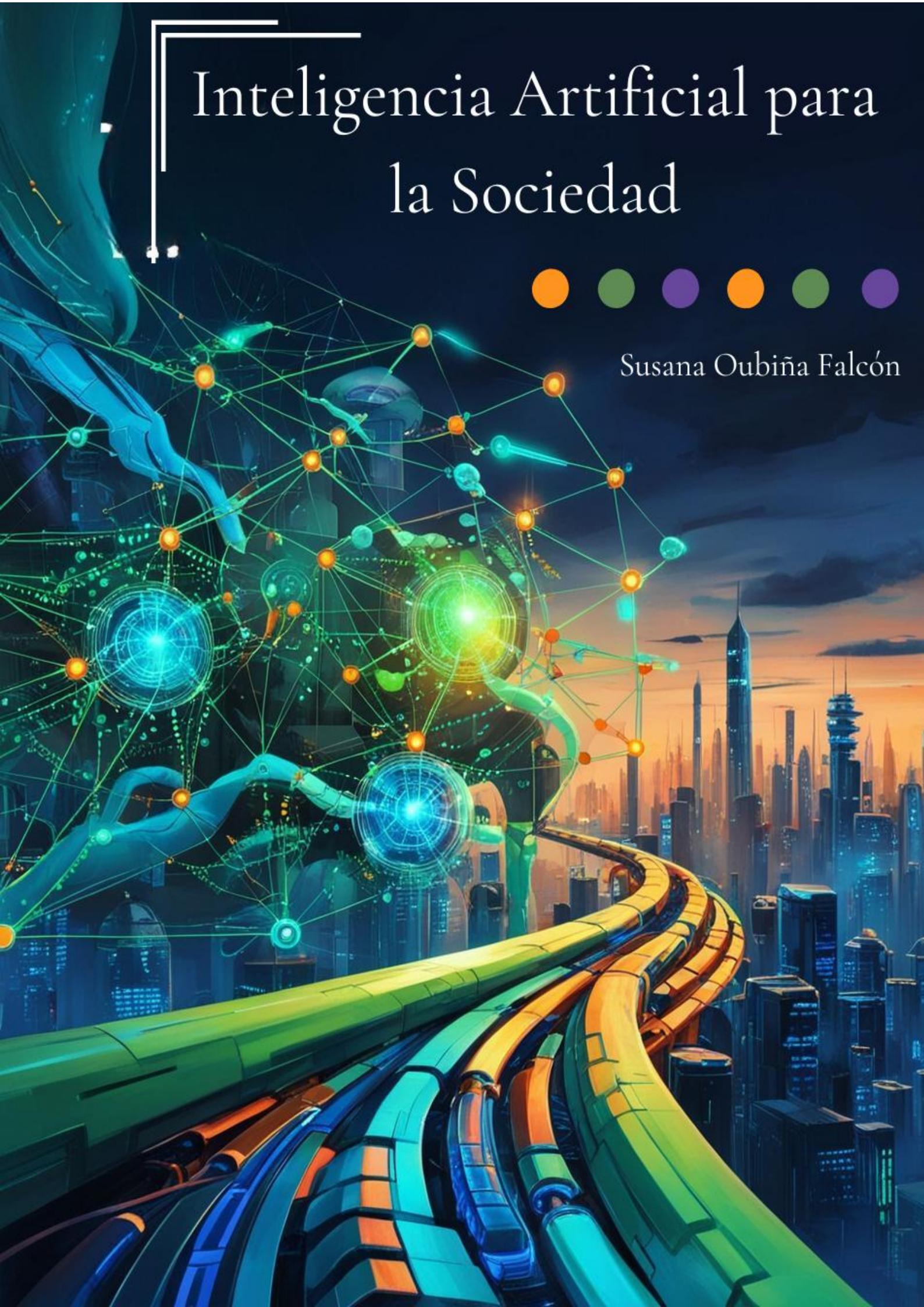


Inteligencia Artificial para la Sociedad



Susana Oubiña Falcón



Índice

1. Introducción	2
2. ¿Qué es la inteligencia y el aprendizaje?.....	2
2.1. ¿Qué es la Inteligencia Artificial?.....	4
2.2. Historia de la IA.....	5
2.3. Evolución de la IA.....	11
2.4. ¿Qué no es Inteligencia Artificial?	15
3. Fundamentos más relevantes de la IA	17
3.1. KDD (Knowledge Discovery in Database).....	17
3.2. Machine learning (Aprendizaje automático)	23
3.3. Paradigmas de aprendizaje	27
3.3.1. Aprendizaje supervisado	29
3.3.2. Aprendizaje no supervisado	32
3.3.3. Aprendizaje reforzado	34
3.3.4. Las Redes Neuronales Artificiales (RNA) y el Deep Learning	36
4. El impacto de la IA.....	47
4.1. Situación actual (Aplicaciones)	47
4.2. Retos que debe solucionar la IA.....	50
4.3. Conclusiones sobre la situación actual de la IA	53
5. Bibliografía.....	57

1. Introducción

En este documento vamos a estudiar los conceptos básicos que rodean el campo del aprendizaje automático a partir del análisis masivo de datos y sus aplicaciones, en lo que se denomina *Inteligencia Artificial*.

Primero nos centraremos en una introducción conceptual e histórica al campo de la IA, junto con la definición más técnica de los elementos básicos de un sistema de este tipo: *KDD* y *Data Mining*, *Machine Learning (ML)*, *los tres paradigmas del aprendizaje*, *Redes Neuronales de aprendizaje (RNA)* y *Deep Learning*. Hablaremos de los problemas que suscita esta nueva tecnología, los retos de futuro que se le plantean y sus aplicaciones actuales. Es decir, nos centraremos en el impacto de la IA, desde el punto de vista de la ética, aspectos legales (responsabilidad) y sostenibilidad.

2. ¿Qué es la inteligencia y el aprendizaje?

Llamamos **inteligencia** a la facultad de la mente que nos permite aprender, entender, razonar, tomar decisiones y formarnos una idea determinada de la realidad.

Actividad inicial

¿Qué es la inteligencia?

Hable con la persona que está a su lado.

Comúnmente, las personas describen la inteligencia como la capacidad de aprender y adaptarse / reaccionar a nuevas situaciones..

Si nos centramos en la facultad de **aprender**, podemos decir que es el proceso a través del cual se modifican y adquieren habilidades, destrezas, conocimientos, conductas y valores. Y esto se produce como resultado del estudio, la experiencia, la instrucción, el razonamiento y la observación. También, se puede definir el **aprendizaje** como un proceso de cambio relativamente permanente en el comportamiento de una persona, generado por la experiencia.

Este proceso de cambio supone un cambio conductual, que debe ser perdurable en el tiempo y ocurre a través de la práctica o de otras formas de experiencia (por ejemplo, observando a otras personas).

El aprendizaje no es una capacidad exclusivamente humana, compartimos esta facultad con otros seres vivos.

Actividad 1: El pedazo de papel “inteligente”

Actividad 1

¿Es inteligente este pedazo de papel?

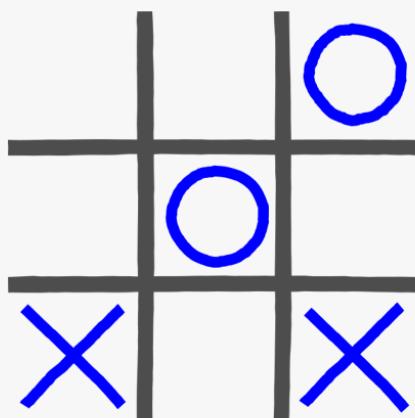
1

El pedazo de papel "inteligente"

Normas

Mi inteligente pedazo de papel nunca ha perdido un juego de ceros y cruces (3 en raya)

- Se necesitan 2 voluntarios.
- 1 jugará en nombre de los humanos.
- 1 seguirá las instrucciones dadas por la hoja de papel inteligente.



2

Tras la ejecución del juego, responde a la cuestión:

- ¿ El pedazo de papel es "inteligente" o "parece inteligente"?
 - El pedazo de papel tenía instrucciones escritas en él. ¿Eso hace que el pedazo de papel sea inteligente?
 - ¿En qué se basa el papel para que nunca pierda?

2.1. ¿Qué es la Inteligencia Artificial?

No es fácil definir la **inteligencia artificial (IA)**, ya que es un concepto en constante evolución. Según las máquinas van adquiriendo nuevas capacidades, la definición se va modificando para poder diferenciarla de la inteligencia humana. Este es el motivo por el que no hay una única definición definitiva.

- Según la R.A.E, la IA es una “*disciplina científica que se ocupa de crear programas informáticos que ejecutan operaciones comparables a las que realiza la mente humana, como el aprendizaje o el razonamiento lógico*”. Podemos entender que el término inteligencia artificial se aplica cuando una máquina **imita** las funciones cognitivas que los humanos asocian con otras mentes humanas, como por ejemplo percibir, razonar, aprender y resolver problemas.
 - Si nos centramos en esa capacidad de **aprender**, es donde actualmente se están produciendo grandes avances que están revolucionando la inteligencia artificial, ya que por primera vez las máquinas (programas informáticos) están adquiriendo la capacidad de aprender a partir del análisis de grandes cantidades de datos.

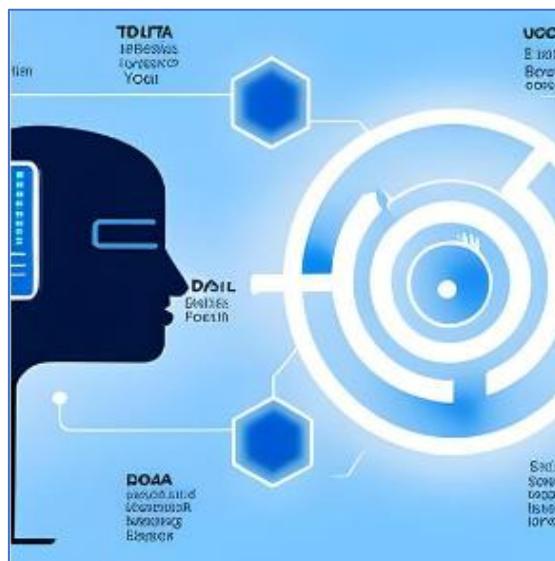
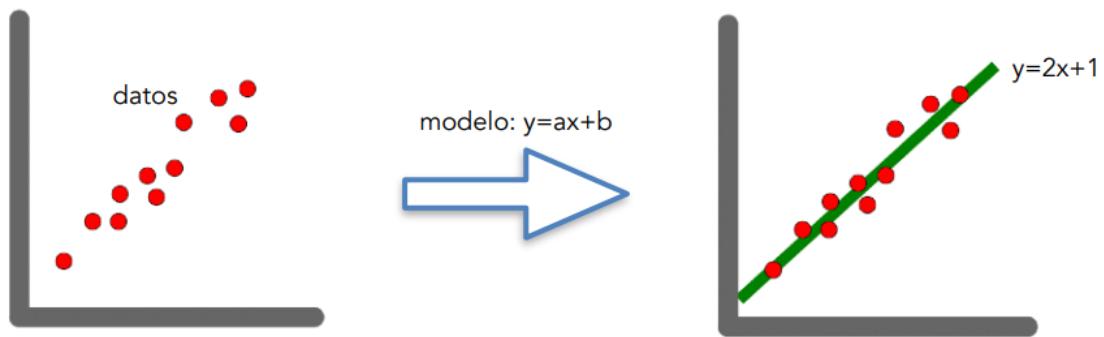


Imagen creada en canva con la aplicación "Texto a imagen"

Enfrentando a estos datos a una serie de modelos predefinidos (que no son más que funciones matemáticas) podremos encontrar y extraer de ellos una serie de patrones que servirán para conocer su comportamiento y poder tener así una explicación aproximada de la realidad.

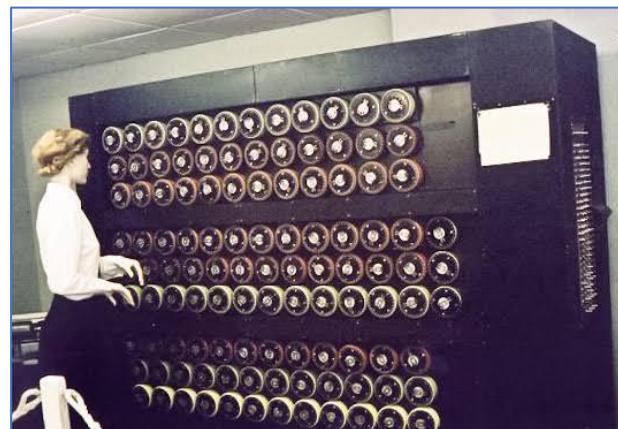
Por ejemplo, si tenemos esta nube de puntos (cada punto es un dato), es fácil entender que una recta (modelo) se ajusta bastante a cómo se distribuyen esos puntos. Solo nos queda encontrar qué valores de a y b (que llamaremos parámetros de la recta) nos dan la recta que mejor se ajusta a esa nube de puntos. Es decir, la recta que menor error posee.



Esto es, en esencia, lo que hace el machine learning: darle datos al sistema informático y este nos devolverá los parámetros que, con un menor error, ajustan esos datos a una función matemática.

2.2. Historia de la IA

Las ideas fundamentales para construir máquinas inteligentes se remontan, varios siglos antes de Cristo, a la lógica y algoritmos de los griegos y a las matemáticas de los árabes. El concepto de obtener razonamiento artificial aparece en el siglo XIV, pero no es hasta la aparición de un artículo científico realizado en 1950 por Alan Turing, matemático británico, que propuso la pregunta «*¿puede pensar una máquina?*», cuando se abre la puerta a una nueva disciplina de las ciencias de la información.



Máquina de Turing

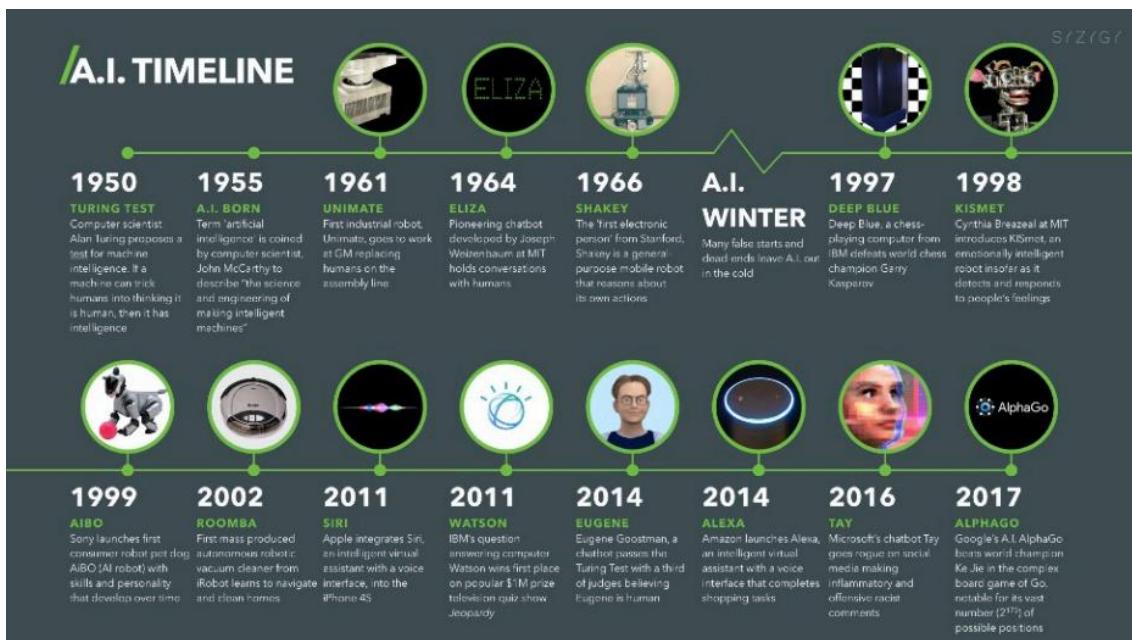
En 1956 se celebró la [Conferencia de Darmouth](#), que fue organizada por John McCarthy (Dartmouth College), Marvin L. Minsky (Harvard University), Nathaniel Rochester (I.B.M. Corporation) y Claude E. Shannon (Bell Telephone Laboratories). En ella procedieron a conjeturar que todos los aspectos del aprendizaje o cualquier otro rasgo de la inteligencia pueden, en principio, ser descritos de una forma tan precisa que es posible crear una máquina que los simule.

Esta conferencia fue el inicio real del trabajo y de las investigaciones en el campo de la IA y es donde John McCarthy habló por primera vez del término *inteligencia artificial*. En esta conferencia se hicieron previsiones tan triunfalistas a diez años, que



jamás se cumplieron, lo que provocó el abandono casi total de las investigaciones durante años.

En los años 80 apareció Japón como nuevo jugador con los sistemas expertos, pero tampoco se llegó a grandes avances. Todo este tiempo se conoce como **el invierno de la IA**. Pero, acercándose al año 2000, con el boom de internet (www) y la disponibilidad que internet nos ofrece relativo a la adquisición de grandes cantidades de datos, unido a la mejora de los equipos informáticos (mejores microprocesadores), provocó el renacimiento o el resurgir de la IA.



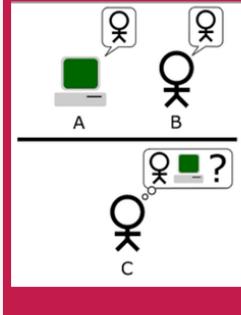
Si analizamos esta [línea de tiempo](#) y ponemos el foco en las IA creadas para enfrentarse a humanos en diferentes juegos de mesa, como [el ajedrez o el Go](#), podemos entender como ha sido la evolución desde que terminó el invierno de la IA.

Actividad 2: Crea una línea de tiempo con AI sobre los chatbots (desde ELIZA al Chat GPT)

Actividad 2
Crea una línea de tiempo con AI sobre los chatbots (desde ELIZA al Chat GPT)
≡

- La línea de tiempo anterior nos menciona a [Eliza](#) y a [Tay](#). ¿Qué eran Eliza y Tay? Ambos fueron chatbots.
- Utiliza [History Timelines](#).

La Prueba de Turing [sienta las bases](#) para el avance de los chatbots.



A
B
C

Prueba o Test de Turing

La "interpretación estándar" de la prueba de Turing, en la cual la entidad C (el interrogador), debe determinar qué entidad —A o B?— es una computadora y cual un ser humano. El interrogador (C) se limita releer las respuestas a las preguntas escritas para tomar esa decisión.

Utiliza History Timelines. (te ayudará esta [guía](#))

Actividad 3: [Historia de los chatbots desde ELIZA al Chat GPT](#) (Descarga el PDF)

- Lee mi información e interactúa en los links que te muestro en el documento.
- Reflexiona y valora ambas informaciones (Actividad 2 y Actividad3) ¿De cuál has aprendido más?, ¿Qué has sacado de interesante en cada una de las actividades 2 y 3?

Actividad 4: Investiga los chatbot Gemini (Bard) y Claude

Actividad 5 (para pensar y razonar)

- Da tu opinión sobre la información que te proporciono en los siguientes links:

[Amores modernos: ¿y si me enamoro del chatbot?](#)

[Cómo tener 20 mil novios y no morir en el intento: créate un clon digital \(y hasta ganarás dinero\)](#)

["Mi novio es un chatbot": Xiaolce, el sistema de inteligencia artificial que ha enamorado a China](#)

Actividad 6: Crea tu propio Chatbot



Crea un Chatbot con la siguiente finalidad:

- Debe ser capaz de interactuar con personas mayores
- Debe ser capaz de ayudarle en algo

Necesitaremos crear un “árbol de decisión”

Usaremos el enfoque “Top down” que es el que se basa en reglas

Dos enfoques para desarrollar la IA

TOP DOWN A partir de reglas

Aplicando las leyes de la física (dinámica) se obtiene la siguiente fórmula para saber la velocidad que alcanza un cuerpo al caer de una altura h :

$$v = \sqrt{2gh}$$

BOTTOM UP A partir de datos

Se suelta la pelota desde varias alturas, cuantas más mejor, y se mide la velocidad. Con el conjunto de datos obtenido se construye un **modelo matemático** que genere el resultado para cualquier altura.

$$v = M(h)$$

Para crear tu chatbot usarás la aplicación [Landbot](#)

1. [Construye un chatbot con IA en 3 minutos](#)
2. [Crea tu propio chatbot en minutos. Tutorial de Landbot](#)
3. [Cómo crear tu chatbot en 10 minutos](#)
4. [Guía Landbot](#)

Ejemplos de árbol de decisión (modelo Top down):

[Árbol de decisión simple - v2 on Scratch \(mit.edu\)](#)

[Akinator, el genio de la web](#)

Saber más: Otra plataforma que me parece sencilla a la hora de crear ChatBots es [Voiceflow](#).

En el siguiente [link](#) se explica cómo utilizarla, creando varios ejemplos.

Ajedrez y Go

Objetivo: matar al Rey



Objetivo: ganar territorio usando el menor número de piezas

Entre los años 1996 y 1997, las miradas de medio mundo se centraron en un tablero de ajedrez. A un lado, el ruso Garry Kasparov, campeón desde 1985 hasta 1993 y considerado como el mejor jugador de todos los tiempos. Al otro, Deep Blue (1996) y Deeper Blue (1997), dos superordenadores diseñados por IBM, capaces de calcular 100 millones de posiciones por segundo (Deep Blue) y 200 millones (Deeper Blue).

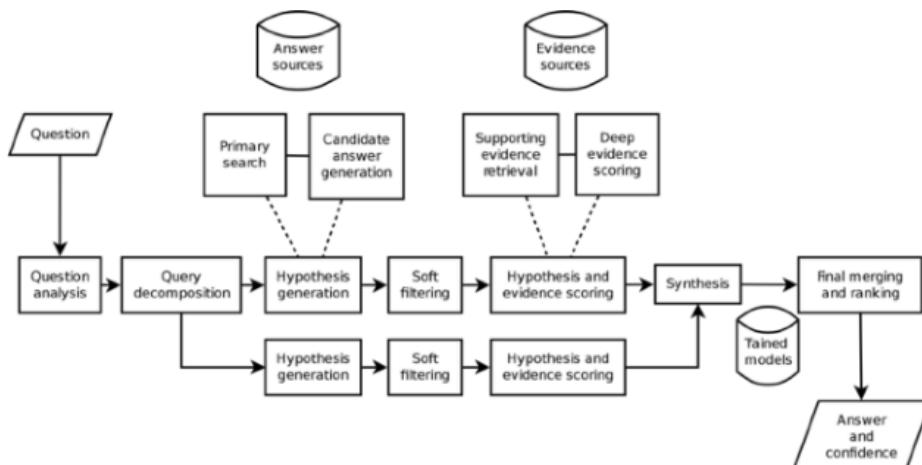
Deep Blue no venció, pero si Deeper Blue en el año 1997. La derrota del mejor jugador de mundo de ajedrez contra el supercomputador de IBM fue un hito que popularizó la idea de que las máquinas sí podían pensar. Sin embargo, no hubo mucha inteligencia en estos ordenadores, sólo "fuerza bruta". Su hardware de 30 nodos (cada uno con 30 microprocesadores P2SC de 120 MHz, ampliados con otros 480 procesadores VLSI) servía la suficiente potencia electrónica para que su software pudiera calcular 200 millones de operaciones por segundo, lo que le permitía, a su vez, escoger siempre la mejor opción entre las miles de partidas que tenía almacenadas en su memoria. Simplemente **imitaba** una solución tomada por un humano.





Supercomputador Watson de IBM

El siguiente paso lo dio otro supercomputador de IBM llamado Watson, quien en 2011 ganó un popular concurso televisivo en USA enfrentándose a los dos mejores humanos, respondiendo a preguntas que se le hacían en lenguaje natural. En esta ocasión, Watson no sólo logró ganar gracias a su superior velocidad de cálculo, sino que también se sirvió de una inmensa base de datos (en la que se incluyeron encyclopedias, diccionarios, tesis, artículos de noticias, y obras literarias, al igual que bases de datos externos, taxonomías y ontologías) que le convertía en una máquina "sabedotodo", capaz de responder a cualquier pregunta antes que ningún humano. El avance de Watson consistía en la capacidad de buscar información de forma rápida en diferentes fuentes y entender preguntas con juegos de palabras y dobles sentidos. **Era capaz de procesar el lenguaje natural.**

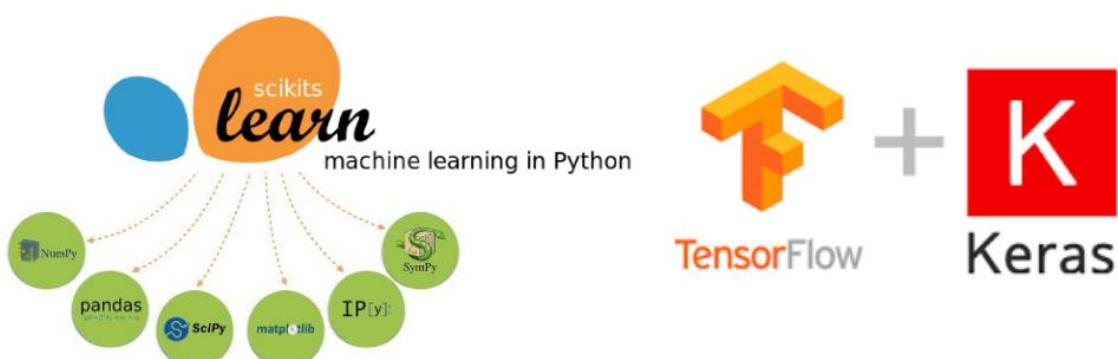


La arquitectura de alto nivel empleada por la tecnología DeepQA, que IBM utilizó específicamente para Watson

Pero, ninguno de estos supercomputadores de IBM demostraron realmente un ápice de inteligencia artificial. La máquina que venció a Kasparov se basaba en movimientos predefinidos y cálculos de probabilidades, nada más, mientras que Watson se valía de una enorme base de datos para recabar la información que necesitaba. **No había aprendizaje.**

Y esto ha sido así hasta que llegó el software de Google llamado AlphaGo Zero, que se centra en un juego de mesa, el Go (¿Cómo jugar al Go?), un juego chino de tablero estratégico para dos jugadores, cuya complejidad es, según los expertos, mucho mayor que la del ajedrez.

AlphaGo Zero **aprende por sí mismo**, empezando desde el juego aleatorio, sin supervisión o uso de datos humanos. Este programa, en tan sólo tres días, no sólo aprendió las reglas y fundamentos básicos, sino que definió nuevas líneas, aperturas y patrones que los jugadores humanos jamás habían llegado si quiera a plantearse. Los resultados fueron demoledores: ganó 100 partidas de 100. No fue derrotado ni una sola vez. La capacidad de aprendizaje de estos algoritmos está siendo y seguirá siendo una revolución de la que vamos a ser testigos y si queremos, también seremos protagonistas. La buena noticia es que existen multitud de librerías para Python que nos facilitan el trabajo con este tipo de algoritmos de **aprendizaje automático**, desde las más sencillas regresiones lineales hasta redes neuronales para aprendizaje profundo.



En noviembre del 2022, la empresa OpenAI dio a conocer ChatGPT. Tras su lanzamiento, la sociedad se dio cuenta de su utilidad y de la importancia que tendría en un futuro inmediato: *la IA estaría presente si o si en nuestro futuro*. Curiosamente, la familia GPT y otras como BERT, T5, etc.., están basadas en una arquitectura de Transformers, creada en el 2017, y que no despuntaron hasta que el Chat GPT, a finales del 2022, revolucionó el mundo.

2.3. Evolución de la IA

Si nos fijamos en la historia anterior y analizamos el presente, podemos clasificar la IA por **capacidades**. Es decir, por el grado en que la IA puede imitar los procesos del pensamiento humano. Este nivel de inteligencia se clasifica en: IA débil, IA general o fuerte y superinteligencia.

- **IA Débil:**

Se refiere a sistemas que realizan tareas específicas. Pueden llegar a superar al humano, pero en una tarea concreta. No tienen la capacidad de razonar, aprender o adaptarse a nuevas situaciones de la misma manera como lo haría un ser humano. Eso sí, ejecuta perfectamente esa tarea o tareas específicas.

Ejemplos: Deep Blue y Alpha Go

Ojo: Esta IA se denomina débil porque no analiza las consecuencias que puede tener su implantación, tal cual lo hacemos los humanos, considerando un análisis más amplio de efectos, relaciones e implicaciones. Por eso, es una opción que podemos considerar peligrosa si no se utiliza con prudencia.

- **IA Fuerte:**

Se refiere a sistemas que tienen la capacidad de realizar cualquier tarea intelectual que un ser humano pueda hacer. Tienen la capacidad de razonar, aprender o adaptarse a nuevas situaciones de la misma manera como lo haría un ser humano. Es decir, sería capaz de analizar cualquier situación y deducir el conjunto de acciones más adecuadas para dicha situación y contexto. Sería capaz de conducir un coche o resolver una ecuación matemática o mantener una conversación, entre muchos ejemplos. Esta IA puede aprender de las situaciones.

Ejemplos: [Estamos lejos de conseguirlo](#) (entiendo que cada vez estamos más cerca)

- **Superinteligencia:**

La Superinteligencia Artificial se asocia con una IA actualmente hipotética en la que los sistemas de IA no sólo podrían realizar tareas que normalmente requieren inteligencia humana, sino que también superarían a los humanos en prácticamente todos los aspectos de la inteligencia.

Actualmente, la superinteligencia **es una posibilidad teórica, más que una realidad práctica**.

Hoy en día solemos clasificar los tipos de IA atendiendo a **si son capaces o no de adaptarse al entorno** en: IA Clásica e IA Moderna o generativa.

- **IA Clásica:** Están bien delimitadas y resuelven ciertos problemas de forma ejemplar, pero **no son capaces de adaptarse a su entorno** (y mucho menos de lograr que el entorno se adapte a ellas).

Ejemplo: Siri, que ha sido entrenado para seguir unas reglas específicas y no crea nada nuevo.

- **AI Moderna:** Utilizan **agentes inteligentes** para resolver problemas reales **en un entorno determinado**.

Ejemplo: El ChatGPT

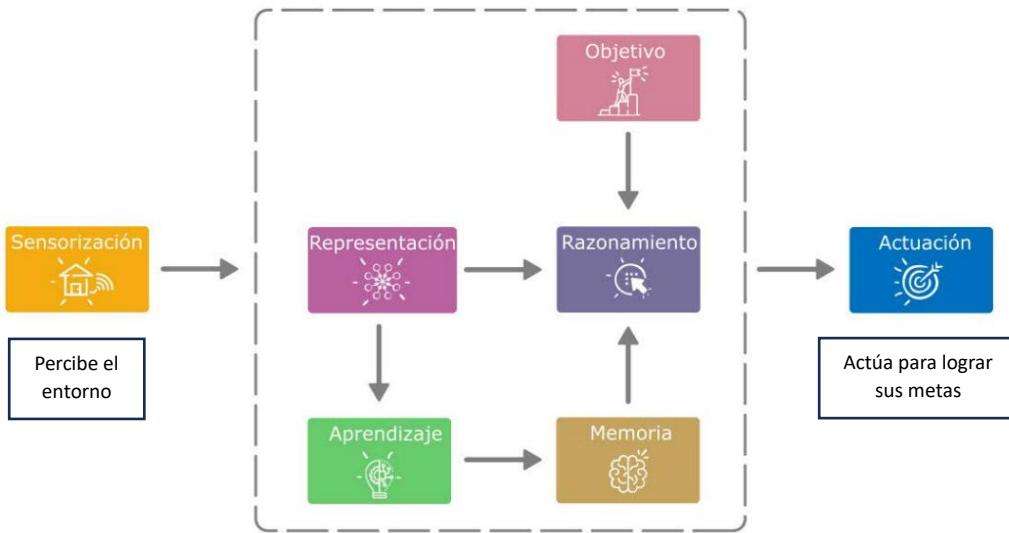


IA Moderna o Generativa

Para definir de qué está compuesto un ecosistema de inteligencia artificial (IA), podemos comenzar considerando el diagrama que se muestra arriba, el cual contiene 4 bloques principales y sus relaciones. En la parte superior podemos ver el **Entorno**, que puede ser real o virtual, y es donde opera el sistema de IA. En la parte derecha del diagrama, podemos ver un bloque de **Detección**, que representa que el sistema de IA tiene sensores que proporcionan información del entorno. Además, el bloque de **Actuación** en la parte izquierda representa que tiene actuadores que le permiten modificar el entorno en cierta medida a través de las acciones aplicadas en él. Finalmente, en la parte inferior del diagrama podemos ver el bloque del **Sistema de IA**, que se dedica a seleccionar, dependiendo de la información proveniente del bloque de Detección o Sensorización, la actuación más apropiada que debe ejecutarse en el entorno para cumplir con los objetivos del sistema.

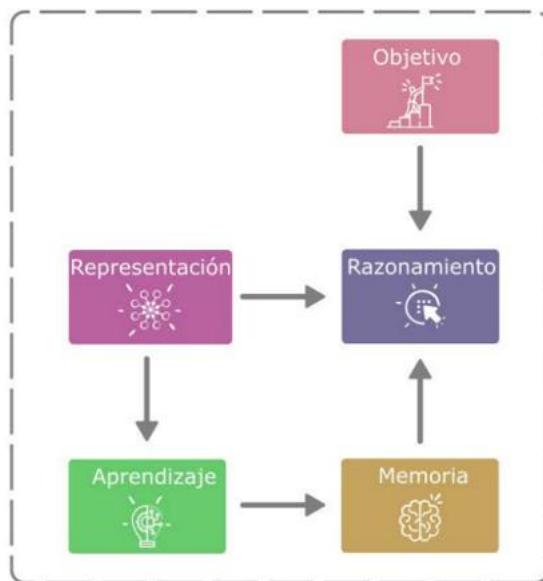
Como veis, es cíclico porque las acciones que ejecuta el sistema de IA **modifican el entorno**, y se deben obtener nuevos valores de detección para percibir tales cambios, proporcionando al sistema de IA nueva información sobre la adecuación de dicha acción para lograr sus objetivos. Este proceso continúa a lo largo del tiempo, por lo que el sistema de IA **puede aprender** de su experiencia para mejorar la selección de acciones.

La complejidad del bloque de **Sistema de IA** (agente inteligente) puede variar ampliamente, estableciendo el nivel de procesamiento de la información que se puede alcanzar y, en consecuencia, limitando de forma más o menos óptima las decisiones que se pueden tomar. El siguiente diagrama contiene un esquema de los elementos básicos que se pueden encontrar en un bloque de Sistema de IA:



Podemos observar que está compuesto por 5 bloques, vinculados a Detección (Sensorización) y Actuación como se explicó anteriormente. Estos 5 bloques forman el **Agente Inteligente** y son los siguientes:

1. **Objetivo:** un sistema de IA debe buscar cumplir algún objetivo. Ese objetivo puede ser impuesto por el diseñador o puede ser aprendido por el sistema, lo que implica un mayor nivel de inteligencia.
2. **Representación:** la información proveniente de los sensores debe almacenarse en el sistema computacional bajo algún tipo de representación computacional, lo cual es muy importante para simplificar los procesos restantes.
3. **Aprendizaje:** todo sistema de IA debe aprender de la experiencia. Esta característica es muy relevante, ya que proporciona la propiedad de poder adaptarse a nuevas situaciones y, en consecuencia, ser realmente autónomo.
4. **Razonamiento:** el sistema debe seleccionar la acción a aplicar en el entorno siguiendo algún tipo de proceso de decisión, desde uno muy simple y reactivo hasta enfoques muy complejos que utilicen modelos internos, planificación u otros.
5. **Memoria:** es obligatorio que un sistema de IA tenga diferentes tipos de elementos de memoria para almacenar los modelos que se aprenden, las representaciones y otros datos importantes que podrían ser útiles en el futuro para evitar tener que volver a aprender.



La inteligencia artificial clásica y la inteligencia artificial generativa o moderna son dos enfoques diferentes de IA y pueden ser utilizados en conjunto para proporcionar soluciones más poderosas. La IA tradicional se centra en realizar tareas específicas *utilizando algoritmos y reglas predefinidas*, mientras que la IA generativa tiene la capacidad de crear contenido nuevo *basado en datos de entrenamiento*.

2.4. ¿Qué no es Inteligencia Artificial?

Si recordáis, la primera actividad de este libro que llamé “El pedazo de papel inteligente”, se basaba en un algoritmo simple que seguía unas reglas para que, siempre que el papel comenzara jugando, nunca perdiera: o empataba, o ganaba.

Todos entendimos que ese papel, que se llamaba inteligente, no era IA. **Los algoritmos simples y la programación tradicional** se basan en gestionar condiciones predefinidas y, aunque puedan parecer inteligentes (porque están muy bien programadas), el que es inteligente es el humano que realizó el programa y no el programa en sí, que no tiene capacidad para aprender ni adaptarse.

Lo mismo ocurre con las **automatizaciones** que, como su nombre indica, se basan en una programación preestablecida y estructurada que consigue que esas máquinas realicen esa acción. Cuando creamos **un robot** seguidor de línea, no estamos implementando IA. No hay nada inteligente en seguir una línea, aunque lo haga un robot. Aunque la IA y la robótica son campos distintos, actualmente están más relacionados y hoy en día pueden trabajar juntas,

complementándose para brindar avances significativos en una amplia gama de industrias, desde la manufactura y la atención médica, hasta la seguridad y la exploración espacial. Por explicarlo de una manera simple, con esta combinación conseguiríamos que: la robótica se encargara de diseñar el **hardware** (la carcasa o lo que podríamos llamar robot) y la inteligencia artificial sería el **software** o programa que permite que este robot actúe de forma inteligente.

Los **métodos estadísticos y el análisis de datos** (que estudiaréis en el bachillerato), tampoco es IA. Sólo son partes de las matemáticas que nos ayudan a identificar patrones o correlaciones entre ese gran volumen de datos.

Internet de las Cosas ([IoT](#)) ha emergido en las últimas décadas como un concepto revolucionario que ha cambiado la manera en que interactuamos con el mundo que nos rodea. A medida que el avance de la tecnología ha ido en aumento, cada vez más dispositivos están conectados a Internet, capaces de recopilar y compartir datos en tiempo real. Tenemos claro que IoT e IA son dos tecnologías diferentes, pero si unimos ambos podemos hablar de [AIoT](#) (*Inteligencia Artificial de las cosas*), que son sistemas inteligentes y conectados que son capaces de tomar decisiones por sí mismos, evaluar los resultados de estas decisiones y mejorar con el tiempo.

3. Fundamentos más relevantes de la IA

3.1. KDD (Knowledge Discovery in Database)

Se denomina KDD al proceso que, utilizando una gran cantidad de datos, consigue obtener nuevos conocimientos a partir de un análisis iterativo de los datos, hasta obtener la información necesaria. Para entenderlo, debemos comenzar sabiendo a qué llamamos datos, información y conocimiento:

- Los **datos** son un conjunto discreto de factores objetivos sobre un hecho real, pero que de forma aislada no contienen información. Es necesario analizarlos en conjunto para así, poder sacar de ellos información.

ase	Length	Height	Width	Weight	Quality
4.7	1.8	1.7	1.7	high	
4.5	1.4	1.8	0.9	high	
4.7	1.8	1.9	1.3	high	
4.5	1.8	1.7	1.3	medium	
4.3	1.6	1.9	1.7	medium	
4.3	1.4	1.7	0.9	low	
4.5	1.6	1.9	0.9	very-low	
4.5	1.4	1.8	1.3	very-low	

- La **información** es capaz de cambiar la forma en que el receptor percibe la realidad. Tiene significado y está organizada para algún propósito. Los datos se convierten en información cuando les añadimos un significado, les añadimos un valor nuevo.
- El **conocimiento** existe dentro de las personas como parte de la complejidad humana y llega cuando somos capaces de relacionar la información que nos rodea, realizando así un proceso de aprendizaje que nos permitirá tomar las decisiones más adecuadas en cada momento.

El proceso KDD, también llamado “**descubrimiento de conocimiento en bases de datos**”, se puede resumir en 4 etapas:

1. Recopilación de datos

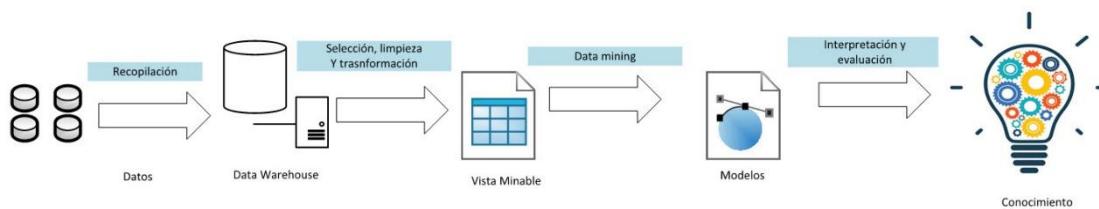
2. Selección, limpieza y transformación de los datos

3. Minería de datos (algoritmos para obtener modelos de aprendizaje)

Los algoritmos más relevantes o utilizados son los siguientes:

- K-nearest neighbors
- Naïve Bayes
- Decision trees
- Support Vector Machine
- K-means
- Linear regression
- Logistic Regression
- **Neural networks**
 - Perceptrons
 - FeedForward
 - RNN/LSTM
 - Convolutional
 - Transformers

4. Interpretación y evaluación de los modelos



Pongamos algún ejemplo:

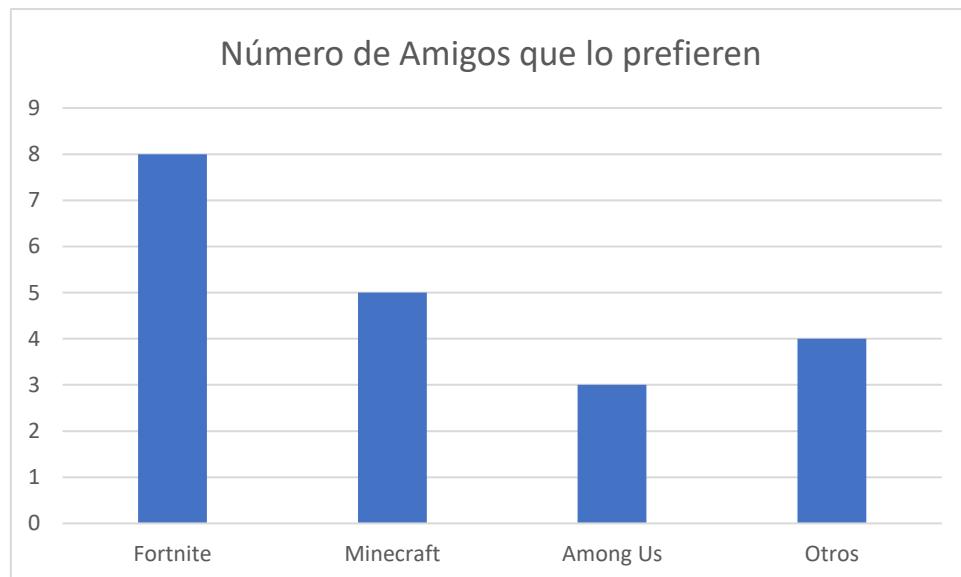
Ejemplo 1: Supongamos que nuestro objetivo es conocer cuál es el videojuego más popular entre las personas de tu edad. Lo correcto sería hacer esta pregunta a miles de usuarios de videojuegos. Para simplificarlo, supongamos que sólo le hacéis esa pregunta a 20 de vuestros amigos y que ellos responden lo siguiente:

Buscamos encontrar los datos relevantes:

Videojuego	Número de Amigos que lo prefieren
Fortnite	8 amigos
Minecraft	5 amigos
Among Us	3 amigos
Otros	4 amigos

A la hora de crear la tabla anterior debemos hacer una limpieza y preprocesamiento de los datos. Es decir, verificamos que los nombres de los juegos estén escritos correctamente y que no haya errores en los datos recopilados.

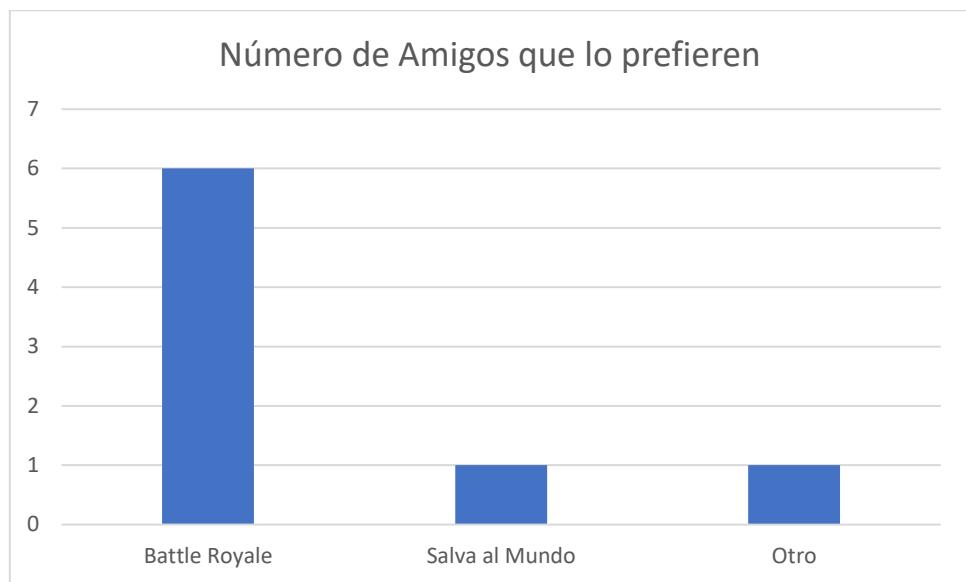
La tabla anterior nos muestra claramente que *Fortnite* es el juego más popular. Pensar que en la vida real, con enormes cantidades de datos, se aplican algoritmos complejos de minería de datos. En un escenario más complejo, podríamos usar algoritmos de clasificación o clustering para identificar patrones en los gustos de juegos.



Pero ¿por qué *Fortnite* es el más popular? Para saber el por qué necesitamos más datos. Realizamos un segundo análisis para obtener más detalles sobre el videojuego "*Fortnite*". Preguntas a los amigos que eligieron "*Fortnite*" sobre cuál es su modo de juego favorito en "*Fortnite*": "*Battle Royale*", "*Salva al Mundo*" u "*Otros modos*". Pensar que el proceso KDD busca obtener "conocimiento" a través de un análisis iterativo de los datos, de ahí que ejemplifiquemos un segundo análisis. Nos encontramos con lo siguiente:

Modo de juego Fortnite	Número de Amigos que lo prefieren
Battle Royale	6 amigos
Salva al Mundo	1 amigo
Otro	1 amigo

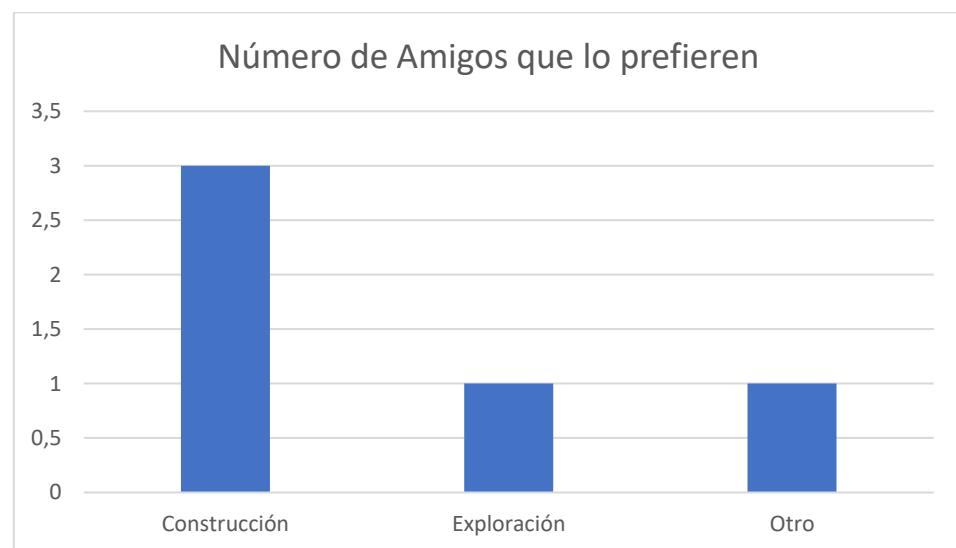
Analizamos los datos y ya sabemos que el más popular es el *Fortnite* y que prefieren jugar al modo de juego *Battle Royale*.



Vamos a hacer un tercer análisis (más iteración buscando obtener más conocimiento) preguntando a los que les gusta Minecraft qué aspecto del juego les gusta más: "Construcción", "Exploración" u "Otros aspectos". Nos encontramos con las siguientes respuestas:

Aspecto del Juego	Número de Amigos que lo Prefieren
Construcción	3 amigos
Exploración	1 amigo
Otro	1 amigo

Con este segundo y tercer análisis estamos interpretando y evaluando mejor los datos, obteniendo más conocimiento.



Ejemplo 2: Supongamos que queremos realizar una inversión de dinero en la Bolsa de valores y comprar acciones de alguna compañía. Partimos de los datos del precio de esas acciones:

Bankinter	6,216€
BBVA	4,764€
Caixabank	2,798€
Cellnex	42,630€
CIE Automotive	21,100€
Enagas	24,000€
Ence Energía	3,946€
Endesa	24,640€
Ferrovial	29,070€
Grifols-A	31,910€

Estos datos (imagen anterior) no nos sirven para nada, no sabemos si son altos o bajos ni que probabilidades tienen de subir o bajar en el futuro. Necesitamos más datos para la fase de “Recopilación”:

Fecha	Último	Apertura	Máximo	Mínimo	Vol.	% var.
19.01.2020	9.656,80	9.674,00	9.679,00	9.627,00	-	-0,23%
12.01.2020	9.681,30	9.506,60	9.709,00	9.466,00	723,000	1,12%
05.01.2020	9.573,80	9.585,40	9.657,00	9.492,70	655,000	-0,78%
29.12.2019	9.616,80	9.672,50	9.705,40	9.518,00	435,000	-0,66%
22.12.2019	9.700,50	9.656,20	9.703,50	9.607,80	257,500	0,28%
15.12.2019	9.676,50	9.607,00	9.662,40	9.667,40	1,000	1,17%
08.12.2019	9.563,70	9.378,70	9.650,80	9.245,10	864,000	1,83%
01.12.2019	9.382,70	9.367,20	9.395,40	9.119,10	781,720	0,33%
24.11.2019	9.352,00	9.300,00	9.404,70	9.288,20	726,110	1,05%
17.11.2019	9.254,70	9.265,10	9.327,00	9.152,10	712,000	-0,07%
10.11.2019	9.251,40	9.370,30	9.436,50	9.132,80	897,200	-1,43%
03.11.2019	9.393,70	9.376,50	9.482,50	9.254,30	944,000	0,70%
27.10.2019	9.328,00	9.432,30	9.455,50	9.208,90	859,300	-1,08%
29.10.2019	9.430,20	9.366,30	9.448,00	9.332,40	735,000	1,08%
13.10.2019	9.339,80	9.229,70	9.494,50	9.195,20	887,200	0,60%
06.10.2019	9.273,80	9.064,70	9.273,80	8.903,80	779,000	3,48%
29.09.2019	9.961,80	9.194,80	9.298,50	8.850,80	824,000	-2,42%
22.09.2019	9.184,10	9.154,80	9.210,80	8.803,80	705,000	0,00%
15.09.2019	9.179,00	9.099,50	9.210,90	8.953,80	1,150	0,45%
08.09.2019	9.137,90	8.987,70	9.183,50	8.947,70	1,110	1,64%
01.09.2019	9.990,10	8.810,80	9.025,80	8.772,80	810,000	2,01%
25.08.2019	8.812,90	8.603,30	8.865,50	8.692,20	640,000	1,09%
18.08.2019	8.849,50	8.726,90	8.805,40	8.618,30	675,000	-0,24%
11.08.2019	8.670,40	8.796,90	8.811,50	8.408,10	1,030	-1,08%
04.08.2019	8.757,80	8.824,00	8.869,00	8.679,80	865,200	-1,57%
28.07.2019	8.897,80	9.281,30	9.296,10	8.870,80	934,000	-3,55%

Esta tabla de la imagen anterior dispone de más valores. Pero esos datos tampoco nos permiten obtener una información de un modo directo y sencillo. Si, es difícil de extraer algo de ella. Debemos analizarlos “Selección de los mismos, limpieza (eliminar datos erróneos) y transformación de los datos”. Si tratamos estos datos y obtenemos una gráfica podremos

hacernos una idea de la evolución del precio de las acciones en un período de tiempo más amplio. Ya tenemos alguna información útil:

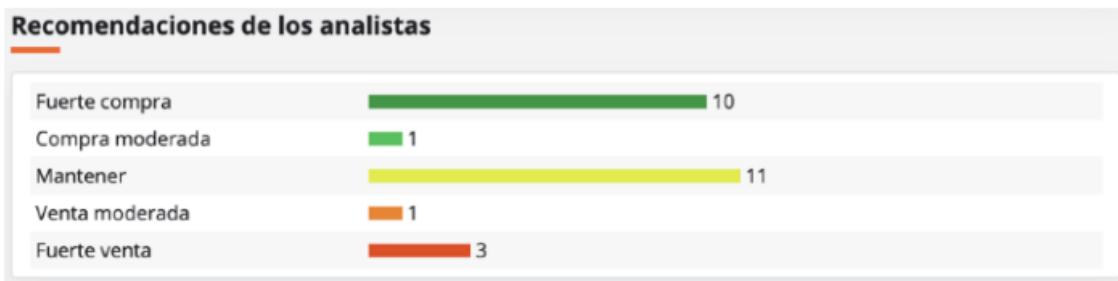


Si ahora analizamos más datos y obtenemos nuevos gráficos, “*data mining*”, podemos llegar a tener un conocimiento del mercado y qué valores merece la pena comprar o si, por ejemplo, es un buen momento para hacerlo:



Sería el momento en que nos podemos considerar expertos o conoedores profundos del funcionamiento de este mercado. Es la fase de “ Interpretación y evaluación de los modelos”. Y

en ese momento es cuando llegamos al conocimiento. En el siguiente ejemplo vemos las recomendaciones de diferentes analistas sobre qué hacer con un determinado tipo de acciones:

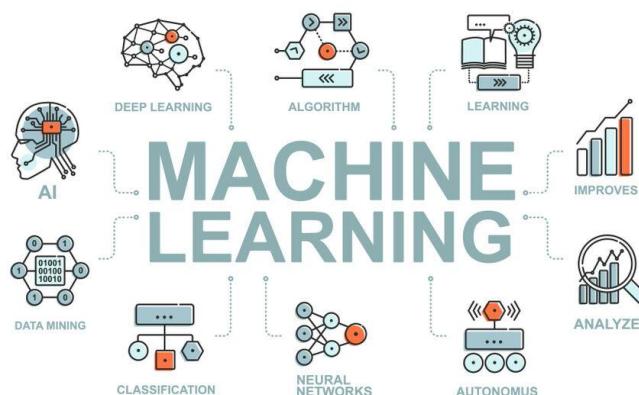


Estos analistas después de obtener muchos datos y analizar muchos gráficos llegan a ser conocedores del mercado y son capaces de hacer previsiones de la evolución futura de las acciones y por lo tanto, tomar la decisión correcta a la hora de comprar o vender acciones. Tienen en su cabeza un modelo del mercado y de su funcionamiento, que les permite entenderlo y hacer una predicción acertada.

El reto de la AI consiste en crear un programa informático que repita el mismo esquema de funcionamiento y sea capaz de aprender y hacer previsiones acertadas en el futuro, en cualquier campo de conocimiento, no solo en la [bolsa](#).

3.2. Machine learning (Aprendizaje automático)

El **aprendizaje automático** (del inglés, *machine learning*) es el subcampo de las ciencias de la computación y a la vez una rama de la inteligencia artificial, cuyo objetivo es desarrollar técnicas que permitan que las computadoras aprendan. Se centra en el desarrollo de algoritmos y modelos que permiten a las computadoras aprender automáticamente a partir de los datos y mejorar su rendimiento en tareas específicas, sin ser programadas explícitamente.

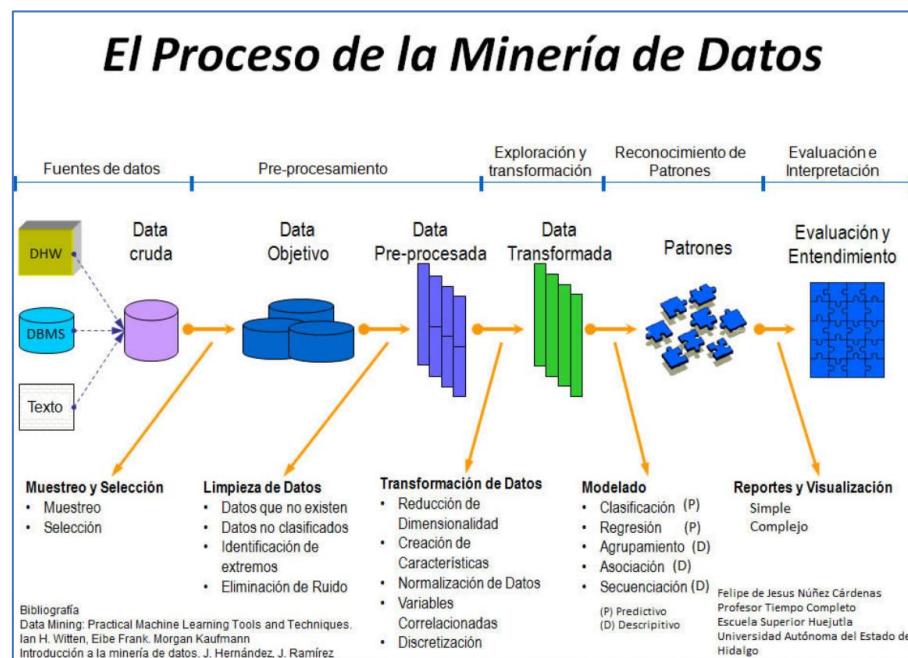


Machine learning:

Machine learning (ML) es una forma de construir sistemas de inteligencia artificial utilizando la estadística, en lugar de escribir las reglas en un programa. Los sistemas de aprendizaje automático se denominan "basados en datos", porque utilizan muchos datos para funcionar.

Los ordenadores son capaces de analizar cantidades masivas de datos (que se conoce con el término *big data*) y de generar modelos a partir de ellos. Buscan relaciones entre los datos e incluso, descartan aquellos que no nos son útiles. A este proceso se le denomina *minería de datos*. Cuando usamos los ordenadores para este fin estamos llegando al conocimiento, el ordenador lo ejecuta de una forma rápida y sencilla. Obviamente, para un humano sería inviable hacerlo directamente, debido a la enorme cantidad de datos a analizar.

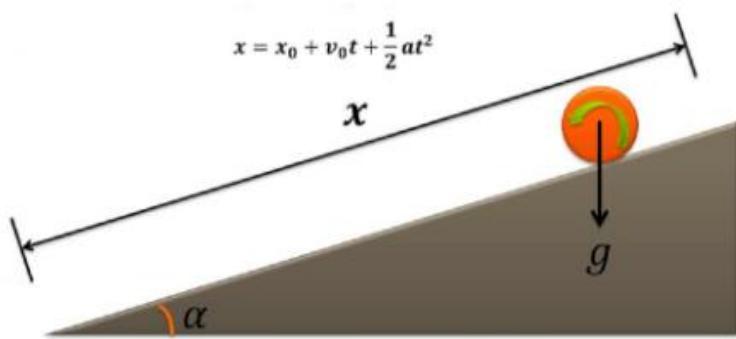
La **minería de datos** o exploración de datos (etapa del "Knowledge Discovery in Databases" o KDD) es un campo de la estadística y las ciencias de la computación asociado al proceso que intenta descubrir patrones o modelos en grandes volúmenes de conjuntos de datos. El objetivo general del proceso de minería de datos consiste en extraer información de un conjunto de datos y transformarla en una estructura comprensible, para su uso posterior.



Lo que realmente hacemos con el *machine learning* es crear modelos o patrones que nos ayudan a simplificar la realidad y que nos permiten hacer predicciones de cierta fiabilidad. ML es una técnica específica utilizada dentro del proceso de KDD para desarrollar modelos.

La ciencia lleva mucho tiempo haciéndolo y esto nos ha permitido convertir todo *el ruido del mundo real* en conocimiento. ¿Cómo? Reconstruyendo la realidad a través de modelos.

Un modelo no es más que una construcción conceptual simplificada de una realidad más compleja. A través de esta reconstrucción somos capaces de entender mejor esta realidad. Convivimos diariamente con multitud de modelos, como pueden ser, un mapa, la escritura o también una ecuación física. Por ejemplo, la ecuación que nos predice la distancia recorrida por un objeto cuando desciende por un plano inclinado:



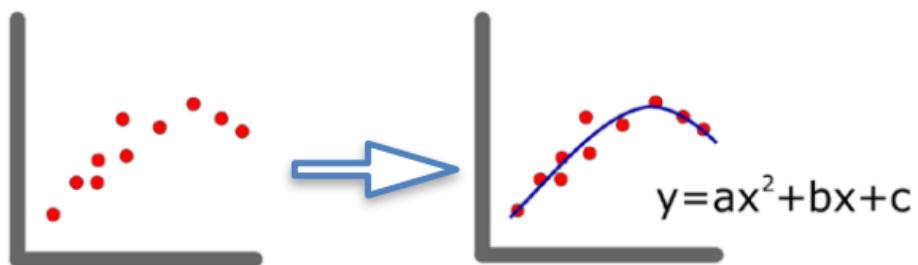
Esta fórmula, a partir de ciertas variables, logra estimar el comportamiento del objeto y aproximar lo que ocurre en la realidad. Pero es una simplificación, lo que realmente estamos calculando es la probabilidad de que el objeto haya recorrido determinada distancia al pasar cierto tiempo. Es decir, tenemos es un **modelo probabilístico**.

Nuestro cerebro aplica esquemas similares a estos modelos probabilísticos y gracias a ellos tenemos capacidades como la de conceptualizar, predecir, generalizar, razonar o aprender. Por eso somos capaces de conducir un coche por una carretera desconocida o diferenciar un perro o un gato sin tener que conocer todas las razas de estos animales. Lo que hacemos en el *machine learning* es **trasladar este método de aprendizaje a un programa** que sea capaz de realizar este proceso, y tenemos 3 elementos clave para conseguirlo, que son: los datos, los parámetros y el error.

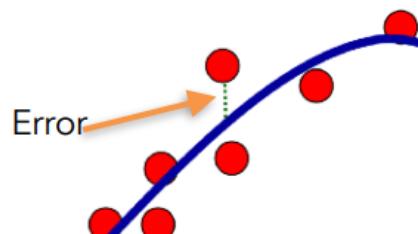
Para ejemplificarlo vamos a simplificar esa enorme cantidad de datos en una tabla muy, muy pero que muy reducida (Ojo, hacemos esto para ejemplificar el proceso):

Datos de entrada	Salida
1	2
2	4
3	6
4	8
10	¿?

- a. Los **datos**, que son nuestra toma de contacto con la realidad, no son más que las mediciones que hacemos de ella, y a partir de estos datos vamos a construir nuestro modelo. Cuantos más datos tengamos, mejor (*Big Data*).
- b. Los **parámetros** se calculan matemáticamente y se asocian a funciones matemáticas. Estos parámetros se ajustan en nuestro modelo que nos devuelve una función que se approxima a la realidad. Por ejemplo, si queremos buscar un modelo que nos separe unos datos de otros, buscaremos una función matemática que se ajuste lo más correctamente posible: Los parámetros serán los valores del polinomio que mejor se ajuste a esos datos.



- c. El **error** es la diferencia entre el modelo y la realidad. Siempre tendremos que crear una función de error que nos diga con qué fiabilidad nuestro modelo se ajusta a la realidad.



3.3. Paradigmas de aprendizaje

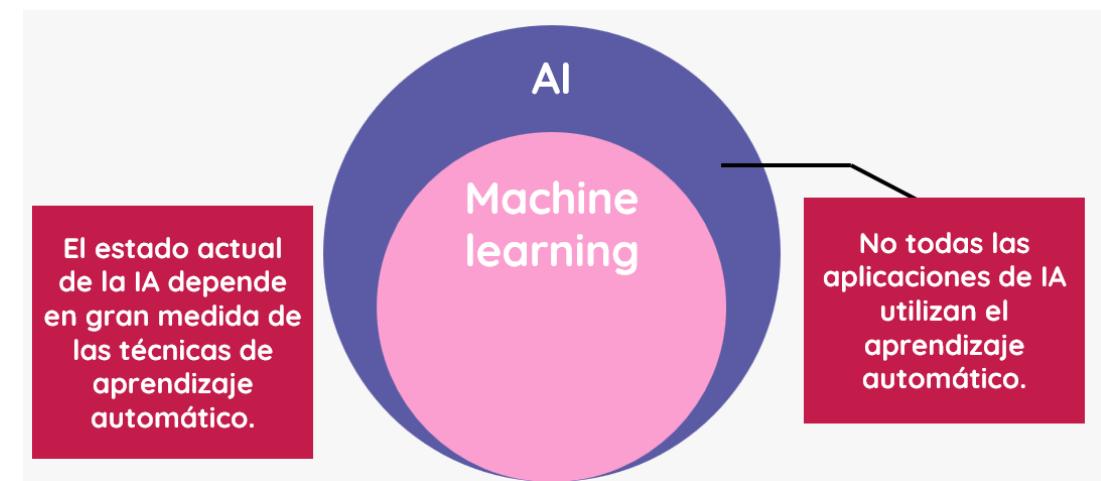
Para aprender, los humanos nos valemos de varias estrategias. Una de ellas es la que usamos cuando aprendemos a jugar al ajedrez o a un videojuego. Conociendo las reglas básicas, vamos probando, en principio un poco aleatoriamente, y memorizamos aquellas estrategias que mejor resultado nos dan y las usamos con mayor frecuencia, y así, poco a poco, mejoramos nuestro nivel.

Otra forma sería la que usamos cuando tenemos que clasificar cosas. Por ejemplo, cuando organizamos nuestra biblioteca de libros. No tenemos una única solución, pero somos capaces de agrupar los libros usando características comunes a algunos de ellos: temática o por tamaños o por autor o por orden alfabético... con diferentes estrategias obtenemos diferentes soluciones, todas igual de válidas. Cuando nos compramos un nuevo libro sabremos a qué grupo debemos añadirlo. Esto lo usamos inconscientemente al ser capaces de reconocer diferentes escrituras, tanto tipográficas como manuscritas, cuando vemos una letra la añadimos automáticamente a uno de nuestros grupos mentales de letras.

Una tercera estrategia es la que conocemos como **método científico** y es la menos intuitiva de todas. Primero tenemos que diseñarlo correctamente y al realizarlo mediremos los resultados jugando con el valor de las variables de entrada. La dificultad reside en encontrar la relación entre esas variables y el resultado obtenido, y para eso nos ayudamos de las matemáticas y sus diferentes herramientas. Eso es lo que hizo Georg Simon Ohm cuando midió la intensidad eléctrica de un circuito y lo relacionó con la resistencia usada y la tensión aplicada. Encontró la relación matemática entre estos tres factores y es lo que hoy conocemos como la Ley de Ohm.

Estas tres formas que tenemos de aprender son las que han trasladado a los sistemas informáticos y han dado lugar a los 3 [paradigmas de aprendizaje](#) que se usan en *machine learning* y son: el **aprendizaje supervisado**, el **aprendizaje no supervisado** y el **aprendizaje reforzado**.





Cuando aplicamos cualquiera de los métodos de aprendizaje debemos dividir ese gran volumen de datos en dos partes: **los datos de entrenamiento y los datos de test**. Primero entrenaremos el algoritmo usando exclusivamente los datos de entrenamiento. Tras hacerlo, obtendremos un modelo o una regla que predice el resultado a partir de las variables de entrada. Pero, para determinar el grado de acierto con el que realmente podremos predecir los resultados, no basta con los datos de entrenamiento. Que un modelo pueda hacer muy buenas predicciones con los datos de entrenamiento no demuestra que pueda generalizarse a cualquier otro dato. Aquí es cuando los datos de test resultan útiles: podemos usar el modelo entrenado para predecir los resultados correspondientes a los datos de test y comparar las predicciones con los resultados reales.

En la siguiente imagen se muestran los tipos de Machine Learning:



Son varias las plataformas con las que podemos realizar prácticas de ML Hoy en día podemos trabajar con las siguientes: [Plataformas de Machine Learning](#).

3.3.1. Aprendizaje supervisado

Cuando hablamos de aprendizaje supervisado nos referimos a descubrir las relaciones entre unas variables de entrada y unas de salida. Es decir, el aprendizaje surge de enseñarles a estos algoritmos cual es el resultado que quieras obtener para un determinado valor de entrada, tras mostrarle muchos ejemplos. Por lo tanto, dispone de las respuestas correctas y la tarea del algoritmo de aprendizaje automático es hallar un modelo que las prediga en función de los datos de entrada. El algoritmo será capaz de dar un resultado correcto, incluso cuando le enseñas valores que nunca había visto antes. Este sistema le hubiese ahorrado mucho trabajo a Ohm.

Importante

Tipos de aprendizaje automático

El entrenamiento requiere datos de **ejemplo**

Aprendizaje supervisado	Aprendizaje no supervisado	Aprendizaje reforzado
1	2	3

Importante

Tipos de aprendizaje automático

Requiere muchos **datos de entrenamiento**

Aprendizaje supervisado	Aprendizaje no supervisado	Aprendizaje reforzado
1	2	3

Por ejemplo, vista la siguiente tabla de valores, ¿qué valor tendremos en la salida si el valor de entrada es 10?

Datos de entrada	Salida
1	2
2	4
3	6
4	8
10	¿?

A partir de los datos de entrada somos capaces de generalizar y entender que, multiplicando por 2 la entrada tenemos la salida, ya tenemos nuestro modelo. Esta es la base del aprendizaje supervisado.

En este ejemplo no es necesario ningún algoritmo ni ordenador para entender la relación entre la entrada y la salida, pero cuando los datos de entrada son múltiples o difíciles de clasificar (estamos hablando de 2000 datos), la cosa se complica y es donde el *machine learning* es de gran utilidad.

Por ejemplo, el sistema de clasificación de los correos electrónicos como spam se basa en el análisis de millones de correos por parte de los usuarios que luego se usan para entrenar un algoritmo que consiga clasificarlos automáticamente. La clave aquí es que **no hay una norma o normas** que nos digan si un correo es spam o no. Es el algoritmo el que crea un modelo propio para clasificarlos.

Por lo tanto, se puede decir que el objetivo del aprendizaje supervisado es el **de crear una función capaz de predecir el valor** correspondiente a cualquier objeto de entrada válido después de haber visto una serie de ejemplos.

Cuando la función nos devuelve un valor numérico nos encontramos ante un problema de regresión, y cuando lo que buscamos es una marca de clase, el problema al que nos enfrentamos es de clasificación.

De este modo, se puede decir que normalmente, en aprendizaje supervisado, tenemos dos posibilidades:

- Regresión:** Predice un dato numérico de salida a partir de las variables de entrada.
- Clasificación:** tenemos datos de varias clases y queremos aprender a clasificar nuevos datos en esas clases de forma automática.

En los proyectos de *Machine Learning* quizás necesitemos **convertir texto a audio o texto a texto en otro idioma**, como las herramientas de traducción automática.

Un conversor de texto a audio con inteligencia artificial es una herramienta muy útil para los estudiantes, ya que les permite convertir textos en archivos de audio para poder escucharlos. La **inteligencia artificial** lee el texto y lo convierte en un archivo de audio de alta calidad.

A continuación os paso algunos enlaces a algunas **herramientas online** relacionadas con esta conversión. Con ellas podéis explorar esta forma de utilizar la inteligencia artificial.

Actividad 7: Selecciona un texto o créalo tú mismo y pruébalo en las distintas páginas que tienes a continuación, todas son gratuitas y te permiten pasar de texto a voz y de voz a texto.

- [Textfromtospeech](#)
- [Readspeaker](#)
- [Ttsreader](#)
- [Texttospeechrobot](#)

Actividad 8 y 9: Realiza las siguientes prácticas de *Machine learning Supervisado*

- Práctica 1: ML Supervisado con datos numéricos: [Videotutorial1](#)
- Práctica 2: ML Supervisado con imágenes: [Videotutorial2](#) (dispones de la base de datos de las imágenes en el entorno EVA o en el Aula Virtual)

Actividad 10: Responde a estas preguntas:

- a. ¿Qué base de datos has creado para la práctica 1?
- b. ¿Por qué se denomina supervisado el aprendizaje de las prácticas anteriores?
- c. ¿Por qué se dividen los datos entre entrenamiento y testeo?
- d. ¿Qué parte es más importante de las tres siguientes: datos de entrenamiento, aprendizaje o evaluación?

Actividad 11: Crea tu propio proyecto de ML supervisado.

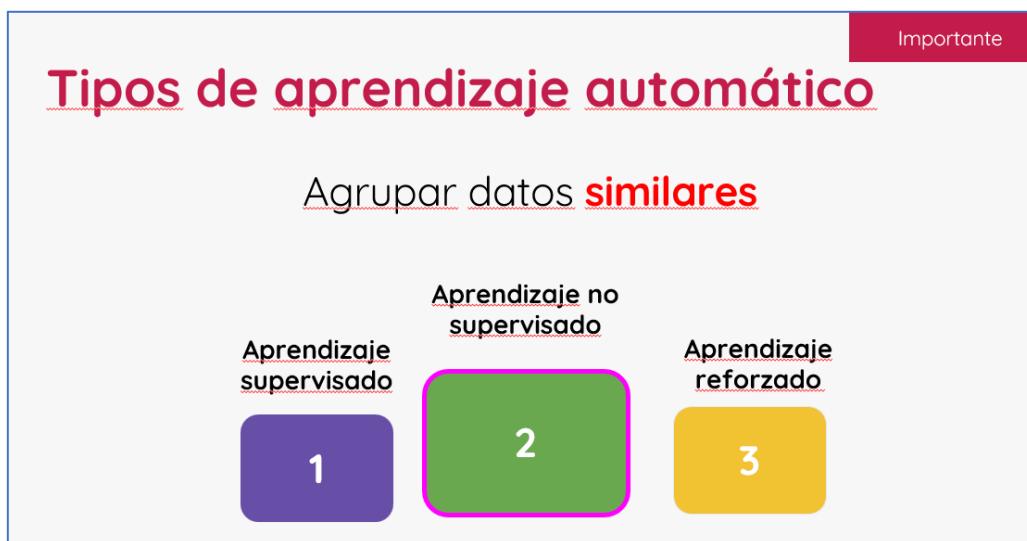
IMPORTANTE: Puedes descargar bases de datos de la web [Kaggle](#)

Actividad 12: Crea tu primer proyecto de ML con audios o sonidos. Utiliza [Teachable Machine](#) para guardar tu base de datos de audios y el clon de scratch llamado [Stretch3](#) para programarlo.

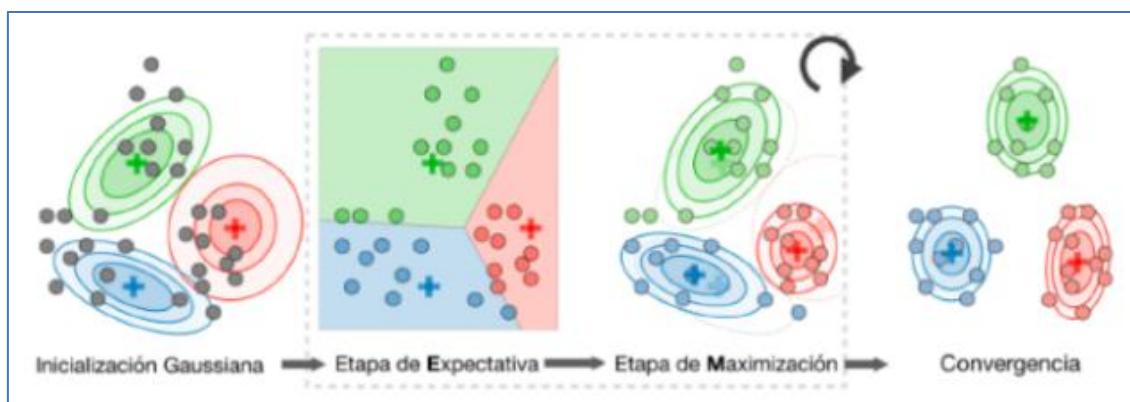
(En el Aula Virtual he subido un documento de guía)

3.3.2. Aprendizaje no supervisado

El paradigma del aprendizaje no supervisado es aquel en que el algoritmo se crea simplemente a partir de los datos de entrada, pero **no hay datos de salida**. No le indicamos al sistema los resultados que queremos obtener porque no se cuenta con las respuestas correctas, y eso cambia bastante las cosas, ya que no podemos desarrollar el modelo adaptándolo a las respuestas correctas según los datos de entrenamiento. En este tipo de aprendizaje, al analizar los datos de entrada se busca encontrar relaciones entre ellos y, de esta forma, el algoritmo los clasificará en varios grupos, igual que nosotros cuando organizamos nuestra biblioteca.



Por ejemplo, si queremos hacer una clasificación de jugadores de baloncesto. Partimos de los datos que los definen como jugadores: altura, peso, porcentaje de tiro, balones perdidos... y todo tipo de estadística de ese deporte. Podemos entrenar a un algoritmo de aprendizaje no supervisado para que los agrupe en función de sus características. De este modo tendremos varios grupos de jugadores similares que podremos relacionar con los que juegan como bases, aleros, escoltas o pívots.

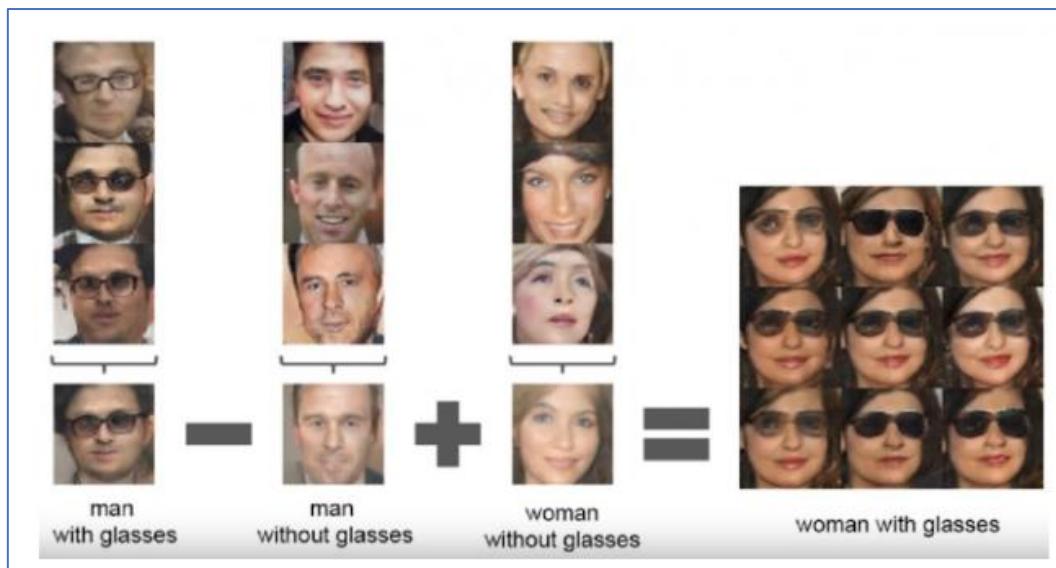


Lo que hace el aprendizaje no supervisado es **buscar patrones de similitud** de los valores de entrada. Un dato muy relevante es que los algoritmos más potentes de este paradigma son capaces de descubrir *cual es la estructura interna que han generado dichos datos*. Se podría decir que son capaces de entender estructuras mentales, denominadas espacios latentes, que le permiten saber, por ejemplo, lo que es un 2 o un 3 (escritos manualmente), pese a que cada persona pueda escribir los números de un modo diferente.

0 0 0 0 0 0 0 0 0 0 0 0 0 0
 1 1 1 1 1 1 1 1 1 1 1 1 1 1
 2 2 2 2 2 2 2 2 2 2 2 2 2 2
 3 3 3 3 3 3 3 3 3 3 3 3 3 3
 4 4 4 4 4 4 4 4 4 4 4 4 4 4
 5 5 5 5 5 5 5 5 5 5 5 5 5 5
 6 6 6 6 6 6 6 6 6 6 6 6 6 6
 7 7 7 7 7 7 7 7 7 7 7 7 7 7
 8 8 8 8 8 8 8 8 8 8 8 8 8 8
 9 9 9 9 9 9 9 9 9 9 9 9 9 9

[Demostración Espacio latente letras](#)

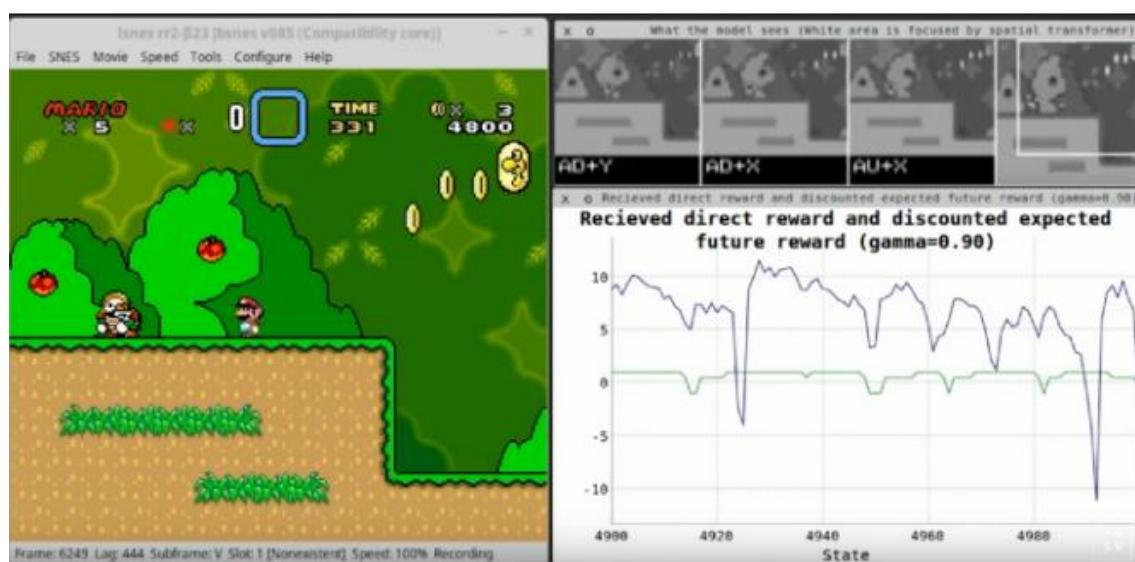
Estos algoritmos incluso están llegando a la posibilidad de operar matemáticamente con estos conceptos abstractos:



Si a la imagen de “hombres con gafas” le restamos “las caras” y le sumamos “caras de mujeres”, nos da como resultado “mujeres con gafas”.

3.3.3. Aprendizaje reforzado

En este caso, la salida que queremos que prediga nuestro algoritmo es una secuencia de acciones que serán ejecutadas dentro de un entorno de simulación que tendrán que realizar una determinada tarea. El algoritmo será recompensado cada vez que logra su objetivo o penalizado si no lo consigue.



IA entrenada para jugar a Mario Bros

El mayor campo de ensayo de estos algoritmos es el de los videojuegos, ya que cumplen todas las premisas que nos permiten entrenar a nuestras IA. Ya tenemos sistemas que juegan al ajedrez, Go, [Starcraft 2...](#)

El [aprendizaje reforzado](#) usa el mundo de los videojuegos como entrenamiento. Pero este no es su objetivo, es en el campo de la robótica donde las aplicaciones de este paradigma de aprendizaje están empezando a dar sus primeros frutos: que un brazo robótico pueda aprender a hacer los movimientos necesarios para realizar una determinada tarea.

Este aprendizaje necesita de un **agente** que lo colocamos en un entorno o ambiente en donde queramos que aprenda a hacer algo. Tenemos que poner reglas para que sepa qué está bien y qué está mal, de modo que, por cada acción que toma el agente, le decimos si su decisión fue correcta o incorrecta; lo retroalimentamos de forma positiva (dándole puntos) o negativa. El algoritmo que usa el agente buscará ajustarse hacia las acciones que le den puntos.

Importante

Tipos de aprendizaje automático

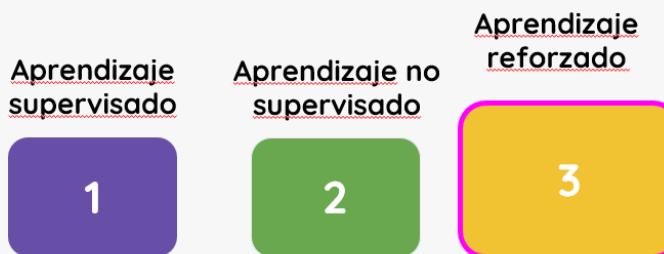
Utiliza **un agente** que tiene acceso directo al entorno



Importante

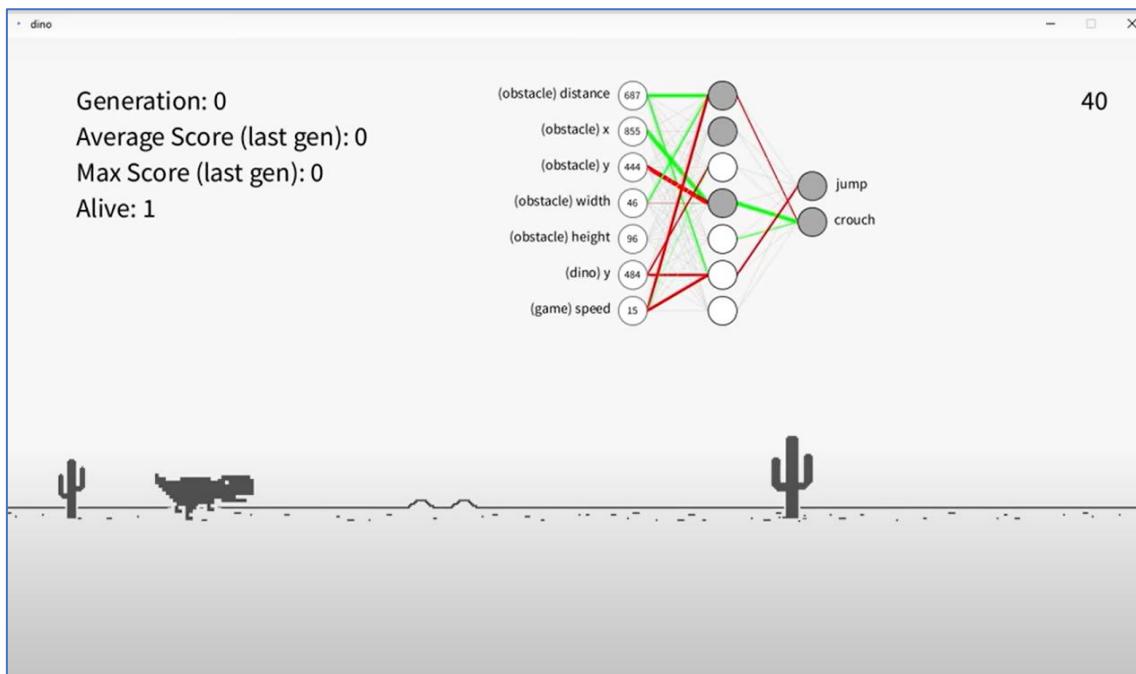
Tipos de aprendizaje automático

Aprender por **ensayo y error**



Si unimos el aprendizaje reforzado con las redes neuronales (deep learning), nos encontramos con el Deep Reinforcement Learning o Aprendizaje reforzado profundo, que se dio a conocer en febrero de 2015 cuando la empresa [Deep Mind](#) publicó un artículo sobre su algoritmo Deep Q-Network (DQN) y donde explicaron que habían conseguido unos resultados increíbles superando el nivel humano en numerosos videojuegos.

En el siguiente [vídeo](#) se muestra, de forma resumida, como se puede implementar una red neuronal que aprenda a jugar al juego del dinosaurio de Google Chrome. El vídeo parte de cero, es decir, crea el juego programándolo con Processing, que es un entorno gráfico de programación basado en Java.



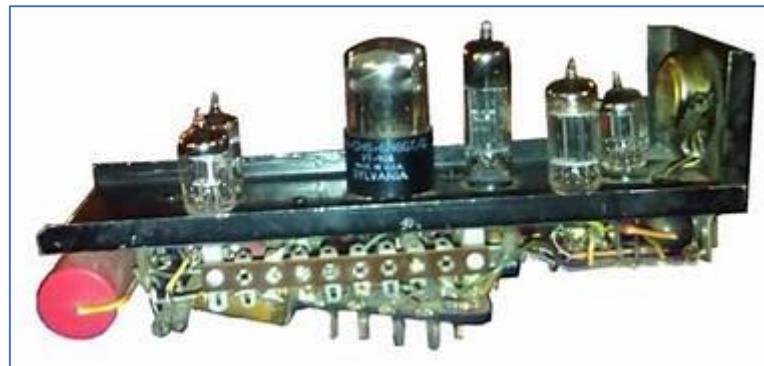
Simulación de “Aprende a jugar a Dino”

3.3.4. Las Redes Neuronales Artificiales (RNA) y el Deep Learning

Estos tres paradigmas de aprendizaje están limitados por la capacidad de los sistemas informáticos de trabajar con una cantidad de datos determinada. Pero creando nuevos modelos computacionales seremos capaces de aumentar esa capacidad. Así nacen las **Redes Neuronales Artificiales (RNA)** y como consecuencia de las mismas, el **Deep Learning** o aprendizaje profundo, que es clave de los actuales y espectaculares avances.

Los paradigmas de aprendizaje supervisado, no supervisado y reforzado pueden ser mucho más potentes si les **añadimos** el modelo computacional de las Redes Neuronales Artificiales (RNA), que es un modelo inspirado en el comportamiento de las neuronas biológicas naturales. Aunque esta idea ya se barajaba desde 1943, no ha sido hasta estos últimos años y gracias al uso de ordenadores con procesadores mucho más potentes, cuando se ha generalizado su uso.

En 1943 Walter Pitts y Warren McCulloch publicaban un influyente estudio en el que describían la neurona artificial, la primera formulación teórica de lo que después se llamaría red neuronal. Sobre este modelo matemático, en 1951 Marvin Minsky y Dean Edmonds diseñarían SNARC, la primera máquina basada en una red neuronal.



SNARC

Una RNA consiste en un conjunto de unidades, llamadas **neuronas artificiales** o **perceptrones**, conectadas entre sí para transmitirse señales. La información de entrada atraviesa la red neuronal, produciendo unos valores de salida. Estas redes neuronales son capaces de aprender de forma jerarquizada. Es decir, la información **se aprende por niveles**. En las primeras capas se aprenden conceptos muy básicos (sus neuronas obtienen sus entradas directamente de los datos) y las siguientes se van especializando en función de las capas anteriores. Esto hace que, cuantas más capas añadamos, mayor es la abstracción que consigue el algoritmo y por lo tanto, pueden tener un mayor conocimiento. Y es lo que se denomina un **aprendizaje profundo** o **deep learning**.

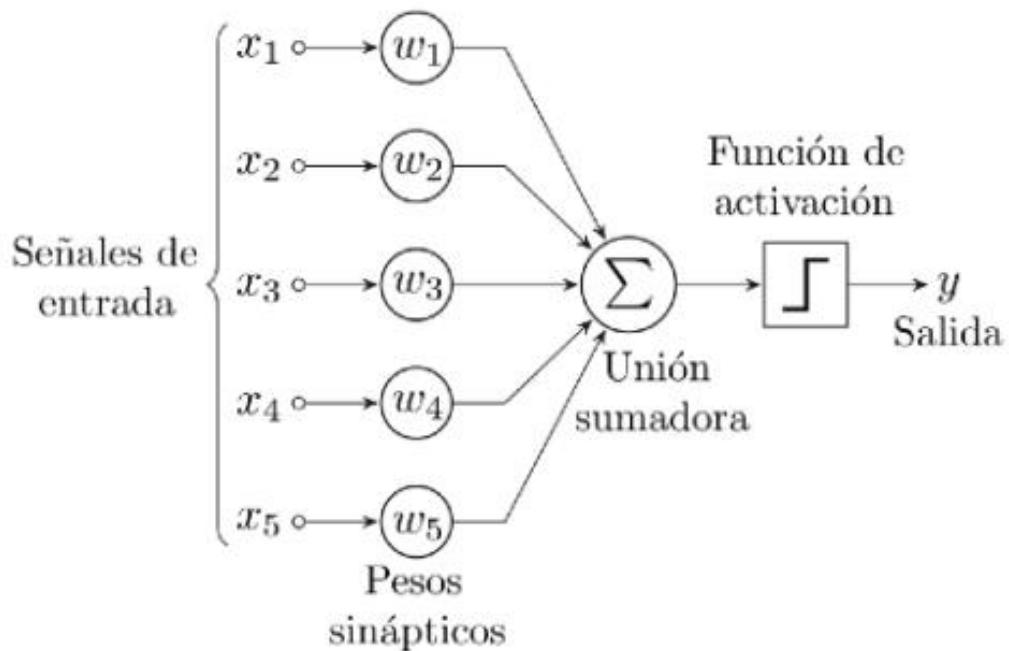
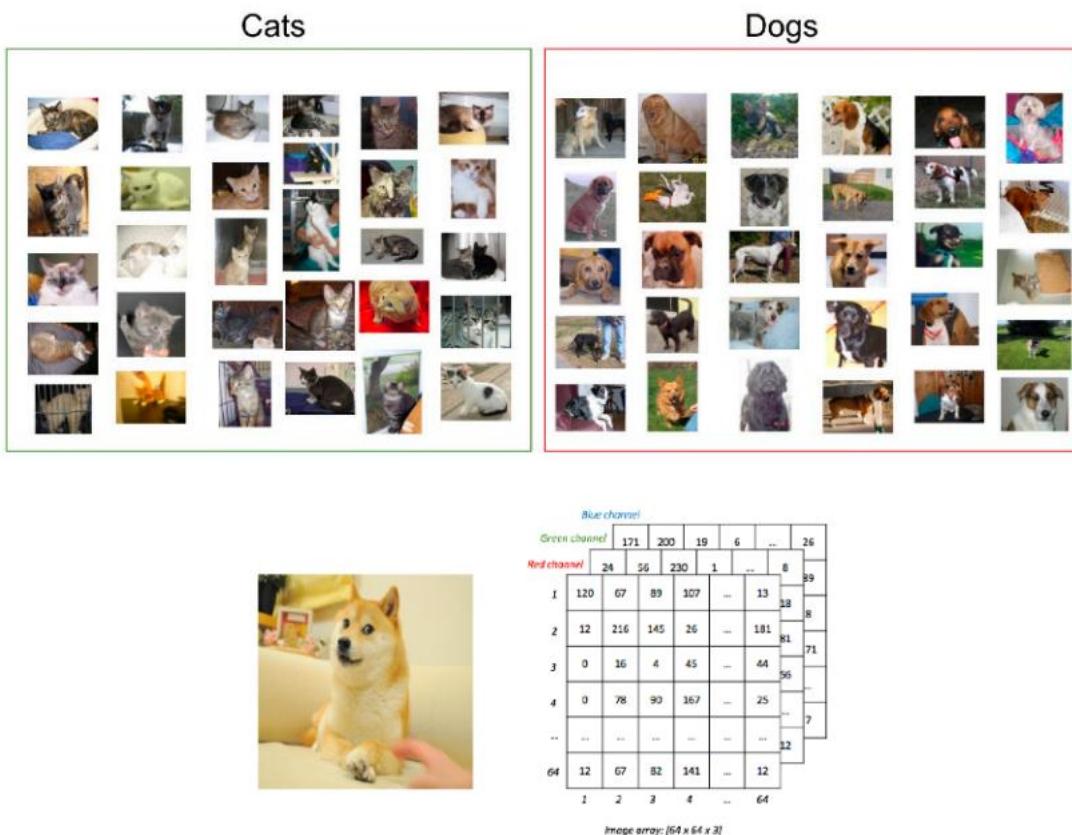


Diagrama de un perceptrón o neurona con cinco señales de entrada.

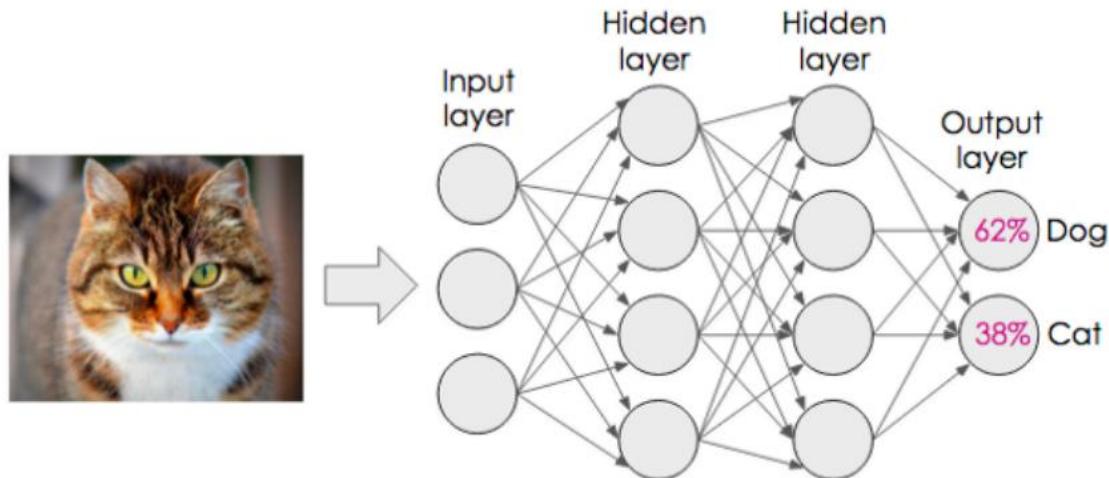
Las redes neuronales se utilizan para resolver una amplia variedad de tareas. Por ejemplo: la visión por ordenador, el reconocimiento de voz, la predicción bursátil, la conducción autónoma, la generación de texto, la traducción de idiomas, el análisis genético o el pronóstico de enfermedades. Todas, difíciles de solucionar sin este tipo de tecnología.

Vamos a explicarlo mejor tomando como ejemplo el reconocimiento de imágenes. Veámoslo.

Supongamos que partimos de una cantidad de fotos etiquetadas de perros y de gatos. Esas imágenes podemos enseñárselas a nuestra red neuronal para que sea capaz de identificar a un perro o a un gato en una imagen nueva, imagen que nunca antes hubiese visto. El sistema dividirá cada imagen en una matriz de pixeles y analizará el color de cada pixel, así como, comprobará la etiqueta de esa imagen para ver si es un perro o un gato.



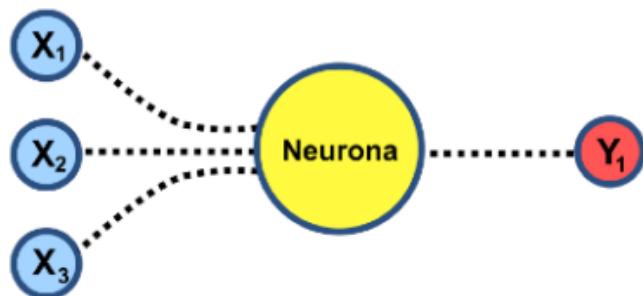
En la imagen de arriba observamos que ha dividido cada imagen en 64x64 píxeles y al color de cada pixel le asigna un valor entre 0 y 255 analizando el color rojo, verde y azul. Este análisis se hará en diferentes niveles o capas, donde cada una se especializará en resolver la tarea a partir de la información enviada por la capa anterior.



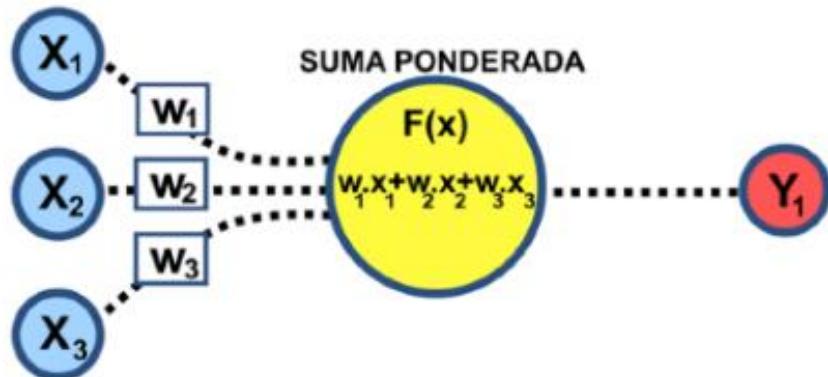
Cuanta más capas tengamos, mejor será nuestro algoritmo y mejores predicciones hará. Si le preguntásemos al terminar este proceso a nuestra red neuronal ¿qué es un gato?, nos contestaría con algo así (ver imagen siguiente): Sería la abstracción creada por la red neuronal de lo que para ella es un gato:



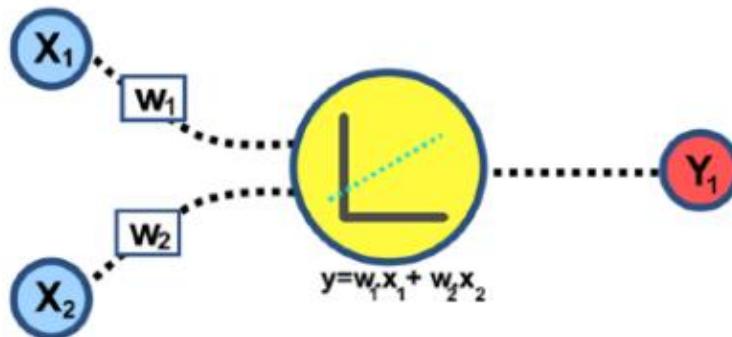
El punto de partida de estas redes neuronales es una **neurona o perceptrón**: una [neurona artificial](#). Esta neurona se forma a partir de unos datos de entrada (X) y nos dará una salida (Y). Realmente, aunque le llamemos neurona, no es más que una [función matemática](#):



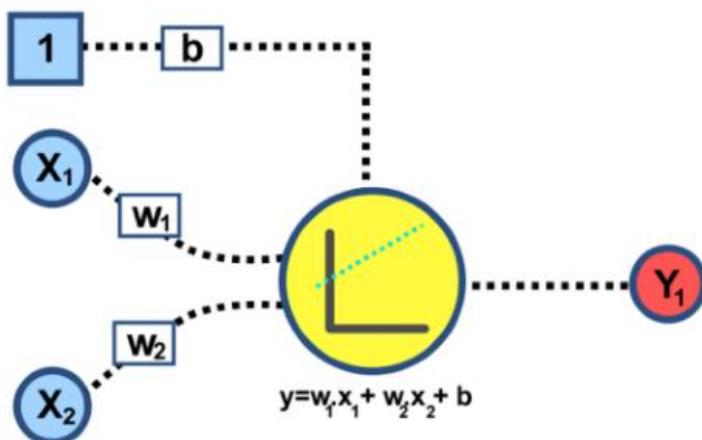
Esta función no es más que una suma ponderada de los valores de entrada: se le asigna un peso (W) a cada una de las entradas (X):



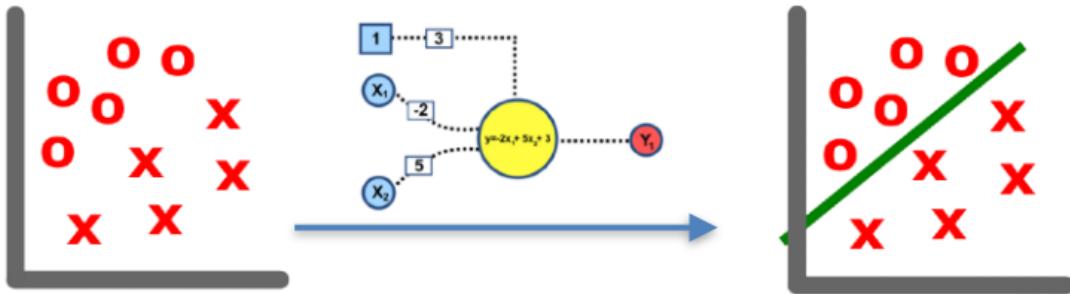
Esta función se corresponde con una recta, ya que si sumamos el peso (w) de cada entrada nos quedará una función tipo $y = a \cdot x + b$:



Aunque realmente nos falta añadirle el término independiente b :

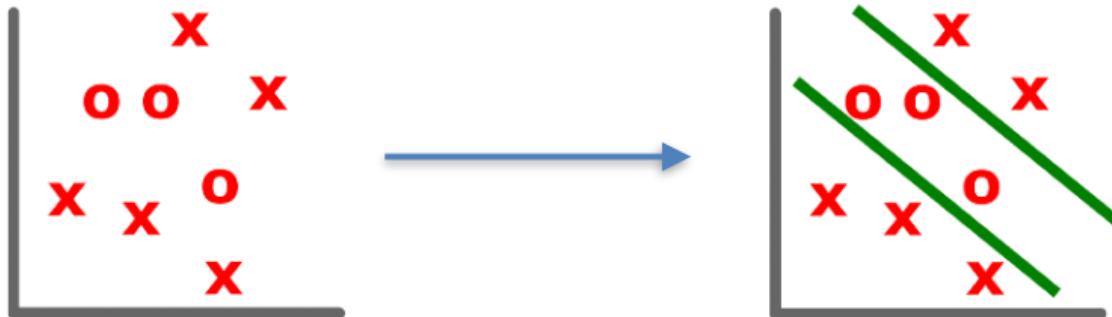


Por ejemplo, si tenemos un conjunto de datos como los que se ven en la gráfica izquierda de la siguiente imagen, podemos buscar que nuestra neurona calcule la recta que es capaz de separar unos datos de los otros (ver figura derecha):



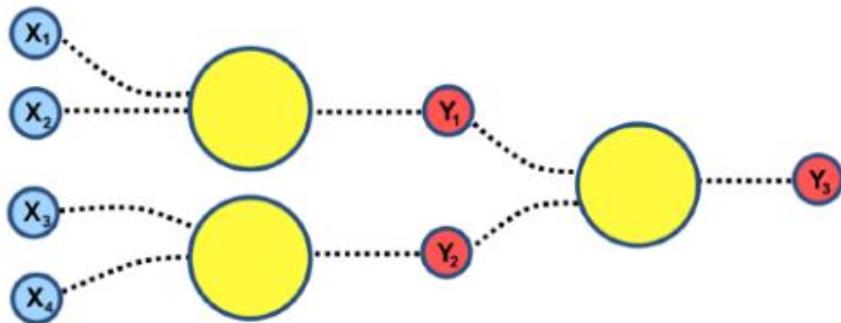
Pero ¿y si queremos separar con una función de este tipo unos datos como los que se muestran en la imagen izquierda de la siguiente figura?

Resulta que ahora no nos llegará con una sola recta (neurona). Necesitaremos otra neurona que genere una segunda recta y así poder solucionar el problema:



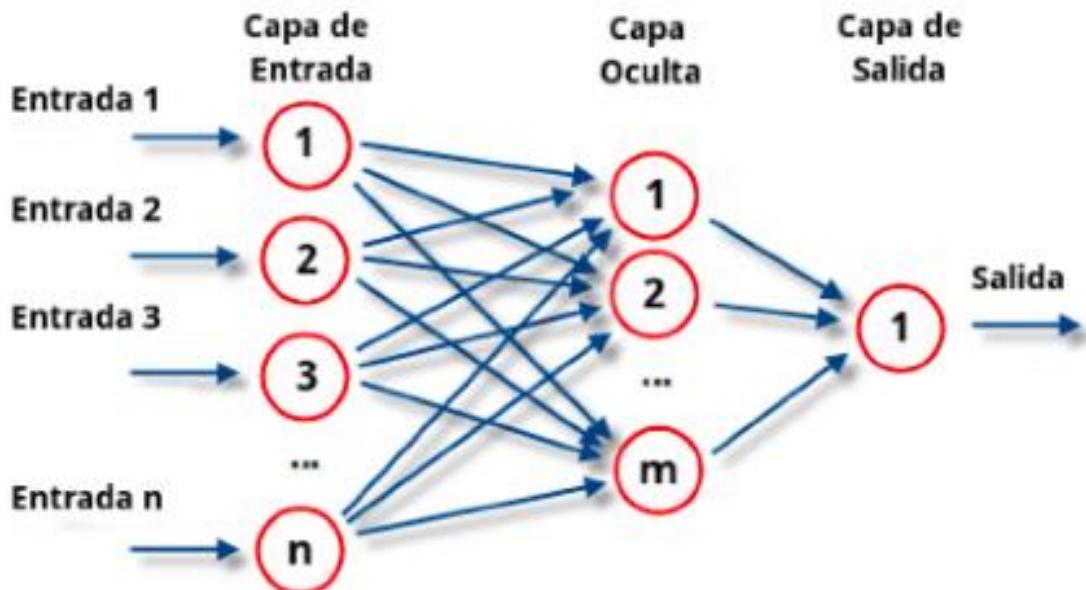
Colocando [varias neuronas trabajando juntas](#) tenemos un conocimiento jerarquizado, donde dos neuronas crearán dos rectas que separan los datos y una tercera unirá las salidas de las neuronas anteriores para darnos una única salida final:

Cada neurona formará una capa, y cuantas más capas tengamos más complejo será el conocimiento que elaboremos. Esta profundidad en la acumulación de capas es lo que le da nombre al **aprendizaje profundo**, el **deep learning**.



Las capas pueden clasificarse en tres tipos:

- **Capa de entrada:** Constituida por aquellas neuronas que introducen los patrones de entrada en la red. En estas neuronas no se produce procesamiento.
- **Capas ocultas:** Formada por aquellas neuronas cuyas entradas provienen de capas anteriores y cuyas salidas pasan a neuronas de capas posteriores.
- **Capa de salida:** Neuronas cuyos valores de salida se corresponden con las salidas de toda la red.



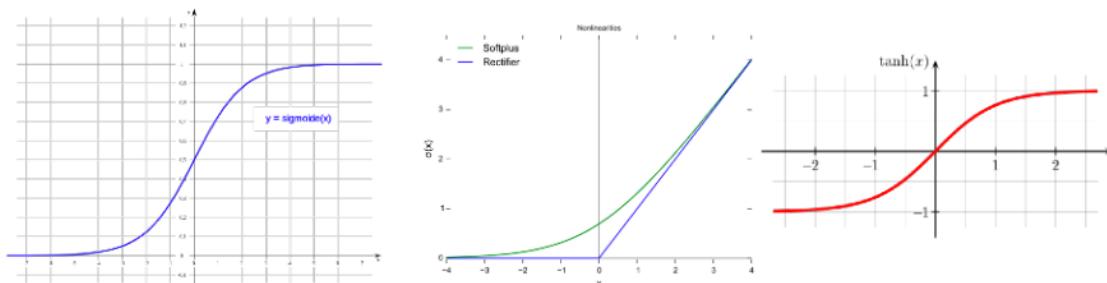
Pero, al conectar tantas neuronas en varias capas nos encontramos con un gran inconveniente: *cada neurona realiza una regresión lineal*. Es decir, calcula la función de una recta. Si sumamos todas estas rectas, matemáticamente obtendremos como resultado una sola recta, **que vendría a ser lo mismo que tener una sola neurona, por lo que la red colapsa (perderíamos toda la información de cada neurona de la red) y no sería útil.**



Si queremos evitar este problema tenemos que hacer que la suma nos proporcione algo diferente a una línea recta y para ello tenemos que hacer que cada una de las rectas sufra algún tipo de manipulación que las distorsione:



La solución es lo que se denomina **función de activación**, que se encarga de crear una distorsión **no lineal** en cada una de las neuronas. De esta manera, la suma de todas las funciones nos dará una nueva función completamente diferente: Existen diferentes funciones de activación que se usarán en función del tipo de problema que tengamos que solucionar. Así nos podemos encontrar con una función sigmoide, la relu (rectificadora), o la tanh (tangente hiperbólica):



$$y = \frac{1}{1 + e^{-x}}$$

$$f(x) = \max(0, x)$$

$$\tanh x = \frac{e^x - e^{-x}}{e^x + e^{-x}}$$

Las funciones de activación solucionan el *colapsamiento* de la red neuronal.

A modo de resumen, debemos tener claro los conceptos como: peso, sesgo y función de activación. La función de las neuronas es recibir las entradas, procesarlas (realizando una serie de operaciones matemáticas) y producir una salida. Estas operaciones matemáticas se resumen en tres:

- Calcula la suma ponderada de las entradas multiplicadas por sus respectivos **pesos** entre conexiones.

$$\sum_i^3 w_i \cdot x_i = w_1 \cdot x_1 + w_2 \cdot x_2 + w_3 \cdot x_3$$

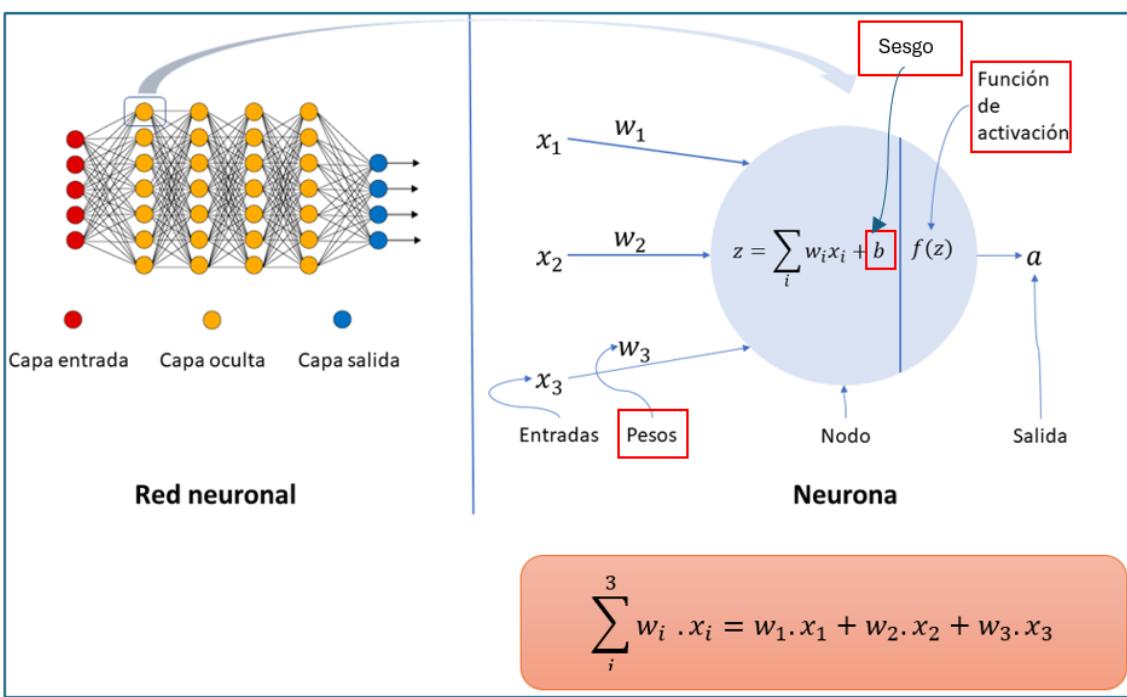
Las neuronas se conectan entre si mediante conexiones (en la imagen de abajo estamos viendo neuronas **densas**, porque se conecta con todas las siguientes). Cada una de estas conexiones tiene un **peso** asignado y que podemos entender como el valor numérico que representa la importancia de la conexión.

- b. A esa suma se le añade un valor numérico que recibe el nombre de **sesgo** y que le ayudará a ajustar la salida y permitir que la red aprenda patrones más complejos.

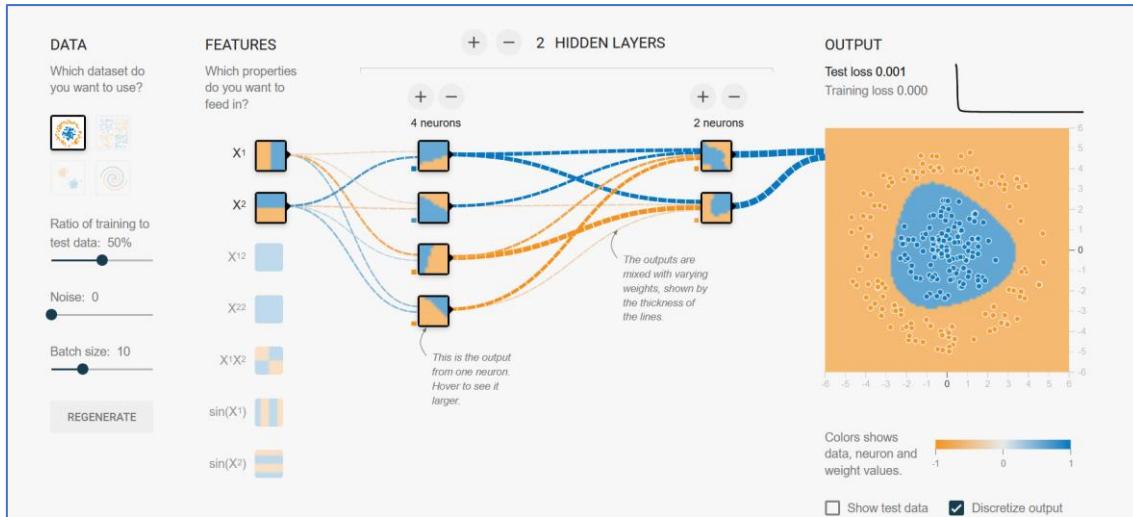
$$z = \sum_i^3 w_i \cdot x_i + b = w_1 \cdot x_1 + w_2 \cdot x_2 + w_3 \cdot x_3 + b$$

- c. A esa función resultante de sumas de productos más el sesgo y que llamé **z**, se le aplica una **función de activación**, buscando introducir NO linealidades y permitiéndole aprender.

$$\text{Salida} = a = f(z)$$



Actividad 13: Podemos crear nuestra primera red neuronal en la aplicación online de Google, “*Playground tensorflow*”. En el siguiente [link](#) explico cómo podemos jugar con las RNA:



Redes neuronales: <https://playground.tensorflow.org/>

Ya solo nos queda un último problema que solucionar y consiste en calcular el error de nuestra red neuronal. Ese error se podría calcular por el método del [error medio cuadrático](#) o por el de [descenso del gradiente](#). Sería más correcto decir que lo que buscamos es minimizar el error de nuestra red y para ello se ha desarrollado el **método de la propagación hacia atrás de errores** o **backpropagation**, que no es más que un algoritmo que se basa en el método de descenso del gradiente.

En este método la señal de salida se compara con la salida deseada y se calcula una señal de error para cada una de las salidas. Las salidas de error se propagan hacia atrás, partiendo de la capa de salida, hacia todas las neuronas de la capa oculta que contribuyen directamente a la salida. Sin embargo, las neuronas de la capa oculta solo reciben una fracción de la señal total del error, basándose aproximadamente en la contribución relativa que haya aportado cada neurona a la salida original. Este proceso se repite, capa por capa, hasta que todas las neuronas de la red hayan recibido una señal de error que describa su contribución relativa al error total ([vídeo de backpropagation](#)).

Redes neuronales: El algoritmo (backpropagation)

- La red neuronal con el valor de sus pesos es el modelo de ML
- El algoritmo de ML lo que hace es buscar el valor de los pesos óptimo para que la red sea capaz de clasificar correctamente los datos de entrenamiento y también nuevos datos parecidos pero distintos a aquellos (generalización)

Algoritmo de aprendizaje

- Se dan valores aleatorios a los pesos de la red. Así ya se tiene un primer modelo, que no clasifica muy bien, pero funciona.
- Los datos de entrenamiento se reparten en varios grupos (bunches) y para cada grupo:
 - Se calcula la salida de la red para cada dato del grupo (predicciones del modelo)
 - Se comparan las predicciones anteriores con la clasificación real (etiquetas)
 - Se usa el algoritmo **backpropagation** para actualizar el valor de los pesos, de manera que el error cometido en las predicciones se haga mínimo

Se repiten de forma iterativa estos pasos hasta que el error cometido sea aceptable.

En la siguiente web disponéis de 17 problemas cuya solución tiene diferentes niveles de dificultad dejando incluso crear vuestro propio modelo neuronal.

Actividad 14: Resuelve alguno de los retos de la [Caja de arena de Neuron \(cmu.edu\)](http://www.cs.cmu.edu/~tom/arena/)

Can I make a peanut butter and jelly sandwich? I need both peanut butter and jelly.

Solve for Outputs

Inputs

Have Peanut Butter	Have Jelly
0	0
0	1
1	0
1	1

Outputs

Activation Σ	Current Output $\Sigma > 1$	Desired Output
0	0	0
1	0	0
1	0	0
2	1	1

Activation: Σ

How does this work?

Uno de los ejemplos resuelto

4. El impacto de la IA

El impacto de la IA en la sociedad puede considerarse positivo y a la vez negativo. Por un lado, ofrece enormes beneficios, como la mejora de la productividad, la creación de nuevos productos y servicios, y el avance en sectores como la salud y el medio ambiente, entre otros. Sin embargo, también plantea importantes desafíos, desde la pérdida de empleos hasta problemas éticos y de sostenibilidad.

4.1. Situación actual (Aplicaciones)

Todos estos avances en las tecnologías de las redes neuronales, combinadas con la digitalización de la sociedad, la mejora en el hardware como las tarjetas gráficas con [GPU](#) y el hecho de que grandes empresas tecnológicas empezaran a invertir mucho dinero en investigación, señala el año 2012 como el del resurgir de la Inteligencia Artificial.

Por ejemplo, **DeepMind** impulsora del aprendizaje reforzado en 2010, o la empresa **OpenAI**, la que está detrás de Chat GPT, DallE o Whisper, fue fundada en 2015.

De todos modos, no ha sido hasta el 2020 cuando se han podido ver los resultados en nuevas aplicaciones y funcionalidades que nos han provocado sorpresa, admiración y finalmente cierto miedo o vértigo, incluso en personas que trabajan en ese campo, y que ha llevado a plantearse la necesidad de una [moratoria de unos meses](#) en su desarrollo para recapacitar sobre sus consecuencias.

Algunos de estos avances son el archiconocido Chat [GPT](#), pero también otras [IAS generativas](#) de imágenes, vídeos, música, ... sistemas que han sido capaces de abstraer conceptos y de generar *entidades* que antes no existían; ser “creativas”.

Gracias a los últimos avances en los sistemas de redes neuronales, como son las redes neuronales [convolucionales](#), los [transformers](#) o las [redes antagónicas generativas](#) (GANs) estamos viendo cosas cada vez más sorprendentes. El año 2022 ha sido el año de la IA, pero eso no significa que no lo sea el 2023. Las aplicaciones de la IA están **revolucionando sectores** como la ciencia, la tecnología, la búsqueda de nuevos materiales, la [medicina personalizada](#) y sectores como las finanzas, la logística o las comunicaciones.

El año que se hizo viral [DALL-E](#), la IA que creaba imágenes a partir de unas pocas palabras a modo de sugerencias. También fue el año del **ChatGPT**, un chat en el que podemos **entablar conversación con una inteligencia artificial** para preguntarle cualquier cosa y obtener, en tiempo

real, en teoría, respuestas satisfactorias. Detrás de estos dos proyectos de inteligencia artificial se encuentra **OpenAI**, una de las empresas más punteras en el sector de la IA.

Una de las aplicaciones de la inteligencia artificial que más han generado debate durante 2022 se centra en la creación de imágenes y vídeos a partir de texto o recursos gráficos concretos como pueden ser dibujos sencillos. Así, en cuestión de unos pocos minutos, es posible generar imágenes con gran calidad. **Stable Diffusion** fue una de las IAs especializadas en esta tarea más populares junto con DALL-E. Su utilización gratuita y/o de pago ha dado que hablar porque sus creaciones se generan en minutos y ofrecen una gran calidad artística. Sin embargo, existe el debate en Internet sobre si las imágenes creadas con IA **son arte o no**, ya que no han sido generadas por humanos. Y es más, aunque el maravilloso mundo de la tecnología y la inteligencia artificial pretende crear un producto que “mejore la vida de las personas”, somos nosotros, las personas, las que a veces hacemos mal uso de estas tecnologías y **convertimos una utilidad en un problema**. Todos hemos oido hablar de los *deepfakes*, que se han convertido en un tema candente y de gran preocupación en los últimos años. Es importante distinguir la diferencia entre “información errónea” y “desinformación” en el ancho internet:

- La información errónea se relaciona con información incorrecta que se comparte con la creencia de que es verdad, mientras que la desinformación es información incorrecta compartida **intencionalmente** para engañar.

A veces, las IA se *mal usan* con esta intención, engañar. Ya sea en textos, imágenes, audios o vídeos. En los siguientes links podéis entender cómo podemos diferenciar *deepfakes* de audio e imágenes:

- [Deepfakes de audio y clones de voz](#)
- [¿Cuál de los rostros es real?](#) Despues de interactuar, entra en el menú “[Aprender](#)” y verás qué señales deja la IA a la hora de crear esas imágenes.

El proyecto más relevante en la temática de IA en Galicia es el [proyecto “Nós”](#). Este proyecto se divide en tres partes aportándonos una aplicación de reconocimiento de voz (ASR), una de síntesis de voz también llamada texto a voz (TTS) y un traductor neuronal multilingüe diseñado específicamente para el gallego y a cuyas demos podemos acceder desde la opción de [recursos](#) del propio proyecto. La demo “[Traductor Nós](#)” está preparada para trabajar con los idiomas, gallego, castellano e inglés. Las demos de aplicaciones relacionadas con las dos tecnologías de IA utilizan redes neuronales y ponemos interactuar con ellas en los siguientes links: [TTS](#) y [ASR](#).



Relacionado con la cultura, en 2023 Google ha recreado, con ayuda de la IA, la mente del inventor, científico y artista Leonardo da Vinci en una plataforma que nos permite viajar por sus invenciones con más de mil imágenes detalladas de sus obras. En esta web el usuario podrá conocer la vida de Leonardo utilizando los diferentes juegos y test que nos desvelarán detalles de su vida e incluso, recrear su pensamiento. La plataforma [Inside a Genius Mind](#) es la que alberga este proyecto.

Algunas aplicaciones más relevantes de estos últimos años son:

- [Whisper](#) (de voz a texto y dejó de estar abierta a los usuarios en junio de 2023)
- [Plegamiento de proteínas](#) (AlphaFold 2)
- [Procesamiento Lenguaje Natural](#) (NLP)
- [De texto a vídeo: Gen2](#) de la empresa [Runway](#)
- [Generador de música, suno](#)
- [chatGPT4 y los plugins](#)
- [Bing+chatGPT](#)
- [Copilot y Codex](#)
- [Vídeo generado artificialmente NVDIA](#)
- Generadores de imágenes: [Dalle2](#), [MidJourney](#) o [Stable Diffusion 2](#). Actualmente, [Playground AI](#) nos permite crear 1000 imágenes libres de pago por día. [BlueWillow](#) es un generador de ilustraciones. [Leonardo AI](#) crea imágenes con un estilo diferente.

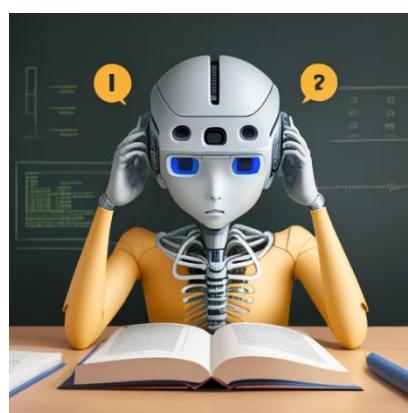


Imagen generada con Playground AI (Inteligencia artificial, educación)

Las aplicaciones que integran IA han tenido que ser entrenadas. Muchas de ellas parecen de forma gratuita y, cuando su modelo funciona y la aplicación ofrece resultados más precisos (la IA no se equivoca y deja de tener “*alucinaciones*”), restringen su versión gratuita e incluyen versiones de pago. Por internet podemos encontrar repositorios de herramientas y aplicaciones de IA, así como de bibliotecas.

Algunos repositorios en los que podemos buscar herramientas por categorías, funcionalidades y precios, son los siguientes:

- [There's an AI for that](#)
- [AIFINDY](#)
- [Pantheon Almanac](#)

El chat GPT en la versión actual, aunque ha mejorado mucho sus respuestas en relación a sus modelos anteriores, a veces nos aporta resultados incorrectos. Pensar que siempre responde a nuestras preguntas, y da igual la dificultad de las cuestiones que le planteemos. Cuando se equivoca, decimos que ha tenido una “alucinación”. Estos errores podemos minimizarlos intentando que rzone, usando técnicas de *prompting* como añadir la muletilla “Piensa, paso a paso....” a la pregunta que le hacemos al chat, pero esto no es suficiente. La empresa Open AI ha estado trabajando en estas limitaciones y ha sacado la nueva versión del Chat GPT que llama [OpenAI o1](#) (disponible en Chat GPT Plus) y que le permitirá pensar sus respuestas (razonar) antes de responder, por lo que no es tan rápido a la hora de contestar. Según la compañía, o1 “pensará” antes de ofrecer una respuesta, “**al igual que lo haría una persona**”. Incluso van más allá, dicen que con el paso del tiempo [refinará su pensamiento y también reconocerá sus errores](#). Digamos que, para acciones comunes, ChapGPTo4 es más rápido y funciona bien pero, para problemas más complejos, la solución óptima pasa por usar OpenAI o1, una herramienta para el futuro. El modelo 01 Pro puede procesar imágenes (y no sólo texto), pero su visión computacional no trabaja bien y se equivoca. Tras la salida del modelo OpenAI o1, la empresa ha seguido trabajando y hoy podemos utilizar el modelo [OpenAI o3](#) que se presenta como un buen solucionador de tareas científicas y técnicas de una gran complejidad y mejorando mucho el procesamiento de imágenes.

4.2. Retos que debe solucionar la IA

Los deslumbrantes avances de la IA no deben cegarnos a la hora de hacer un análisis crítico de estos sistemas. El mundo ya está cambiando debido a la aparición de todas estas IAs y estos cambios van a ser profundos, equiparables a los que se produjeron con la aparición de la

máquina de vapor, la electricidad o internet. Pero, para que estos cambios sean positivos debemos solucionar algunos aspectos que tienen que ver, por ejemplo, con los datos que usamos al entrenar estos sistemas, con las decisiones que toman y con los sesgos que puedan contener.

Así como, por ejemplo, los sindicatos son una respuesta a los problemas de la primera revolución industrial, las normas ISO a la segunda o el teletrabajo y el comercio global de la tercera, queda por saber qué pasará con esta cuarta. Sin duda, la privacidad, la ética, la ciberseguridad o la huella de carbono son problemas que ya están encima de la mesa.



DallE2, Midjourney o StableDiffusion fueron entrenados con miles de imágenes publicadas en la red, casi todas ellas con derechos de autor, pero todas las imágenes generadas son libres de derechos, [¿es correcto hacer esto así?](#). Y los sistemas de reconocimiento facial, [¿son perfectos? ¿Y qué impacto ambiental tiene entrenar estos sistemas?](#) Ya están abiertos una [serie de debates](#) y de reclamaciones que demuestran que es necesario **legislar y regular** este nuevo sector, de manera que sean explicables, confiables y responsables. Y también verdes y open source.

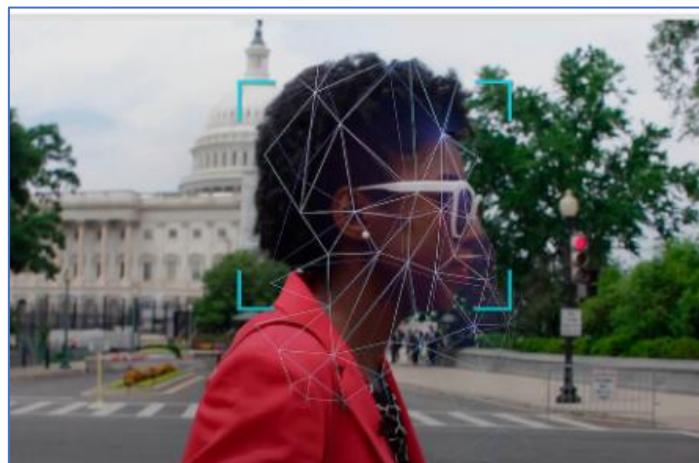
1. La **explicabilidad** se refiere a la capacidad de comprender y explicar cómo funciona un sistema de IA y cómo toma decisiones. Esto es importante porque, a medida que la IA se vuelve más compleja y se utiliza en áreas críticas como la salud o la seguridad, es necesario que los usuarios puedan comprender cómo se están tomando las decisiones y cómo se están utilizando los datos. La falta de explicabilidad puede generar desconfianza en los usuarios y en la sociedad en general.

2. La **confiabilidad** se refiere a la capacidad de los sistemas de IA para tomar decisiones precisas y consistentes en diferentes situaciones. Es

importante que los sistemas de IA se prueben rigurosamente para asegurarse de que sean precisos y confiables, ya que las decisiones equivocadas pueden tener consecuencias graves.



3. La **responsabilidad** se refiere a la necesidad de garantizar que los sistemas de IA sean éticos y justos. Esto incluye la protección de la privacidad y la seguridad de los datos, así como la eliminación de sesgos y discriminación en los sistemas de IA. A pesar de los avances, la ética de la IA no se puede automatizar. La supervisión humana sigue siendo esencial para garantizar que las decisiones tomadas por sistemas automáticos sean justas y éticas.



4. Los **algoritmos verdes** nacen de la necesidad de reducir la huella de carbono en el proceso de entrenamiento de estos sistemas. No existen mediciones exactas de cuanta energía consume cada año a nivel global la inteligencia artificial, pero las estimaciones sobre aspectos concretos permiten hacerse una idea. Según investigadores de la Universidad de Massachusetts Amherst, entrenar a un solo modelo de IA tiene unas emisiones parecidas a las de cinco coches en toda su vida útil. Por ello, es fundamental desarrollar **algoritmos más eficientes y tecnologías sostenibles** que reduzcan el consumo de recursos y contribuyan al objetivo de **sostenibilidad** ambiental.

Ya hay herramientas como [Code Carbon](#) que nos permiten estimar la huella de carbono asociada con el algoritmo del entrenamiento de modelos de IA, promoviendo prácticas más ecológicas en el sector.

5. Algunos sistemas actuales de IA son **open source**, y lo que pretenden es abrir a toda la comunidad las herramientas necesarias para desarrollar estos sistemas. Esto tiene una serie de ventajas y beneficios para la sociedad en cuanto a:
 - **Transparencia:** al ser open source, el código de la IA está disponible públicamente, lo que aumenta la transparencia y la claridad en torno a cómo funciona la tecnología. Esto significa que los usuarios pueden comprender mejor cómo se toman las decisiones, detectar errores y hacer sugerencias de mejora.
 - **Colaboración:** El desarrollo de una IA open source permite que la comunidad contribuya a su mejora y evolución. Esto significa que más desarrolladores pueden trabajar en el código, lo que a su vez puede llevar a una mayor innovación y mejores resultados.
 - **Adaptabilidad:** Los usuarios pueden personalizar y modificar el código para satisfacer sus necesidades específicas, lo que permite una mayor flexibilidad en la implementación y uso de la tecnología.
 - **Accesibilidad:** una IA open source puede ser utilizada por una amplia gama de personas y organizaciones, incluidas aquellas con recursos limitados. Esto puede democratizar el acceso a la tecnología y aumentar su adopción.
 - **Seguridad:** Más desarrolladores pueden revisar y auditar el código en busca de posibles vulnerabilidades o debilidades de seguridad. Esto puede ayudar a garantizar que la tecnología sea más segura y confiable.

4.3. Conclusiones sobre la situación actual de la IA

La **Inteligencia Artificial (IA)** ha dejado de ser una disciplina futurista y se ha integrado profundamente en la vida cotidiana, afectando a múltiples aspectos de la sociedad. En la actualidad, la IA influye en la economía, el empleo, la educación, la salud, la política, el entretenimiento y más. Sin embargo, aunque su impacto es indiscutible, todavía hay mucho por explorar, tanto en sus aplicaciones actuales como en los posibles efectos que podría generar en el futuro.

Impacto actual de la IA en la sociedad

1. **Transformación de la economía y los mercados laborales:** La **automoción** y la **automatización** están entre los sectores más transformados por la IA. En el ámbito industrial,

la IA está optimizando las cadenas de producción y ofreciendo soluciones más rápidas y eficientes. Sin embargo, uno de los impactos más significativos es el cambio en la **naturaleza del trabajo**. Muchas tareas repetitivas y manuales están siendo reemplazadas por robots y sistemas automáticos, lo que plantea una disyuntiva importante: la **desaparición de empleos** tradicionales frente a la creación de nuevos puestos que requieren habilidades más especializadas, como el manejo y mantenimiento de sistemas de IA.

2. **Transformación en la salud:** En la medicina, la IA ha permitido avances significativos en el diagnóstico y tratamiento de enfermedades. Los algoritmos de IA son capaces de analizar grandes volúmenes de datos médicos para identificar patrones que podrían pasar desapercibidos por los médicos humanos. Esta capacidad está revolucionando la forma en que se diagnostican enfermedades, optimizando el tiempo de los profesionales de la salud y mejorando la precisión. Además, la IA se está utilizando en la **medicina personalizada**, adaptando tratamientos y terapias a las necesidades individuales de los pacientes.
3. **Cambio en el entretenimiento y la cultura:** Las plataformas de **streaming** como Netflix, Spotify y YouTube utilizan sistemas de recomendación basados en IA que analizan los hábitos de consumo de los usuarios para ofrecer contenido personalizado. Estos sistemas no solo mejoran la experiencia del usuario, sino que también cambian la forma en que las empresas crean y distribuyen contenido. Además, la IA está comenzando a utilizarse en la **creación de arte**, generando música, imágenes, y literatura, lo que plantea nuevas preguntas sobre la autoría y el valor del arte generado por máquinas.
4. **Mejora de la experiencia del usuario en tecnología:** Los **asistentes virtuales** como Siri, Alexa y Google Assistant están basados en IA y han transformado la manera en que interactuamos con nuestros dispositivos. Estas tecnologías permiten la interacción por voz y están integrando la IA en dispositivos domésticos, vehículos y en el uso del internet. La integración de la IA hace que los servicios sean más rápidos, intuitivos y accesibles, cambiando la forma en que gestionamos nuestras vidas cotidianas.
5. **Impacto en la política y la democracia:** En el ámbito político, la IA ha introducido tanto oportunidades como riesgos. Por un lado, los **algoritmos de análisis de datos** permiten a los partidos políticos entender mejor las necesidades de los votantes y personalizar sus mensajes, lo que puede ayudar a mejorar la relación entre los ciudadanos y sus representantes. Sin embargo, la manipulación de estos algoritmos para influir en las

elecciones, como se ha observado en el caso de las **fake news** o en las campañas de desinformación, también representa un riesgo para la democracia y la transparencia.

Impactos futuros de la IA en la sociedad

1. **Automatización y el futuro del trabajo:** A medida que los sistemas de IA continúan evolucionando, se espera que la automatización afecte a más industrias y sectores. En el futuro, podrían ser reemplazadas muchas más profesiones que actualmente requieren habilidades cognitivas o creativas, como las que desempeñan los diseñadores gráficos, traductores, o incluso ciertos roles de gestión. Sin embargo, a medida que algunas tareas sean automatizadas, surgirán nuevas oportunidades de empleo, especialmente en áreas como el desarrollo de IA, la ética de la IA, y la ciberseguridad. La **recualificación** de la fuerza laboral será esencial para mitigar los impactos negativos.
2. **Desafíos éticos y sociales:** Con el creciente uso de la IA, surgen preocupaciones sobre su **impacto ético** y su influencia en la **privacidad** de los individuos. Los sistemas de IA están basados en grandes volúmenes de datos, lo que genera preocupaciones sobre la recopilación y el uso de datos personales. La falta de **transparencia** en cómo se toman las decisiones automatizadas y el posible sesgo en los algoritmos son dos cuestiones que deberán abordarse de manera urgente. Es probable que en el futuro se desarrollen **marcos legales** más estrictos y normativas internacionales para garantizar que los sistemas de IA respeten los derechos fundamentales y no perpetúen desigualdades.
3. **IA explicativa y controlada:** A medida que los sistemas de IA se vuelven más complejos, será necesario que los algoritmos sean **explicables y transparentes**. En aplicaciones críticas como la justicia, la salud o la seguridad, los usuarios y las autoridades deben poder comprender cómo los sistemas de IA toman decisiones. Para ello, se investigará cómo crear modelos que no solo sean eficaces, sino también fáciles de auditar y comprender, lo que podría cambiar la forma en que los sistemas son regulados y supervisados en el futuro.
4. **IA y sostenibilidad:** En el futuro, la IA jugará un papel crucial en la **sostenibilidad** y la lucha contra el **cambio climático**. Los sistemas de IA pueden ayudar a optimizar el uso de recursos en diversas industrias, reduciendo el desperdicio de energía y mejorando la eficiencia. Además, se prevé que la IA contribuya a la **gestión inteligente de las ciudades**, promoviendo un desarrollo urbano más sostenible, mejorando la gestión de residuos y gestionando las infraestructuras de manera más eficiente. Sin embargo, también existen preocupaciones sobre el impacto ambiental de la IA, especialmente en cuanto a la energía consumida.

durante el entrenamiento de modelos complejos, lo que impulsará el desarrollo de **algoritmos más ecológicos** y eficientes.

5. **IA general y los desafíos de la superinteligencia:** En el largo plazo, uno de los mayores objetivos de la investigación en IA es la creación de una **IA general (AGI)**, que sea capaz de realizar cualquier tarea cognitiva humana, independientemente del contexto. Aunque aún estamos lejos de alcanzar este nivel de inteligencia artificial, los avances continúan. La creación de una IA general podría tener un impacto disruptivo, modificando por completo la estructura social, económica y política. En este escenario, surgirían debates sobre el **control** de las superinteligencias y la **seguridad** en su implementación, ya que podría ser más difícil prever las consecuencias de su uso.

El impacto de la Inteligencia Artificial en la sociedad es tanto prometedor como desafiante. En el presente, está transformando sectores claves y mejorando muchas facetas de nuestra vida diaria, desde la salud hasta la forma en que interactuamos con las máquinas. No obstante, su futuro está lleno de incertidumbres y retos, como la automatización del empleo, los dilemas éticos y la sostenibilidad. A medida que la IA evoluciona, será fundamental establecer regulaciones y marcos éticos para garantizar que sus beneficios sean aprovechados de manera justa y responsable. Solo entonces podremos asegurarnos de que la IA sea una herramienta para el bien común y el progreso social.

[Conclusiones escritas por chatGPT 4o]

Existen [herramientas](#) para detectar si un texto ha sido o no generado por una IA o por un humano. **AI Text Classifier** de OpenAI nació como una respuesta a la creciente preocupación sobre la capacidad de los modelos de lenguaje avanzados, como **GPT-3** y **GPT-4**, para generar textos altamente coherentes y naturales que podrían confundirse con escritos humanos. Con el avance de la inteligencia artificial, se comenzaron a plantear una serie de preguntas y problemas relacionados con el **uso ético** de estos modelos, especialmente en el ámbito educativo, académico y profesional. El proyecto se olvidó en 2023 por problemas relacionados con su precisión y efectividad (falsos negativos o falsos positivos).

El clasificador considera que es **probable** que el texto sea generado por IA.

Como vemos, este clasificador de textos **no era totalmente fiable** y sabe que puede dar falsos positivos, sobre todo a la hora de identificar textos de IA, por lo que siempre nos dirá que ha **sido probablemente** creado por uno u otro, pero no afirmará nada de forma rotunda.

Actualmente, existen diversas herramientas para clasificar si un texto ha sido generado o no por una IA. Algunas son las siguientes:

- [**GPTZero**](#) (una herramienta popular para detectar texto generado por IA).
- [**CopyLeaks**](#) (que ofrece una herramienta de detección de texto generado por IA y de plagio).
- [**Turnitin**](#) (utiliza herramientas para detectar plagio y contenido generado por IA).

5. Bibliografía

[1] *Apuntes Machine Learning Xornadas de Apetega*, de Jorge Gómez Suárez

[2] [*Inteligencia artificial desde cero, con Python y Tensorflow*](#), de Ringa Tech

[3] [*Aprende Inteligencia Artificial*](#), de DotCSV

[4] [*Experience IA*](#), desarrollado en colaboración con [*Raspberry Pi Foundation*](#) y [*DeepMind*](#)