-shot	<b>88.1%</b> 0-shot	<b>94.2%</b> 0-shot	<b>94.4%</b> 0-shot
Visual question answering <i>MMMU(val)</i>	<b>68.3%</b> 0-shot CoT	<b>59.4%</b> 0-shot CoT	<b>69.1%</b> 0-shot CoT	<b>62.2%</b> 0-shot CoT
Chart Q&A <i>Relaxed accuracy (test)</i>	<b>90.8%</b> 0-shot CoT	<b>80.8%</b> 0-shot CoT	<b>85.7%</b> 0-shot CoT	<b>87.2%</b> 0-shot CoT
Document visual Q&A <i>ANLS score, test</i>	<b>95.2%</b> 0-shot	<b>89.3%</b> 0-shot	<b>92.8%</b> 0-shot	<b>93.1%</b> 0-shot





# WOW!



**PARTE II**

**CLASE**

**8**