

# Introducción al mundo de la Inteligencia Artificial.

## Una perspectiva multidisciplinar

*F. J. Serón*  
*Septiembre de 2024*

### Índice

Presentación.....	2
Tipos de problemas/decisiones en nuestro mundo.....	3
Problemas.....	6
La inteligencia artificial.....	13
¿De dónde podemos extraer modelos para la inteligencia artificial? .....	16
Simbolistas .....	18
Bayesianos.....	25
Analogistas .....	33
Evolutivistas.....	35
Conexionistas .....	39
Refuerzo .....	45
Otro tipo de taxonomía basada en las técnicas de aprendizaje de máquina. ....	48
Algoritmos supervisados. ....	51
Algoritmos no supervisados. ....	52
Algoritmos combinados por conjuntos (ensemble).....	53
Inconclusión final .....	55

## Presentación

Estamos en la primera mitad del año 2024 y es casi imposible ignorar los asombrosos avances de la inteligencia artificial en estos días. Esto es especialmente cierto en el campo de la inteligencia artificial generativa, donde vemos un número creciente de modelos impresionantes que pueden crear imágenes, texto, vídeo y música. Estos desarrollos han capturado la imaginación popular y las empresas e instituciones están luchando por determinar cómo utilizar la IA en su organización.

*¿Cuál es el resultado? Algunas empresas están luchando por determinar cómo utilizar la inteligencia artificial, mientras que otras descubren que el panorama actual de la inteligencia artificial es complejo y difícil de navegar.*

**¿Cuál es el objetivo de este texto?** *Presentar un conjunto de modelos de la inteligencia artificial que nos han permitido y nos permiten resolver problemas y avanzar en el conocimiento.*

**¿Para qué?** *Antes de responder, déjeme pedirle que haga la siguiente reflexión. Imagino que en su cabeza tendrá la estructura organizativa de la empresa o institución en la que desarrolla su labor, ya sea estudiantil o profesional, y sería capaz de describirla e incluso de explicarla con profundidad. Pero ahora vea si tiene en su cabeza algo análogo cuando piensa en la inteligencia artificial. Puede que su estructura mental sea el conjunto vacío, o sea muy estereotipada, del tipo ciencia ficción, o tenga una visión borrosa.*

*Lo que pretendo con estas páginas, es que pueda pasar de cualquiera de ellas a por lo menos una visión desordenada y en el mejor de los casos a una visión coherente.*

**¿Qué vamos a conseguir?** *Mejorar la capacidad para discernir como se relacionan entre sí los aspectos informacionales relativos a la IA para captar su estructura y significado. Es decir, para poder conocer los fundamentos, tener perspectiva y saber seleccionar.*

## Tipos de problemas/decisiones en nuestro mundo

*No hay que ser muy lince para comprobar que vivimos en un mundo muy complejo en el que:*

- *Hay que resolver problemas muy difíciles*
- *Hay que tomar decisiones*
- *Hay que construir modelos de pensamiento que nos ayuden*

**Para resolver problemas difíciles** lo que necesitamos es:

- *Datos e información*
- *Comprender las relaciones y conexiones entre los diferentes trozos de información (Conocimiento)*
- *Comprender qué tipo de conocimiento aplicar (Sabiduría)*
- ...

**La toma de decisiones** es el proceso mediante el cual se realiza una elección entre las opciones o formas que se tienen para resolver diferentes situaciones de la vida (problemas) en diferentes contextos y dominios. A veces se busca la elección de la mejor opción, entonces los siguientes términos pueden ayudar a tomar la decisión según el resultado que se busque:

- *Maximizar: Tomar la mejor decisión posible.*
- *Satisfacer: Elegir la primera opción que sea mínimamente aceptable satisfaciendo de esta forma una meta u objetivo buscado.*
- *Optimizar: La que genere el mejor equilibrio posible entre distintas metas.*

Y para finalizar hay que **construir modelos de pensamiento**, Un modelo es una representación de la realidad sobre la que podemos hacer inferencias. Lo relevante aquí es que un modelo debe permitirnos hacer predicciones sobre datos no observados a partir de unos ciertos datos conocidos.

Un modelo puede ser muy simple de partida y si es necesario, hacerlo evolucionar transformándolo poco a poco en muy complejo. Es decir, inicialmente suelen ser simplificados y si con su uso se observa un buen comportamiento, posteriormente se van refinando. Con el paso del tiempo, se pueden convertir en productos confiables, pero para eso hay que dejar pasar tiempo. Normalmente se trabaja con más de un modelo.

Resumiendo, los modelos proporcionan un medio para transformar datos e información en conocimiento. Imponen lógica a nuestro pensamiento y proporcionan condiciones bajo las cuales la intuición puede navegar. Dicho conocimiento puede ser interpretado y entendido por seres humanos e incluso por máquinas. Utilizando este proceso de crear y mejorar modelos nos va haciendo mejores pensadores.

Por ejemplo, el modelo que se utiliza para analizar el calentamiento global, es un modelo que va evolucionando conforme se van recogiendo más datos y patrones de comportamiento diferentes que con el paso del tiempo se van añadiendo al modelo para que sus predicciones sean cada vez más próximas a la realidad.

Las razones para construir un modelo son las siguientes:

- *Explican (diferente a predecir)*
- *Guían a través de la colección de datos*
- *Iluminan comportamientos dinámicos*
- *Sugieren analogías dinámicas*
- *Descubren nuevas preguntas*
- *Promueven un hábito mental científico*
- *Acotan los resultados en rangos plausibles*
- *Ofrecen análisis de opciones críticas en tiempo real*
- *Demuestran disyuntivas/sugieren mejoras*
- *Desafían la robustez de la teoría frente a perturbaciones.*
- *Exponen incompatibilidades en los datos*
- *Forman de manera práctica*
- *Disciplinan el método de dialogar*
- *Educán al público en general*
- *Revelan que lo aparentemente simple puede ser complejo y que lo aparentemente complejo puede ser simple.*
- ...

Pero no hay que olvidar que Los modelos son simplificaciones necesarias. Los modeladores afirman que "todos los modelos son incorrectos", pero a pesar de ello "algunos" son útiles. Lo que es una gran verdad Las inexactitudes de los modelos obligan a dos tipos de comportamientos:

- *El primero es que constantemente deben refinarse y mejorarse.*
- *El segundo, debido al hecho de que los modelos no son perfectos, se deben conseguir colecciones de modelos diferentes.*

*En sistemas complejos, los modelos simples no funcionan. Hacer modelos no es fácil. Para que tenga sentido nuestro mundo, una persona necesita una multitud de modelos para conseguir crear un diálogo interno entre la lógica y el comportamiento del mundo. Cuantos más modelos se dispone, más posibilidades hay para recombinar sus partes y producir mejores modelos, lo que vamos a necesitar para enfrentarnos a la creciente complejidad del mundo.*

## Problemas

Un problema es una situación cuya solución desconocida debe obtenerse a través de algún método.

**Los problemas se pueden clasificar** de muchas maneras, Una de ellas: es la siguiente:

- **Convergentes**, que se suelen caracterizar por ser lógicos o estructurados que tienen respuestas únicas y definidas. Por ejemplo, los teoremas de existencia y unicidad de solución de algunos problemas matemáticos.
- **Divergentes**, que se caracterizan por el hecho de que, ante un mismo problema, diferentes personas encuentran diferentes soluciones. Por ejemplo, El encargo, a varias compañías de publicidad, de hacer un anuncio televisivo de un producto concreto. Cada propuesta será diferente del resto.

Además, un mismo problema puede presentarse en diferentes contextos. Entendiendo por contexto el marco sobre el que nos movemos, sentimos, pensamos y decidimos, e influye directamente en nosotros, aunque no siempre nos demos cuenta.

Los **contextos** también se pueden clasificar en función de varios parámetros como son:

- La observabilidad:
  - Totalmente observables o también denominados contextos de certeza. En estos casos las opciones de solución van a causar siempre resultados conocidos e invariantes. Por ejemplo, la dinámica de un péndulo simple bajo pequeñas perturbaciones.
  - Parcialmente observables, en estos casos la información para solucionar el problema es completa, es decir, se conoce el problema y las posibles soluciones, pero no se conoce con certeza cual se va a producir, lo que se manifiesta en que las posibles opciones de solución tienen cierta probabilidad.

- Ambiente de incertidumbre, Con esa expresión nos referimos a aquellos contextos en los que hay una alta inestabilidad, y existen muchos factores que no están bajo control. En estas ocasiones no pueden asignarse probabilidades a los eventos posibles. Por ejemplo, suelen ser contextos típicos que surgen después de que se hayan producido crisis
- La variabilidad en la respuesta del sistema:
  - Determinista. Siempre que se realiza la misma acción se obtiene el mismo resultado.
  - Estocástico. En el resultado de la acción interviene la incertidumbre o el azar.
- Comportamiento de dependencia:
  - Episódico. Una acción no influye en las acciones posteriores. Las experiencias son independientes.
  - Secuencial. Una acción influirá en las decisiones futuras.
- Comportamiento temporal:
  - Estático. El entorno no cambia con el paso del tiempo.
  - Dinámico. El entorno evoluciona
- En función del tipo de variables observables:
  - Discreto. Las variables sólo pueden tomar ciertos valores.
  - Continuo. Las variables pueden tomar cualquier valor.
- ...

Ejemplos de tipos de contextos en el caso de varias actividades conocidos:

- Juego del ajedrez: Totalmente observable. Determinista. Secuencial. Estático. Discreto.
- Juego del póker: Parcialmente observable. Determinista. Secuencial. Estático. Discreto.
- Diagnóstico médico: Parcialmente observable. Estocástico. Secuencial. Dinámico. Continuo.
- Análisis de imagen: Totalmente observable. Determinista. Episódico. Dinámico. Continuo.

Otro enfoque que permite caracterizar a los problemas, es el que cae bajo el **concepto de dominio**, que consiste en el conjunto de conceptos interrelacionados que es necesario conocer para poder intentar resolver el problema.

Los dominios también se pueden clasificar en diferentes categorías:

- Simples. Cuando se pueden identificar fácilmente las causas y los efectos. Por ejemplo, el juego del tres en raya.
- Complicados. Puede haber múltiples soluciones correctas para una misma problemática. Se suelen seguir buenas prácticas. Por ejemplo, el juego del ajedrez.
- Complejos. Las soluciones encontradas rara vez son replicables, con los mismos resultados, a otros problemas similares. Normalmente el número de variables que intervienen en el problema suelen ser muchas. Para su resolución, se requieren niveles altos de creatividad, innovación, interacción y comunicación. En estos casos lo adecuado es ensayar y probar. La razón es que las relaciones de causa y efecto no se repiten y son impredecibles. Por ejemplo, la crianza de un hijo.
- Caóticos. Estos problemas son los más amenazantes y se dan cuando surgen periodos de crisis en los que los hechos son inciertos, los datos se discuten, las opiniones son diversas y las decisiones hay que tomarlas de manera urgente.



- Desordenados. Nos movemos en el espacio desordenado cuando no sabemos en qué dominio estamos. Se considera que se está en una zona peligrosa, ya que no podemos caracterizar de alguna manera las situaciones ni determinar la forma de actuar.
- En este caso las personas involucradas tienen que interpretar las situaciones y actuar según preferencias personales, confiando en que la suerte los conduzca hacia alguno de los otros dominios citados previamente. Ejemplos son las situaciones en las que se hacen presentes el pánico, el estrés y la incertidumbre en grandes muchedumbres, desconciertan a cualquiera.

La herramienta fundamental la denominamos **razonamiento**, que es la capacidad para resolver problemas, extraer conclusiones y aprender de manera consciente de los hechos, estableciendo conexiones lógicas necesarias entre ellos.

Tipos de razonamientos del ser humano o como ha creado el camino para hacer inferencias a partir de información conocida son:

- Deductivo. Consiste en la aplicación correcta de las relaciones lógicas entre enunciados que llevan a conclusiones válidas (aunque no tienen por qué ser ciertas). Este proceso psicológico parte de unas premisas y razona para llegar a una conclusión basada en y apoyada por estas premisas. Si el razonamiento se hizo correctamente, resulta en una deducción válida: la verdad de las premisas asegura la verdad de la conclusión. La relación entre las premisas y las conclusiones de un argumento deductivo generalmente se conoce como "consecuencia lógica". Y tiene tres características esenciales: es necesaria, formal y conocible a priori.
  - Es necesaria en el sentido de que las premisas de los argumentos deductivos válidos hacen la conclusión necesaria: es imposible que las premisas sean verdaderas y la conclusión falsa, independientemente de cualquier otra circunstancia.
  - Es formal en el sentido de que depende solo de la forma o la sintaxis de las premisas y la conclusión. Esto significa que la validez de un argumento en particular no depende de los contenidos específicos de este argumento. Si es válido, entonces cualquier

argumento con la misma forma lógica también es válido, sin importar cuán diferente sea en el nivel de sus contenidos.

- Es conocible a priori en el sentido de que no es necesario ningún conocimiento empírico del mundo para determinar si una deducción es válida. Por lo tanto, no se necesita realizar ninguna forma de investigación empírica.

Por ejemplo:

- Regla general: "Todas las bolas de la caja X son blancas".
  - Caso: "Estas bolas provienen de la caja X".
  - Resultado: "Estas bolas son blancas".
- Inductivo. Se parte de información particular para llegar a una conclusión general. Es una forma de razonamiento en que la verdad de las premisas apoya la conclusión, pero no la garantizan. La lógica inductiva estudia las maneras de medir la probabilidad de que una conclusión sea verdadera, así como las reglas para construir argumentos inductivos fuertes.

A diferencia de los razonamientos deductivos, en los razonamientos inductivos no existe acuerdo sobre cuándo considerar un argumento como válido. De este modo, se hace uso de la noción de «fuerza inductiva» que hace referencia al grado de probabilidad de que una conclusión sea verdadera cuando sus premisas son verdaderas. Así, un argumento inductivo es fuerte cuando es altamente improbable que su conclusión sea falsa si las premisas son verdaderas.

Por ejemplo:

- Caso: "Estas bolas provienen de la caja X".
- Resultado: "Estas bolas son blancas".
- Regla general: "Todas las bolas de la caja X son blancas".

Es decir, mientras que las bolas que saque de la caja X sean blancas, no hay motivo para pensar que todas las bolas de la caja X no son blancas.

- **Abductivo.** Se parte de conjeturas espontáneas de la razón. Son como destellos de comprensión. Parte de un conjunto incompleto de hechos, información y conocimiento. Construye una hipótesis extrayendo la conclusión que mejor se justifica o explica a partir de los hechos conocidos. Charles Sanders Peirce la llama una conjetura. Esa conjetura busca ser, a primera vista, la mejor explicación, o la más probable. En el caso de una deducción se obtiene una conclusión «q» de una premisa «p», mientras que el razonamiento abductivo consiste en explicar «q» mediante «p» considerando p como hipótesis explicativa. Por ejemplo, el diagnóstico médico a partir de pruebas clínicas.

Otro ejemplo:

- Resultado: "Estas bolas son blancas".
  - Regla general: "Todas las bolas de la caja X son blancas".
  - Caso: "Estas bolas provienen de la caja X".
- **Analogía.** Su resolución consiste en traer a la memoria casos del pasado, estableciendo una analogía entre las características de la situación actual y las características de situaciones anteriores. La función del pensamiento es de comparación. En este caso las experiencias pasadas conllevan a establecer una generalización que permite recordar métodos para resolver problemas actuales.

El razonamiento por analogía es un tipo especial de razonamiento inductivo según el cual se utilizan similitudes perceptibles como base para inferir alguna otra similitud que aún no se puede percibir. El proceso de inferencia por analogía conlleva darse cuenta de las propiedades comunes de dos o más cosas y, partiendo de esta base, inferir que también comparten alguna propiedad más.

Por ejemplo:

- P y Q son similares respecto a las propiedades a, b y c.

- Se ha observado que P tiene otra propiedad x.
- Por lo tanto, Q posiblemente también tenga la propiedad x.

Por supuesto, este razonamiento no afirma que las dos cosas sean idénticas, solo que son parecidas, y puede proporcionarnos pruebas fehacientes de la conclusión, pero esta conclusión no es producto de una cuestión de necesidad lógica. Determinar la firmeza del razonamiento requiere tener en cuenta algo más que la forma: el contenido también debe ser objeto de examen.

- ...

Otra forma de clasificar los tipos de razonamiento sería la siguiente:

- Monótono, una vez se decide una conclusión, esta permanece fija e invariable, aunque se agregue nueva información a la ya existente.
- No monótono, tras añadir nueva información a la ya existente se pueden invalidar conclusiones previas. Es del tipo del razonamiento basado en el sentido común.

Como ya se ha dicho, a la hora de resolver problemas son necesarios disponer de **datos**. Y debemos recordar siempre que:

- ¡Sin datos no hay inteligencia artificial!
- ¡Sin datos suficientemente buenos no hay inteligencia artificial!
- ¡Siempre es preferible la calidad a la cantidad!

Los datos se pueden clasificar de maneras diversas, en nuestro caso se va a optar por la siguiente:

- Los datos estructurados, vienen usualmente presentados en forma de tabla, es decir dentro de una estructura predefinida, de ahí su nombre.
- Los datos no estructurados. Como por ejemplo el texto escrito o datos que provienen de procesos perceptuales, es decir cuando queremos analizar imágenes, videos o audio.

## La inteligencia artificial

En estos momentos es un área multidisciplinaria que, a través de ramas de conocimiento como las ciencias de la computación, la matemática, la lógica, la neurociencia, la filosofía, las ciencias sociales... estudia la creación y el diseño de sistemas capaces de resolver problemas cotidianos por sí mismos, utilizando como paradigma la inteligencia humana.

De partida, vamos a decir que nos fijamos en cómo **pensamos**, o **funcionamos**, o **interactuamos** y nos elegimos como modelo, patrón o ejemplo para construir sistemas con algoritmos inteligentes.

Lo que se pretende construir son sistemas con las siguientes propiedades características:

- **Autonomía:** Entendida como la capacidad para ejecutar tareas en situaciones complejas sin la dirección constante del usuario.
- **Adaptabilidad:** Entendida como la capacidad para mejorar la ejecución de las tareas aprendiendo de la experiencia.
- **Transversalidad:** Entendida como ideas que tienen un amplio conjunto de aplicaciones en todos los ámbitos del conocimiento humano. Vendría a ser como "el bálsamo de fierabrás" del Quijote, que servía para curar cualquier tipo de enfermedad.

Pero hay que distinguir lo que se puede pretender buscar, es decir cuál es el comportamiento del sistema con inteligencia artificial que se busca. Normalmente, se investigan sistemas que sean útiles para resolver problemas prácticos sin preocuparnos demasiado por si son inteligentes o simplemente se comportan como si lo fuesen. Con base en esta última afirmación se podrían distinguir dos tipos de sistemas:

- **IA general o IA fuerte**, los sistemas verdaderamente inteligentes (un deseo).
- **IA estrecha o IA débil**, los sistemas que se comportan como si fueran inteligentes, sin serlo (la realidad actual).

Para aclarar un poco más esa diferencia hagamos la siguiente analogía. Si repasan sus conocimientos literarios, cinematográficos, científicos y tecnológicos, enseguida llegarán a la conclusión de que, al ser humano, de siempre, le hubiese gustado volar.

La única solución que el ser humano tiene es construir un modelo de lo que es volar, y como muchas otras veces, dirige su mirada hacia la Naturaleza. En ella observamos que existen unos animales con capacidad de vuelo que denominamos pájaros. Si se analiza su estructura se deduce que elementos importantes para volar son las alas, que se baten y las plumas.

A partir de ahí, hemos ido evolucionando artefactos con capacidad de vuelo, pero el que se lleva la palma es el avión. Observen que, en este caso, los aviones tienen alas, pero no las baten, y no tienen plumas. Lo que tienen son reactores que impulsan al avión en el aire, la forma diferente del ala por debajo y por encima, y unos dispositivos denominados alerones que permiten cambiar la geometría de las alas. Cuando el avión es impulsado, la velocidad del aire por debajo del ala es diferente de la velocidad del aire por encima. Dependiendo de la colocación de los alerones, se puede conseguir que la presión por debajo sea mayor que la presión por encima de las alas, por lo que el avión asciende, o se puede conseguir que la presión por encima sea mayor que la presión por debajo, por lo que el avión desciende.

Conclusión se parte de la observación de los pájaros y con el paso del tiempo se llega a la construcción de un avión. Los dos vuelan, pero las soluciones son diferentes.

Volvamos al tema que nos ocupa en este apartado que son los sistemas con inteligencia artificial. Todos ellos se basan en dos ideas, la existencia de los computadores y la existencia de los algoritmos que son lo que le dice al computador lo que tiene que hacer.

¿Qué es un algoritmo? Un algoritmo es un conjunto prescrito de instrucciones, bien definidas, ordenadas de la manera adecuada, y finitas. Que, si se ejecutan, convierten los datos de un problema en una solución, es decir, recordemos que los algoritmos resuelven problemas.

Un ejemplo de algoritmo cuyo ejecutor es un ser humano, es utilizar una receta de cocina para conseguir construir un plato determinado sin ser cocinero. Observe que una buena receta cumple con los requisitos que acabamos de indicar.

Una receta es un conjunto prescrito de instrucciones, bien definidas, ordenadas de la manera adecuada, y finitas. Que, si se ejecutan, convierten los condimentos en un plato.

Ahora bien, la pregunta inmediata dentro de este contexto es ¿Qué es un algoritmo inteligente? Se afirma que un algoritmo es inteligente, cuando es ejecutado por un sistema artificial capaz de seguir las instrucciones. Y se comporta de tal manera que esa tarea de ser realizada por humanos, se diría que se requiere inteligencia.

Un ejemplo podría ser un coche autónomo que funcione sin conductor. Esa máquina, a parte de los elementos mecánicos, no es más que un conjunto de computadores dedicados y redundantes que gracias a sus diferentes algoritmos, son capaces de

realizar ciertas tareas como: ver, interpretar, tomar decisiones, acelerar, girar, frenar, encender las luces, ... Obsérvese que su comportamiento, en el caso de buen funcionamiento, hace lo mismo que haría un ser humano que condujera dicho automóvil.

Ahora háganse la siguiente pregunta, ¿han visto algún animal diferente de un ser humano con la capacidad de conducir un coche? La respuesta es que no. ¿Por qué? Porque para conducir un coche hace falta un ser humano que es inteligente. Por lo tanto, si el coche autónomo hace lo mismo que un ser humano por que es inteligente, el coche autónomo también es inteligente a la hora de realizar esa tarea.

Es evidente que ambas inteligencias son diferentes, la del ser humano es general, ya que sabe hacer más cosas con su inteligencia, y el coche autónomo sólo sabe conducir. Pero desde el punto de vista de la conducción, ambos son inteligentes.

## ¿De dónde podemos extraer modelos para la inteligencia artificial?

Como ya se ha dicho previamente, para generar modelos que nos puedan servir para construir sistemas de inteligencia artificial podemos fijarnos en el funcionamiento con los sistemas naturales que manifiestan inteligencia, por ejemplo, en nosotros, lo que significa en intentar comprender cómo **pensamos**, o cómo **funcionamos** o cómo **interactuamos** nosotros con el mundo. Además, también podemos fijarnos en cómo **actúa** la Naturaleza para crear la inteligencia, es decir fijarnos en su proceso evolutivo.

Pasemos a describir mejor cada uno de esos senderos a transitar o mejor expresado, analicemos esos paradigmas

- Tipo a). Recoge las técnicas que se fijan en los procesos de **pensamiento** desarrollados y formalizados durante el proceso histórico del ser humano. Hay tres tipos:
  - Las técnicas simbólicas
  - Las técnicas bayesianas
  - Las técnicas analogistas
- Tipo b). Las técnicas que se fijan en cómo **funciona** el proceso evolutivo de la vida en la Tierra:
  - Las técnicas evolucionistas
- Tipo c). Las técnicas que se fijan en cómo **funcionamos**, es decir intentan reproducir la estructura del cerebro.
  - Las técnicas conexionistas
- Tipo d). Las técnicas que se fijan en como **interactuamos** con el exterior.
  - Las técnicas de refuerzo



Antes de seguir, me gustaría hacer dos comentarios:

- En realidad, la inteligencia artificial intenta traspasar a las máquinas los modelos que surgen de los procesos de toma de decisiones del ser humano generados por su inteligencia natural.

¡Pero sigue habiendo diferencia de capacidades!

Como, por ejemplo, el sentido común, la capacidad de generalización, la consciencia, ...

- Cada una de las seis tribus que acogen esos planteamientos, tiene un conjunto de creencias básicas, y un conjunto de problemas particulares por los que se preocupa más. De hecho, no hay una técnica universal para crear sistemas inteligentes.

Ahora empecemos a describir con un cierto detalle cada uno de esos 6 paradigmas.

## Simbolistas

El paradigma simbólico intenta construir sistemas inteligentes que imiten a los humanos en lo que a pensamiento racional se refiere.

Son máquinas que tienen, en la lógica, el pilar de su funcionamiento. A través de cálculos, razonan y después actúan.

La inteligencia artificial simbólica es el nombre colectivo para todos los métodos de la inteligencia artificial que se basan en representaciones de alto nivel "simbólico" de los problemas, por ejemplo; la lógica matemática y la búsqueda.

El paradigma simbólico es el enfoque bajo el que se comenzó a desarrollar la inteligencia artificial, en el que la inteligencia se considera la resultante de agrupar conceptos para formar estructuras, las cuales se manipulan para producir enunciados. La técnica, entonces, se centra en la sistematización formal de ese proceso.

Este grupo, piensa que toda la inteligencia se puede reducir a manipular símbolos, es decir, opera con representaciones abstractas de la información, de la misma manera que los matemáticos resuelven ecuaciones, reemplazando expresiones por otras expresiones. Para automatizar el proceso tan sólo se requiere tener a disposición una máquina de Turing universal: La tribu considera que puede hacer cualquier cosa, incluso piensan que se puede aprender partiendo de cero.

Su algoritmo principal es la deducción, que sirve para descubrir el conocimiento que falta para poder alcanzar la solución de un problema del que se tienen datos e información.

Lo que caracteriza a este paradigma es:

- Los diseñadores programan explícitamente todo el conocimiento
- Utilizan símbolos y variables
- Fortaleza: En el trabajo con problemas lógicos, es fácil de saber qué hace bien o mal, arreglar lo que no funciona y conseguir confianza en los resultados.

- Debilidades: Necesita que los datos sean completamente deterministas. Tiene problemas para tratar con la incertidumbre, y los conceptos reales suelen ser difíciles de definir mediante reglas lógicas concisas.

Ejemplo 1: El siguiente razonamiento es correcto

- Se sabe que
  - Los planetas son esféricos
- Si se verifica que
  - La Tierra es un planeta
- entonces
  - La Tierra es esférica
- Por lo tanto, es cierto que
  - La Tierra es esférica

Ejemplo 2: El siguiente razonamiento también es correcto

- Se sabe que
  - anvyoinvst
- Si se verifica que
  - exxnnnttnn
- entonces
  - hgbiugvy
- Por lo tanto, es cierto que
  - hgbiugvy

Ejemplo 3: El siguiente razonamiento también es correcto

- p
- $p \rightarrow q$
- 
- ∴ q

Comentario: Los tres ejemplos simbólicamente se pueden representar de la misma manera, lo habitual es el tercero. ¿Entiende ahora el porqué del nombre de paradigma simbólico?

Si formalizamos un poco más podemos dar las siguientes definiciones para este tipo de paradigma.

- Un sistema formal o sistema lógico es un sistema abstracto compuesto por un lenguaje formal, axiomas, reglas de inferencia y a veces una semántica formal que se utiliza para deducir o demostrar teoremas y dar una definición rigurosa del concepto de demostración.
  - Lenguaje formal: es un lenguaje cuyos símbolos (alfabeto) son primitivos y las reglas para unir esos símbolos (sintaxis) están formalmente especificadas. A una cadena de símbolos formada de acuerdo a la gramática se le llama una fórmula bien formada.
  - Axioma: es una proposición tan clara y evidente que se admite sin demostración.
  - Reglas de inferencia o reglas de transformación: es una función que toma premisas, analiza su sintaxis y devuelve una conclusión. Una regla de este tipo conserva la verdad, que es una propiedad semántica. La acción de la regla de inferencia es puramente sintáctica.
  - Semántica formal: Una interpretación de un lenguaje formal es básicamente una asignación de significados a sus símbolos, y de condiciones de verdad a sus fórmulas bien formadas. Las interpretaciones formales asignan significados inequívocos a los símbolos primitivos, y valores de verdad a las fórmulas.

Las cualidades fundamentales de un sistema lógico son:

- La **consistencia** es la propiedad que tienen los sistemas formales cuando no es posible deducir una contradicción dentro del sistema. Es decir, no es posible deducir una fórmula y su negación.
- La **decidibilidad**: Cuando para cualquier fórmula en el lenguaje del sistema, existe un método efectivo para determinar si esa fórmula pertenece o no al conjunto de las verdades del sistema. Es decir, en un número finito de pasos puede decidir si la fórmula es válida o no. Cuando una fórmula no puede ser probada verdadera ni falsa, se dice que la fórmula es independiente, y por lo tanto el sistema no es decidible.

- La **completitud** es la propiedad que tienen los sistemas formales cuando todas las fórmulas lógicamente válidas del sistema pertenecen al sistema lógico. O lo que es lo mismo, si un algoritmo puede derivar cualquier consecuencia posible.

No es habitual encontrar sistemas lógicos que verifiquen las tres propiedades a la vez.

A lo largo de la evolución histórica de la lógica, se han ido desarrollando diferentes tipos de lógicas, aquí tan sólo citaremos algunas de ellas:

- Lógicas clásicas, como la lógica proposicional, la lógica de primer orden y la lógica de segundo orden.
- Lógicas modales como la deóntica, la temporal, la epistémica y la doxástica.
- Lógicas no clásicas como la difusa, la cuántica, la no monotónica, la intuicionista, ...
- Programación lógica, programación lógica inductiva, programación lógica abductiva.

Ejemplo: Ver ( [https://es.wikipedia.org/wiki/Prueba\\_ontológica\\_de\\_Gödel](https://es.wikipedia.org/wiki/Prueba_ontológica_de_Gödel) ) Allí se puede encontrar la prueba siguiendo argumentos formales para la existencia de Dios propuesto por el matemático Kurt Gödel (1906-1978). Utiliza lógica modal. Si los axiomas en los que se basa fueran ciertos, la conclusión sería cierta. El problema surge cuando se analizan los axiomas, con los que algunas personas pueden no estar de acuerdo que sean ciertos.

Comentarios finales sobre este tipo de paradigma cuando se utiliza para ayuda a la toma de decisiones:

- Aunque tiene varias ventajas, como la precisión, la explicabilidad y la adaptabilidad, también tiene varias desventajas, como la limitación por la calidad (exactitud) y la cantidad de los datos de entrada y la complejidad de los problemas que se intentan resolver.
- Si los problemas son demasiado complejos, existe la posibilidad de que no se pueden modelar fácilmente en términos de reglas lógicas, y se tiene el posible límite del nivel de formación o la capacidad humana

para crear reglas lógicas o un excesivo requerimiento de una gran cantidad de tiempo y de recursos.

Un subconjunto del paradigma simbólico se conoce como algoritmos de búsqueda para la resolución de problemas y planificación. Se caracterizan por el hecho de que el conocimiento que se tiene sobre un problema y todas sus posibilidades puede representarse mediante un espacio de búsqueda estructurado, que suele ser un grafo o un árbol. Encontrar una solución a una pregunta relacionada con ese conocimiento, va a consistir en recorrer la estructura hasta encontrar lo que se puede considerar como una solución.

La formalización de este tipo de problemas se realiza del modo siguiente:

- Hay que conocer lo que se denomina el espacio de estados, que se refiere al conjunto de situaciones posibles.
- Las transiciones son los movimientos posibles entre un estado y otro (acciones directas y que puedan ejecutarse con una única acción).
- Los costes no todas las transiciones son similares, pueden diferir en aspectos que hagan que unas sean preferibles a otras.
- Definición del estado inicial.
- Definición del objetivo a alcanzar.

Para que se hagan una idea, la utilización de *Google Maps* para ir de una zona geográfica de una ciudad a otra zona geográfica, sería un problema de búsqueda.

- El espacio de estados estaría formado por los extremos de las calles.
- Las transiciones serían los movimientos posibles entre un estado y otro, que no son las mismas para una persona caminando, circulando con un coche o utilizando un medio público de transporte.
- Los costes, dependiendo del procedimiento seguido para circular, son o los tiempos necesarios para moverse entre un estado u otro permitido, o bien la distancia recorrida.
- Hay que conocer el punto de partida y el punto de llegada para obtener la secuencia de acciones que nos lleva a una solución.

Ejemplo, ver: (<https://naukas.com/2015/10/05/maridos-celosos-misioneros-zombis/>)

- En él se puede encontrar la solución al siguiente problema: Tres señores y tres criados deben cruzar el río en una barca en la que se permite que viajen una o dos personas. Solo hay un problema, los caballeros tienen motivos para sospechar que los sirvientes han conspirado entre ellos para robar a los señores y matarlos. Por lo tanto, los caballeros son conscientes de que en ningún momento puede haber más criados que señores en alguna de las orillas, puesto que los matarían. ¿Cómo podrían cruzar el río de la manera más eficiente?

Los algoritmos de búsqueda pueden utilizar diferentes estrategias de búsqueda de una solución. A menudo hay varias formas diferentes de resolver uno de esos problemas, aunque algunas son preferibles desde el punto de vista del tiempo, el esfuerzo, el coste u otros criterios. Además, dependiendo del tamaño del problema, diferentes técnicas de búsqueda pueden llevar a diferentes soluciones.

A modo de información sobre los nombres que reciben esas posibles estrategias podríamos hablar de:

- Procesos de búsquedas sistemáticas
  - Búsquedas ciegas o no informadas.
    - Búsqueda en anchura
    - Búsqueda de coste uniforme
    - Búsqueda primero en profundidad
    - Búsqueda en profundidad limitada
    - Búsqueda en profundidad iterativa
    - Búsqueda bidireccional
    - ...
  - Búsquedas informadas mediante heurísticas.
    - Búsqueda primero el mejor
    - Búsqueda voraz
    - Búsqueda primero el mejor A\*
    - Búsqueda primero el mejor IDA\*
    - Búsqueda recursiva primero el mejor
    - Búsqueda SMA\*
    - ...

- Procesos de búsquedas locales
  - Ascenso de colinas y sus variaciones
  - Temple simulado.
  - Algoritmos genéticos.
  - Optimización mediante enjambres de puntos.
  - Optimización por colonias de hormigas.
  - Gota de agua inteligente.
  - Algoritmo de búsqueda gravitacional.
  - ...



## Bayesianos

Las limitaciones de la resolución de problemas basados en datos, exige, cuando nos enfrentamos a problemas de gran complejidad, un esfuerzo por generar nuevos modelos y además disponer de una gran cantidad de datos.

Pero a veces no siempre es posible encontrar tales condiciones. Es por ello, que el enfoque probabilístico es una alternativa al no estar limitado por la cantidad de datos y permitir la generación de modelos estadísticos mediante programación probabilística.

La teoría de la probabilidad es la rama de las matemáticas que estudia fenómenos aleatorios, es decir aquellos cuyo resultado es, a priori, impredecible, por ejemplo, el lanzamiento de una moneda. También se ocupa de los llamados procesos estocásticos, es decir, aquellos sobre los que existe una secuencia cambiante de eventos que, en principio, no son deterministas, por ejemplo, el movimiento de un grano de polen en el agua.

En lugar de información perfecta, nos solemos encontrar un abanico de posibilidades desconocidas, desde la falta de información (ruido = errores e imprecisiones) hasta el engaño deliberado.

Los métodos basados en la probabilidad tienen la capacidad de hacer frente a la incertidumbre, en otras palabras, la lección clave es que la incertidumbre no está fuera del alcance del pensamiento y la discusión racionales, y la probabilidad ofrece una forma sistemática de abordar dicho problema.

La lección más importante sobre la probabilidad no es el cálculo de probabilidades, sino la capacidad de ver la incertidumbre como algo que, al menos en principio, se puede cuantificar. Esto quiere decir que podemos hablar de la incertidumbre como si fuera un número: los números se pueden comparar («esto es más probable que esto otro») y, a menudo, medir.

Pero la flexibilidad de la probabilidad y de la estadística tiene el coste de la complejidad computacional (en el caso de los modelos multiparamétricos). Si un problema depende de muchos parámetros, ese problema puede tomar un número de posibilidades diferentes que haga intratable el problema.

- Veamos un ejemplo de explosión combinatoria:
  - Se define un cuadrado latino de orden  $n$  a una matriz  $n \times n$  cuyos elementos pertenecen a un conjunto de  $n$  elementos con la propiedad que cada elemento del conjunto aparece sólo una vez en cada fila y cada columna de la propia matriz. El siguiente es un ejemplo de cuadrado latino de orden tres:

1	2	3
2	3	1
3	1	2

Existen 12 cuadrados latinos diferentes que se pueden construir con los valores 1, 2, y 3.

- Ahora observe el número de cuadrados latinos que pueden construir cuando hacemos varias el número de valores. Sólo hemos llegado hasta  $n = 11$ , pero el resultado es desorbitante.

$n$	El número de cuadrados latinos de orden $n$
1	1
2	2
3	12
4	576
5	161,280
6	812,851,200
7	61,479,419,904,000
8	108,776,032,459,082,956,800
9	5,524,751,496,156,892,842,531,225,600
10	9,982,437,658,213,039,871,725,064,756,920,320,000
11	776,966,836,171,770,144,107,444,346,734,230,682,311,065,600,000
12	...

Las decisiones que los humanos consideramos racionales están basadas en inferencias consistentes con nuestra experiencia o con un conocimiento adquirido sobre el comportamiento de fenómenos científicos deterministas. Sin embargo, cuando los fenómenos que observamos cambian de una manera aleatoria y con múltiples estados, como la propagación de un fuego avivado por el viento, nos es mucho más difícil realizar tales inferencias. Entonces no podemos echar mano de la experiencia sino de nuestro conocimiento basado en modelos.

La ventaja del enfoque probabilístico es que esos diferentes tipos de incertidumbre se pueden representar bajo los mismos axiomas y reglas. Cuando este enfoque es utilizado en el Aprendizaje Máquina se habla de Aprendizaje Bayesiano.

Dicho aprendizaje consiste en transformar las distribuciones de probabilidad definidas antes de observar los datos en distribuciones condicionadas por dichas observaciones. Además de por su simplicidad conceptual, el aprendizaje bayesiano permite su representación mediante grafos, lo que facilita una comprensión más intuitiva y visual de los modelos.

Otra característica es su flexibilidad en la composición de modelos más complejos para muchos datos a partir de modelos simples probabilísticos de pocos datos.

Medir probabilidades es difícil, pues lo normal es que haga falta observar repetidamente un fenómeno para extraer conclusiones.

No obstante, mediante la recogida sistemática de datos, podemos valorar de forma crítica afirmaciones probabilísticas, y, a veces, podemos comprobar si nuestros números son correctos o incorrectos.

En el caso de los sistemas de inteligencia artificial, las soluciones se obtienen mediante inferencia probabilística y el algoritmo base es el teorema de Bayes y sus derivados.

El teorema de Bayes es una sencilla regla para actualizar su grado de creencia en una hipótesis cuando usted recibe una evidencia nueva.

Si la evidencia es consistente con la hipótesis, la probabilidad de la hipótesis crece, si no, baja.

Enunciemos dos teoremas fundamentales:

- Teorema de la probabilidad total.

Dados  $n$  sucesos diferentes  $B_i$  con  $i = 1 \dots n$ , para cualquier suceso  $A$  como el de la figura se cumple:

Dados  $n$  sucesos "distintos"  $B_1, \dots, B_n$ , para cualquier suceso  $A$  como el de la figura se cumple:

$$P(A) = \sum_{i=1}^n P(A|B_i) \cdot P(B_i)$$

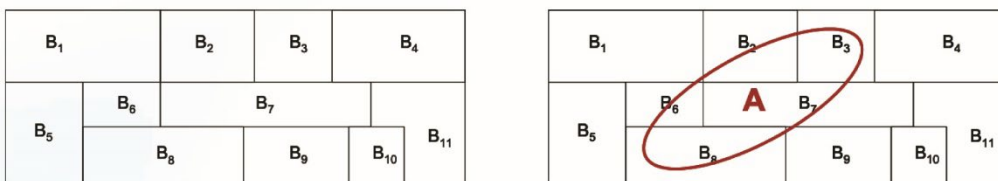


Imagen de Alberto Turón, 2024

En dicha imagen se representan las probabilidades de cada suceso  $B_i$  mediante el área que ocupa, entendiendo que el área total es 1.

- Teorema de Bayes

Dados  $n$  sucesos "distintos"  $B_1, \dots, B_n$ , para cualquier suceso  $A$  como el de la figura se cumple:

$$P(B_i|A) = \frac{P(A|B_i) \cdot P(B_i)}{\sum_{j=1}^n P(A|B_j) \cdot P(B_j)}$$

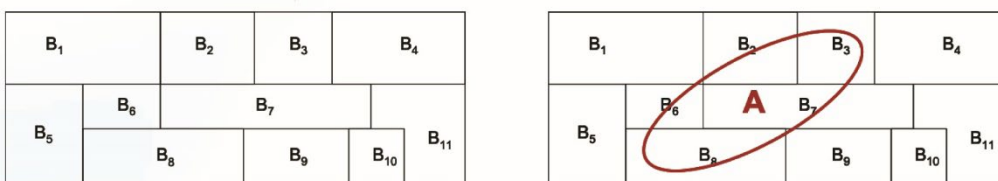


Imagen de Alberto Turón, 2024



¿Cuál es la probabilidad de que el color del taxi que se vio involucrado fuera realmente azul?

La respuesta la vamos a obtener mediante el teorema de Bayes.

- La probabilidad de partida de que el taxi envuelto sea el azul sabemos que es  $P(\text{azul}) = 15/100 = 0.15$
- La probabilidad de partida de que el taxi envuelto sea el verde sabemos que es  $P(\text{verde}) = 85/100 = 0.85$
- El testigo dice que el taxi involucrado era azul, por lo tanto
  - $P(\text{de que el testigo diga que es azul}) = P(T/\text{Azul}) = 80/100 = 0.8$
  - $P(\text{de que el testigo diga que es verde}) = P(T/\text{verde}) = 20/100 = 0.2$
- Aplicando el teorema de Bayes
  - $P(\text{azul}|T) = [P(T|\text{azul}) P(\text{azul})] / [P(T|\text{azul}) P(\text{azul}) + P(T|\text{verde}) P(\text{verde})] = 0.41$
  - $P(\text{verde}|T) = 1 - 0.41 = 0.59$

Otra familia son los algoritmos probabilísticos que intentan construir una distribución de probabilidades de las características de los datos para realizar diferentes tipos de tareas.

Si se tienen muchas evidencias, tal como el resultado de múltiples test. Para combinarlos todos sin sufrir una explosión combinatoria es necesario hacer suposiciones simplificadoras. Los Bayesianos han inventado modelos diabólicamente inteligentes.

Los algoritmos más representativos son:

- Naive Bayes
- Redes Bayesianas
- Adenas de Markov y Cadenas ocultas de Markov
- Modelos de mezcla gaussiana
- ...

La diferencia entre los distintos métodos bayesianos, modelos causales y redes bayesianas, estriba en las hipótesis de independencia condicional entre hipótesis y evidencias. Dichas relaciones se expresan comúnmente mediante un grafo acíclico dirigido.

## Comentarios:

- Redes bayesianas.

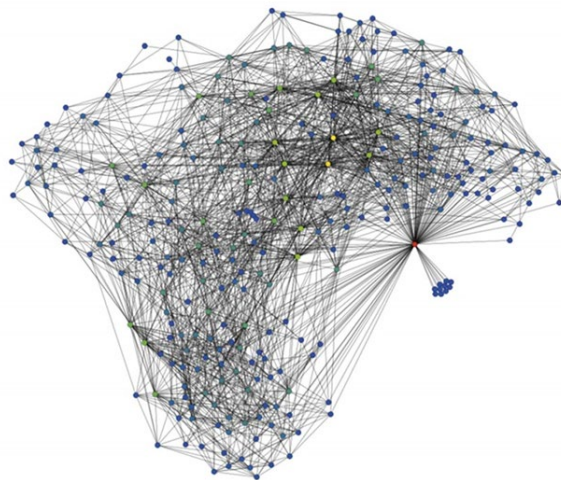
En teoría de la probabilidad, no hay ninguna forma *a priori* de conocer qué variables influyen sobre otras. En general, las distribuciones de probabilidad conjuntas se deben conocer para poder realizar una inferencia correcta. Pero a la hora de construir modelos para el mundo real, si se tuvieran en cuenta todas las posibilidades, esas distribuciones serían muy grandes y en los casos reales no se podrían almacenar en un computador.

Una red bayesiana es un modelo grafo probabilístico (un tipo de modelo estático) que representa un conjunto de variables aleatorias (nodos de un grafo) y sus dependencias condicionales (las aristas) a través de un grafo acíclico dirigido.

Los nodos que no se encuentran conectados representan variables las cuales son condicionalmente independientes de las otras.

Cada nodo tiene asociado una función de probabilidad que toma como entrada un conjunto particular de valores de las variables padres del nodo y devuelve la probabilidad de la variable representada por el nodo.

Observe con detenimiento la siguiente red bayesiana. Una vez establecidas las probabilidades a priori, supóngase que viene nueva información a algún nodo, ¿cómo se modificarían las probabilidades de la red?



- Cadenas de Márkov

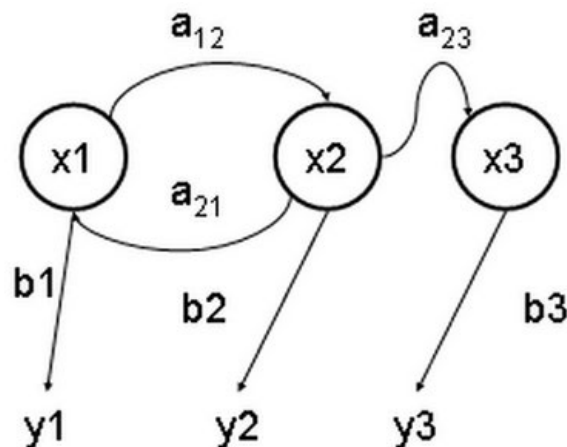
En la teoría de la probabilidad, se conoce como cadena de Márkov o modelo de Márkov a un tipo especial de proceso estocástico discreto en el que la probabilidad de que ocurra un evento depende solamente del evento inmediatamente anterior. Esta característica de incluir una memoria reciente recibe el nombre de propiedad de Márkov en contraste con los eventos independientes que no tienen memoria de ningún evento anterior.

Se dice que una Cadena de Markov es homogénea si la probabilidad de ir del estado "i" al estado "j" en un paso no depende del tiempo en el que se encuentra la cadena, en caso contrario se denomina no homogénea.

- Modelo oculto de Márkov

Un modelo oculto de Márkov es un modelo estadístico en el que se asume que el sistema a modelar es un proceso de Márkov de parámetros desconocidos. El objetivo es determinar los parámetros desconocidos (u ocultos, de ahí el nombre) de dicha cadena a partir de los parámetros observables.

Ejemplo de caracterización de un modelo oculto de Márkov



- $x$  — estados ocultos
- $y$  — salidas observables
- $a$  — probabilidades de transición
- $b$  — probabilidades de salida



## Analogistas

La esencia de la analogía es que dos cosas son similares si ellas coinciden en algunos aspectos. Si eso es así, puede que coincidan en otros.

La clave es aprender a reconocer similitudes entre situaciones y de ese modo inferir otras similitudes. El problema clave es como juzgar la similitud entre las cosas.

Los algoritmos de aprendizaje basado en instancias, almacenan instancias (ejemplos) de los datos de entrenamiento para posteriormente comparar el dato procesado con dichas instancias.

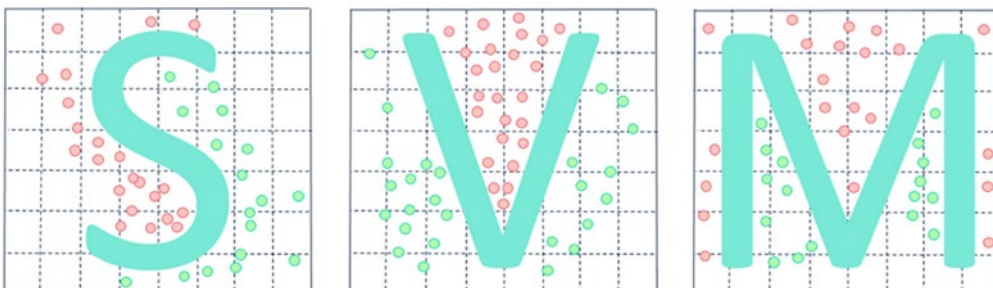
Los algoritmos más usados en esta familia son los:

- *k-vecinos* más cercanos.
- Las *máquinas de soporte vectorial (SVM)*
  - Existen algoritmos de SVM lineales y no lineales.

Comentarios:

- Intuitivamente, una SVM es un modelo que representa a los puntos de muestra en el espacio, separando las clases a 2 espacios lo más amplios posibles mediante un hiperplano de separación definido como el vector entre los 2 puntos, de las 2 clases, más cercanos al que se llama vector soporte. Cuando las nuevas muestras se ponen en correspondencia con dicho modelo, en función de los espacios a los que pertenezcan, pueden ser clasificadas a una o la otra clase.

Ejemplos de resultados obtenidos mediante máquinas de soporte vectorial



- Los SVM son particularmente efectivos en espacios de grandes dimensiones. También existe una amplia variedad de funciones del kernel que permiten separar los datos de diferentes maneras. El tiempo de entrenamiento para una SVM también puede ser muy alto, dependiendo del conjunto de datos
- Entrenar una Máquina de Soporte Vectorial (SVM) implica varios pasos clave. Un resumen del proceso podría ser:
  - Preparación de los Datos: Antes de entrenar una SVM, es crucial preparar los datos. Esto incluye la normalización o estandarización de los datos para asegurar que todas las características tengan la misma escala.
  - Selección del Kernel: El kernel es una función matemática que transforma los datos de entrada a un espacio de características de mayor dimensión. Los kernels más comunes son el lineal, polinómico y el de función de base radial (RBF).
  - Entrenamiento del Modelo: Durante este paso, la SVM ajusta los parámetros del hiperplano de separación para maximizar el margen entre las clases. Esto se hace resolviendo un problema de optimización.
  - Ajuste de Hiperparámetros: Los hiperparámetros, como el parámetro de regularización ( $C$ ) y el parámetro del kernel ( $\gamma$ ), se ajustan utilizando técnicas como la validación cruzada para mejorar el rendimiento del modelo.
  - Evaluación del Modelo: Finalmente, se evalúa el modelo utilizando un conjunto de datos de prueba para asegurar que generaliza bien a datos no vistos.

## Evolutivistas

Los conceptos básicos de este paradigma son:

- Creen que la madre de todo el aprendizaje es la selección natural. Si esto es así, se puede hacer cualquier cosa, todo lo que es necesario es simularla en un computador.
- Se utiliza la biología evolutiva seguida por los organismos complejos.

Tipos de algoritmos evolutivos:

- Genéticos
- Otros algoritmos de optimización bioinspirada
  - Enjambres de partículas
  - Colonias de hormigas
  - Comportamiento de rebaños
  - Crecimiento bacteriano
  - Bandadas y cardúmenes

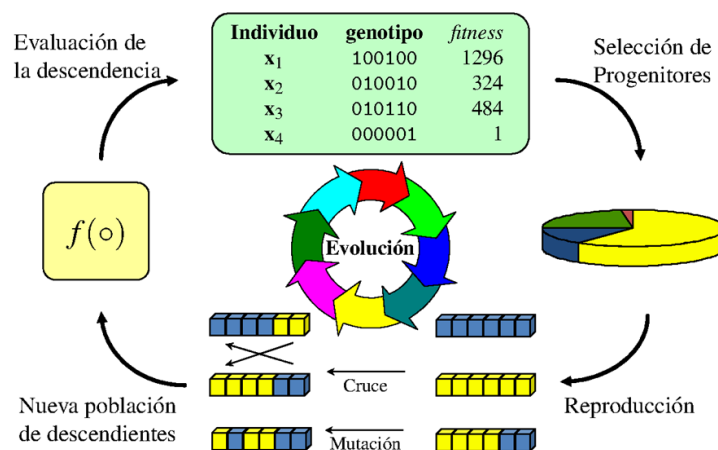
Genéticos:

- Siguiendo la terminología de la teoría de la evolución, las entidades que representan las soluciones al problema se denominan individuos o cromosomas, y el conjunto de estos, población.
- Los individuos son modificados por operadores genéticos, principalmente:
  - El cruce, que consiste en la mezcla de la información de dos o más individuos.
  - La mutación, que es un cambio aleatorio en los individuos.

- La selección, consistente en la elección de los individuos que sobrevivirán y conformarán la siguiente generación.
- De acuerdo con algún criterio, se decide cuáles son los individuos más adaptados, que sobreviven, y cuáles son los menos aptos, que son descartados.

Dado que los individuos que representan las soluciones más adecuadas al problema tienen más posibilidades de sobrevivir, la población va mejorando gradualmente.

- El algoritmo básico completo sería:
  - Generar aleatoriamente una población inicial
  - Calcular la aptitud de cada individuo
    - Puede estar sometido a restricciones
  - Seleccionar probabilísticamente con base en aptitud
  - Aplicar operadores genéticos (cruce y mutación) para generar la siguiente población
    - Los cruces y las mutaciones pueden estar sometidos a satisfacer probabilidades.
  - Repetir todo hasta que cierta condición se satisfaga
- Diagrama:



- Este tipo de algoritmos, no son triviales de implementar, se requiere pericia para la definición de la función de evaluación y la selección de los criterios de mutación entre otros.
- Se utilizan principalmente en problemas con espacios de búsqueda extensos y no lineales, en donde otros métodos no son capaces de encontrar soluciones en un tiempo razonable.
- Deben tenerse en cuenta también las siguientes consideraciones:
  - Si la función a optimizar tiene muchos máximos/mínimos locales se requerirán más iteraciones del algoritmo para "asegurar" el máximo/mínimo global.
  - Si la función a optimizar contiene varios puntos muy cercanos en valor al óptimo, solamente podemos "asegurar" que encontraremos uno de ellos (no necesariamente el óptimo)
  - La alta complejidad requiere funciones de evaluación demasiado costosas.
  - Converge asintóticamente.
  - La mejor solución encontrada puede que no sea la mejor de todas.

#### Algoritmos de comportamiento colectivo

- Enjambres de partículas
  - Es la técnica que estudia el comportamiento colectivo de los sistemas descentralizados, autoorganizados, naturales o artificiales.
  - Están típicamente formados por una población de agentes simples que interactúan localmente entre ellos y con su medio ambiente. Los agentes siguen reglas simples y, aunque no existe una estructura de control centralizado que dictamine el comportamiento de cada uno de ellos, las interacciones locales entre los agentes conducen a la emergencia de un comportamiento global complejo.

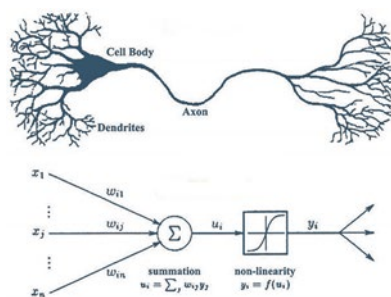
- Como ejemplos naturales se incluyen las colonias de hormigas, el alineamiento de las aves en vuelo, el comportamiento de rebaños durante el pastoreo, el crecimiento bacteriano,...

## Conexionistas

Las creencias de este paradigma son las siguientes:

- El aprendizaje es lo que el cerebro hace, por lo tanto, lo que se tiene que hacer es ingeniería inversa.
- La hipótesis es que la inteligencia emerge a partir de la actividad distribuida de un gran número de unidades interconectadas que procesan información de forma paralela. Estas unidades, por ejemplo, son modelos matemáticos ligeramente aproximados de la actividad eléctrica de las neuronas biológicas.
- Las neuronas artificiales tienen el problema de que, o bien no son modelos fieles a una neurona real, o bien solo son un modelo muy simplificado de una neurona.
- El cerebro aprende reforzando las conexiones entre neuronas, y el problema crucial es descubrir cuáles y cambiarlas en consecuencia.
- Su algoritmo principal es la propagación hacia atrás "backpropagation", que permite comparar la salida de un sistema con la salida deseada para los datos de entrada introducidos y entonces el error detectado permite realizar sucesivamente cambios en las conexiones de cada capa, propagando los errores de delante hacia atrás, de manera que la salida sea lo más próxima a la deseada.

La descripción matemática de una neurona es la siguiente:

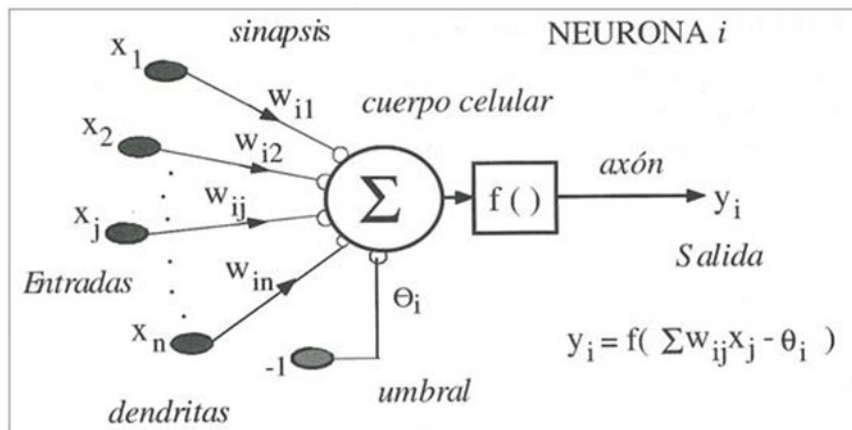


Observemos que la neurona natural tiene los siguientes elementos:

- Las dendritas que forman una cabellera de ramificaciones que recogen información de las terminales de las neuronas que la rodean.

- El cuerpo de la neurona, que recibe todas las informaciones que vienen por las dendritas, las acumula y decide si va a generar una señal de salida que se va a propagar por el axón.
- Posteriormente esa señal se distribuirá por la cabellera asociada a los terminales del axón y se comunicarán a través de las dendritas de las neuronas cercanas.

Ahora observemos la neurona artificial que denominaremos "i" que intenta replicar la geometría y el funcionamiento de la neurona natural.



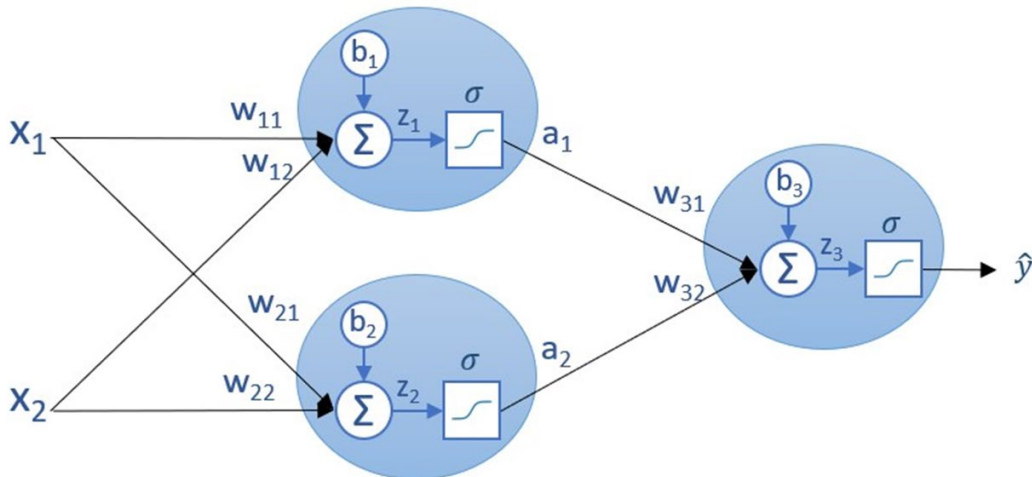
- La forma es similar, tiene las ramificaciones de entrada de la información a la neurona. El valor de cada una de las informaciones que va a recibir la neurona viene representado por los valores numéricos  $X_i$ , con  $i = 1 \dots n$ .
- La importancia de cada información recibida (todas no tienen por qué ser igual de importantes) se representa por los valores numéricos  $W_i$  con  $i = 1 \dots n$ .
- Observe la existencia de un valor denominado  $\theta_i$
- El cuerpo de la neurona recibe toda esa información y la acumula calculando una combinación lineal del modo siguiente

$$\sum_{j=1}^{j=n} w_{ij} x_j - \theta_i$$

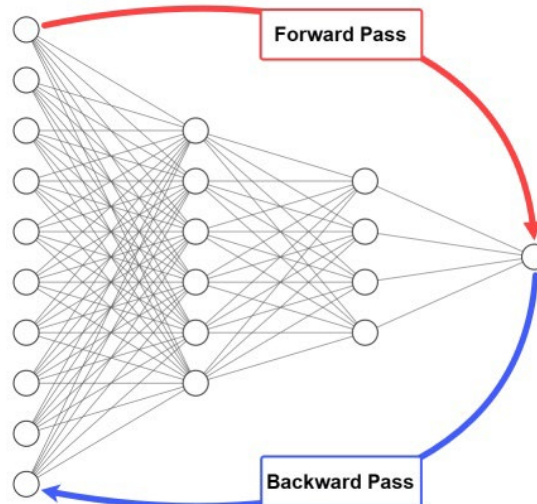


- Ese valor de salida, se pasa por una función no lineal  $f()$  cuya valor se va a corresponder con la información de salida de esa neurona y se comunicará a todas las neuronas que rodean a la neurona "i" .

En la siguiente imagen se puede ver la estructura de una red sencilla formada por dos neuronas de entrada y una de salida.

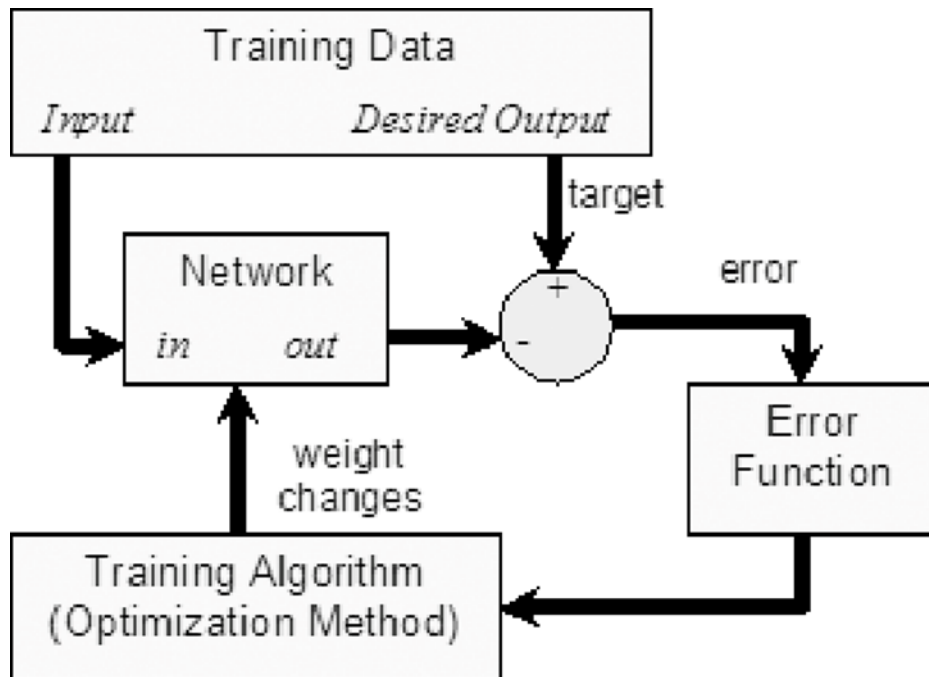


El funcionamiento de una red neuronal que tiene la capacidad de aprendizaje viene descrito en las siguientes imágenes.



- El proceso que se sigue es iterativo, y en cada iteración se realizan dos acciones.
  - El paso hacia adelante (forward pass), en el que a partir de los datos de entrada se realizan todas las operaciones de cálculo y transmisión de cada neurona.

- El paso hacia atrás (backward pass), en el que se modificarán los valores de los pesos.
- La forma de realizar el aprendizaje se puede observar en la siguiente imagen



- Caja "Training Data", que representa dos tipos de datos
  - Input, el conjunto de los datos que se van a utilizar para entrenar a la red.
  - Desired Output, el conjunto de los datos de salida que se deberían obtener a la salida de la red.
- Caja "Network", que representa a la red neuronal, que como ya sabemos recibe información "in" y genera otra información "out".
- El círculo que recibe los datos deseados "desired Output" y los datos de salida "out" de la red. Con ellos se calcula la diferencia, que es lo que se denomina "error" del cálculo que se acaba de realizar.
- Dichos valores se pasan a través de lo que se denomina la función de error, que es la ecuación que se haya definido.

- Y su salida se pasa al Algoritmo de Training (aprendizaje) que mediante un proceso de optimización (descenso del gradiente) más o menos complicado, permite recalculados los nuevos pesos ( $W$ ). A este paso de aprendizaje se suele denominar retropropagación.
- Una vez recalculados los pesos, la red ya está preparada para volver a recibir unos nuevos datos de entrada y de salida deseada.
- El proceso se itera hasta alcanza a la salida un error que se considere adecuado.

Estos sistemas aprenden y se forman a sí mismos, en lugar de ser programados de forma explícita, y sobresalen en áreas donde la detección de soluciones o características es difícil de expresar con la programación convencional.

- Sus características son:
  - Los diseñadores enseñan a una red neuronal todo lo que necesita saber.
  - Fortaleza: Puede trabajar con datos imperfectos.
  - Debilidad: Trabajar con problemas lógicos y tienen un bajo nivel de explicabilidad.
  - Estas arquitecturas requieren una gran cantidad de datos de entrenamiento (del orden de cientos de miles o millones) y son adecuadas cuando tenemos datos no estructurados.
  - La red puede aproximar cualquier función no lineal continua.
  - Puede trabajar con conceptos cualitativos y cuantitativos.
  - Puede modelar sistemas multivariados.
  - En el aprendizaje simbólico, existe una correspondencia unívoca entre los conceptos y los símbolos que los representan. En el caso de los conexionistas, la representación está distribuida. Cada concepto está representado por muchas neuronas, y cada neurona participa en la representación de muchos conceptos diferentes.

Los conceptos y su memoria están representados en el cerebro por ensamblajes de neuronas.

- Otra diferencia entre ellos es que los primeros actúan de forma secuencial y los segundos de forma paralela.
- Los principales representantes de esta familia son las redes neuronales denominadas:
  - Multicapa
  - Redes Convolucionales
  - Redes Recurrentes y Redes Long-Short Term Memory (LSTM)
  - Modelos de difusión
  - Redes Generativas Adversarias
  - Autocodificadores
  - Redes Transformer
  - ...

## Refuerzo

Las creencias de este paradigma son las siguientes:

- El cuerpo interaccionando con el mundo, da forma a la inteligencia y, por lo tanto, sin cuerpo no puede haber inteligencia de tipo general.
- Consiste en diseñar un sistema capaz de interaccionar de forma directa con un entorno y que pueda relacionar las señales que percibe mediante sensores con representaciones internas (modelos del mundo) generadas a partir de lo percibido.
- Esas interacciones conforman las habilidades cognitivas de los agentes, dando lugar a lo que se conoce como cognición situada.

Estos sistemas que actúan como humanos, no imitan nuestra forma de pensar, sino la de interactuar. No se busca que desarrollen complejas habilidades cognitivas, sino que hagan acciones mecánicas de forma más efectiva que nosotros. Los robots controlables, los robots de una cadena de montaje, los personajes virtuales, ...

Ejemplo: Supongamos que vamos a enseñar a un humano a jugar al Pong, el clásico juego de Atari.

Si le mostramos por primera vez el juego a una persona, le podríamos dar una instrucción como esta: "con el teclado puede controlar una paleta, que se mueve hacia arriba o hacia abajo. Su tarea es golpear la bola hasta lograr que su oponente no logre alcanzarla. Cada vez que haga esto obtendrá un punto. Gana el jugador que logre obtener más puntos al final de la partida".

Lentamente, interactuando con el juego, el nuevo jugador humano aprenderá a controlar la paleta, a golpear la bola y a marcar puntos, a través de un proceso de prueba y error, y eventualmente logrará vencer al oponente.

¿Y cómo lograríamos esto con un computador?

- Primero debe "entender" los elementos del juego: hay dos oponentes, un tablero, una bola, dos paletas. Es decir, debe entender el entorno.

- En segundo lugar, debe "entender" lo que está sucediendo en el entorno, por ejemplo, la dirección de la bola o los movimientos de su oponente, (el estado).
- En tercer lugar, dependiendo del estado el agente deberá "aprender" a moverse en el tablero para tratar de vencer a su oponente, (acción).
- La manera de saber si lo hizo bien o mal será a través de la puntuación. Si lo hace mal se verá penalizado, si lo hace bien recibirá un premio (recompensa negativa o positiva).

Resumiendo: la idea detrás del aprendizaje por refuerzo es que "un agente aprenderá de su entorno, mediante la observación de su estado y mediante su interacción a través de una serie de acciones por las cuales recibirá una recompensa".

En esencia hay dos maneras de hacerlo, dependiendo de si el agente conoce al detalle su entorno o sólo parcialmente.

- Aprendizaje reforzado basado en modelos. Cuando se conoce en detalle el entorno y sus reglas de juego tendremos un algoritmo de aprendizaje reforzado basado en modelos. Un ejemplo de un modelo de un entorno es precisamente el juego Go: con antelación el agente puede conocer las reglas del juego, los movimientos que puede realizar, el tamaño del tablero. Con este modelo el agente puede planear con antelación su siguiente movida, y puede analizar las implicaciones de este movimiento o elegir otras alternativas. El sistema Alphazero, desarrollado precisamente por DeepMind en 2017, es un ejemplo de un algoritmo de aprendizaje forzado basado en modelos.

El problema de este tipo de algoritmos es que sólo en contadas aplicaciones se tiene toda la información para construir un modelo del entorno. En la mayoría de las aplicaciones reales sólo se tiene acceso parcial, y en estos casos hablamos de aprendizaje reforzado libre de modelos, al que pertenecen la mayor parte de los algoritmos usados en la actualidad.

- En el caso del aprendizaje reforzado libre de modelos: el agente debe aprender a tomar las decisiones por prueba y error, pues no conoce todos los detalles del entorno. En realidad, sólo tiene acceso a dos elementos: los estados y la recompensa resultante de sus acciones. Un ejemplo de esto es precisamente la inteligencia artificial desarrollada

en 2013 por DeepMind, y que fue capaz de vencer al ser humano en varios juegos de Atari.

Los algoritmos típicos de este paradigma son:

- Algoritmos basados en valores:
  - Q-Learning: Busca aprender la función de valor de acción ( $Q(s, a)$ ) que indica la calidad de una acción ( $a$ ) en un estado ( $s$ ).
  - SARSA (State-Action-Reward-State-Action): Similar al Q-Learning, pero actualiza la función de valor basada en la acción realmente tomada por el agente.
- Algoritmos basados en políticas:
  - Policy Gradient: Optimiza directamente la política en lugar de la función de valor. Ejemplos incluyen REINFORCE y Actor-Critic.
  - Proximal Policy Optimization (PPO): Un método de optimización de políticas que equilibra la exploración y la explotación.
- Algoritmos basados en modelos:
  - Dyna-Q: Combina aprendizaje basado en modelos y sin modelos, utilizando simulaciones para mejorar la política.
  - AlphaZero: Utiliza una combinación de aprendizaje por refuerzo y búsqueda de árboles para juegos como el ajedrez y el Go.
- Algoritmos de aprendizaje profundo por refuerzo:
  - Deep Q-Network (DQN): Utiliza redes neuronales profundas para aproximar la función de valor ( $Q(s, a)$ ).
  - Deep Deterministic Policy Gradient (DDPG): Un algoritmo para entornos continuos que combina DQN y Policy Gradient.

## Otro tipo de taxonomía basada en las técnicas de aprendizaje de máquina.

En el mundo de los algoritmos relacionados con la inteligencia artificial, existen un conjunto de ellos que son capaces de ir mejorando la manera en que ejecutan una tarea dada a medida que acumulan experiencia o datos. Es un subcampo, cuyo funcionamiento tiende casi siempre a implicar algún tipo de estadística. Permiten la solución de problemas que ayudan al juicio experto o la intuición por el aprendizaje inferido a partir del análisis de los datos. Este es un enfoque que permite aprender patrones complejos a partir de datos existentes, y usar estos patrones para realizar predicciones sobre datos nunca antes analizados”.

Un aspecto a tener en cuenta en la utilización de este tipo de algoritmos es el hecho de que las inferencias realizadas por las máquinas están basadas especialmente en la coherencia con los datos, medida en términos de correlación y no en términos de causalidad.

- Cuando los datos de entrada se dicen que están correlacionados con los datos de salida, se está afirmando que ambos pueden estar relacionados.
- Cuando los datos de entrada se dicen que mantienen una relación de causalidad con los datos de salida, se está afirmando que los de entrada son la causa y los de salida son el efecto.

Dentro del aprendizaje automático, existe otro subcampo del anterior, el aprendizaje profundo, caracterizado por la complejidad de sus modelos matemáticos, unido a la capacidad de computación de los ordenadores modernos.

Las posibles características de las variables que intervienen en la representación de los datos son de dos tipos:

- Las características numéricas que representan o bien valores continuos o bien valores discretos que se pueden medir en una escala. Ejemplos de características numéricas podrían ser la edad, la altura, el peso y los ingresos. Las características numéricas se pueden usar directamente en algoritmos de aprendizaje automático.



- Las características categóricas son valores discretos que se pueden agrupar en categorías. Algunos ejemplos de características categóricas son género, color y código postal. Por lo general, las características categóricas deben convertirse en atributos numéricos antes de que puedan usarse en algoritmos de aprendizaje automático. Esto se puede hacer utilizando una variedad de técnicas, como:
  - Codificación One-Hot: Cada categoría se convierte en un vector binario, donde un 1 representa la presencia de la categoría y un 0 su ausencia. Por ejemplo, si tenemos una característica "Color" con categorías "Rojo", "Verde" y "Azul", obtendríamos tres vectores: [1, 0, 0], [0, 1, 0] y [0, 0, 1].
  - Codificación Ordinal: Se asignan valores numéricos a las categorías según su orden. Por ejemplo, "Bajo" = 1, "Medio" = 2 y "Alto" = 3.
  - Codificación de Etiquetas: Asignamos un valor único a cada categoría. Por ejemplo, "Rojo" = 1, "Verde" = 2 y "Azul" = 3.
  - Vectores de Incrustación (Embeddings): Utilizados en procesamiento de lenguaje natural (NLP), estos vectores representan palabras o categorías en un espacio continuo. Se aprenden a partir de datos y capturan relaciones semánticas.
  - Codificación de Frecuencia: Asignamos a cada categoría la frecuencia con la que aparece en los datos. Por ejemplo, si "Rojo" aparece 10 veces y "Verde" aparece 5 veces, tendríamos [10, 5] como vector.
  - Codificación de Target (o Codificación Bayesiana): Utilizamos la variable objetivo (target) para asignar valores a las categorías. Por ejemplo, si estamos prediciendo el precio de una casa, podríamos asignar a cada categoría de "Barrio" el promedio de los precios de las casas en ese barrio.
  - Codificación Helmert: Comparamos cada categoría con la media de las categorías anteriores. Por ejemplo, si tenemos "Bajo", "Medio" y "Alto", el vector para "Medio" sería [0.5, -0.5].

- Codificación Backward Difference: Similar a la codificación Helmert, pero comparamos cada categoría con la anterior en lugar de la media.
- ...

Otra forma de clasificar los algoritmos es basarse en los procedimientos de aprendizaje: En este caso la taxonomía suele ser:

- Supervisado. En este tipo es necesario tener un grupo de datos de entrada y para cada uno de ellos conocer con antelación los datos que deseamos predecir, tarea que se conoce como etiquetado de los datos.

El etiquetado de los datos es uno de los mayores retos del aprendizaje automático, dado que requiere de una toma de datos específica para la etiqueta/respuesta, inicialmente esta tarea es en ocasiones manual, y que obviamente ocasiona costes y puede conllevar errores.

- No supervisado. En éste, sólo tenemos los datos de entrada, pero desconocemos el dato que queremos predecir. En realidad, es un aprendizaje autosupervisado. La máquina se genera sus propios datos de aprendizaje.

Disponer de procedimientos analíticos que permitan minimizar el número de datos etiquetados, y a la vez utilizar datos no etiquetados para mejorar el ajuste, puede generar interesantes beneficios relacionados con la reducción de tiempos y costes vinculados al etiquetado.

- Sea cual sea el procedimiento que se siga, los conjuntos de datos se suelen dividir en tres grupos:
  - De entrenamiento. Conjunto de datos utilizados para entrenar un algoritmo de aprendizaje, esto es, para estimar parámetros y generar una primera respuesta al problema planteado.
  - De validación: Conjunto de datos utilizado tras el entrenamiento del modelo, para refinar las estimaciones de los parámetros del algoritmo de aprendizaje en muestras con variable respuesta.

- De test. Es el conjunto de datos utilizados para evaluar el rendimiento de un algoritmo de aprendizaje, separado de los datos de entrenamiento y validación, y no está disponible en la fase de aprendizaje.
- Por refuerzo. Lo que se busca es usar algoritmos que permitan a un agente aprender a interactuar con su entorno. El proceso implica capacitar a agentes para que tomen decisiones basadas en la retroalimentación del entorno.

### Algoritmos supervisados.

- Si se quiere predecir a partir de datos, una forma de clasificar algoritmos es basarse en lo que se busca:
  - Realizar tareas de clasificación trabajando con características categóricas.
    - La clasificación de Naïve Bayes
    - Los bosques aleatorios
    - La regresión logística
    - Los árboles de decisión
    - Árboles de clasificación
    - Árboles de regresión
    - Máquinas de vectores de soporte
    - Redes neuronales y aprendizaje profundo
    - ...
  - Realizar tareas de regresión trabajando con características numéricas.
    - Regresión lineal (utiliza datos continuos)
    - Regresión por mínimos cuadrados
    - Regresión logística (utiliza datos discretos)
    - Árboles de decisión
    - Redes neuronales y aprendizaje profundo
    - ...

## Algoritmos no supervisados.

- Estos algoritmos descubren agrupaciones de datos o patrones ocultos sin necesidad de ninguna intervención humana.
  - Agrupamiento (o clustering). Estos algoritmos permiten realizar tareas, para lo cual generan agrupaciones basadas en el grado de similitud entre los datos, de forma tal que datos similares pertenecerán a una misma agrupación. Los principales algoritmos en esta familia son:
    - K-means.
    - Mean Shift
    - Fuzzi C-Means
    - DBSCAN
    - Aglomerative
    - ...
  - Reglas de asociación. Estos permiten descubrir relaciones interesantes entre variables en grandes conjuntos de datos. Su objetivo es identificar reglas sólidas utilizando algunas medidas de interés. Los principales algoritmos en esta familia son:
    - FP Growth
    - Eclat
    - A priori
    - FPMax
    - COBWEB
    - Carte
    - Tertius
    - ...
  - Reducción de dimensionalidad. Permiten transformación de datos de un espacio de alta dimensión a un espacio de baja dimensión para que la representación de baja dimensión conserve algunas propiedades significativas de los datos originales, idealmente cercanas a su dimensión intrínseca. Los principales algoritmos son:
    - PCA (Principal Component Analysis)
    - LSA (Light Summary Algorithm) (para la criptografía)
    - SVD (Singular Value Decomposition)
    - LDA (Latent Dirichlet Allocation)

- t-SNE (t-Distributed Stochastic Neighbor Embedding)
- ...

Algoritmos combinados por conjuntos (ensemble).

- En este tipo de algoritmos se construye un modelo a partir de múltiples modelos más simples, con lo cual se busca mejorar el alcance del modelo. Se puede utilizar tanto en algoritmos supervisados como no supervisados, y algoritmos de regresión o clasificación. Los principales algoritmos en esta categoría son los:
  - Algoritmos de Bagging (bolsa)
    - Bosques aleatorios
    - Modelos simples usados de manera paralela.
  - Algoritmos de Boosting (potenciador)
    - XGBoost
    - LightGBM
    - CatBoost
    - AdaBoost
    - Modelos simples usados de manera secuencial.
  - Algoritmos de Stacking (apilamiento)
- Cuando tenemos datos no estructurados (como las imágenes, el video, el texto o el audio) resulta mucho más difícil usar estos algoritmos clásicos para realizar tareas de aprendizaje supervisado o no supervisado. En este caso la alternativa más adecuada son las arquitecturas del Deep Learning. Estas arquitecturas requieren una gran cantidad de datos de entrenamiento (del orden de cientos de miles o millones). Los principales algoritmos son:
  - Perceptrones
  - Redes multicapa
  - Autoencoders
  - seq2seq
  - Redes convolucionales
  - DCNN
  - Redes recurrentes

- LSM
- LSTM
- GRU
- Modelado generativo
- Redes antagónicas RGA o GAN
- Codificadores Variacionales
- Redes Transformer
- ...

## Inconclusión final

A pesar de todo, por todo lo contado, ¡hay que asomarse a una de las nuevas fronteras del conocimiento y por lo menos hay que curiosear y sobre todo intentar comprender lo que hay detrás de la inteligencia artificial. Consiguiendo de paso huir de lo que diría Arthur C. Clarke,

*"Cualquier tecnología suficientemente avanzada,  
si no se entiende,  
es equivalente a magia"*

Recordemos que el objetivo de este texto era intentar pasar de no tener una visión estructural, o tenerla borrosa o desordenada, a una cierta visión coherente.

Me gustaría haber ayudado a conseguirlo. Muchas gracias por su atención.