

Modelos de Inteligencia Artificial

*Conforme a contenidos del «Curso de Especialización
en Inteligencia Artificial y Big Data»*



**Modelos de
Inteligencia_Artificial**

Universidad de Castilla-La Mancha

Escuela Superior de Informática
Ciudad Real

Caracterización de sistemas de Inteligencia Artificial

La inteligencia artificial (IA) es un campo de investigación que existe desde hace más de 50 años y cuyo fin principal es el de crear inteligencia, concretamente, sistemas inteligentes. Aún estamos muy lejos de alcanzar esa denominada inteligencia, aunque ya se han hecho progresos en distintos ámbitos de la IA que podemos encontrar en nuestra vida diaria y que abarcan, por ejemplo, desde nuestros teléfonos móviles hasta sistemas de alarma en el hogar. Esto ha dado lugar a que la investigación en IA se fragmente en varias disciplinas, típicamente independientes. Pero para haber llegado hasta aquí, la IA ha experimentado una progresiva evolución que realmente se ha visto pronunciada en los últimos años.

En este capítulo se introducirán los conceptos fundamentales y la evolución histórica que han permitido la existencia de la IA como tal, algunos campos de aplicación más comunes, aquellas técnicas principales utilizadas en el desarrollo de tecnologías que utilicen IA, y las nuevas formas de interacción basadas en IA que se están utilizando actualmente en las empresas.

0.1. Fundamentos de la IA

0.1.1. Características principales

Antes de enumerar los conceptos que hacen realidad la IA, conviene detenerse a pensar y definir qué es realmente la IA, y es que, la IA es un término que se ha utilizado arbitrariamente en el mundo real para referirse a conceptos muy diversos y que se alejan de la realidad del concepto. Definir qué es la inteligencia daría para un debate filosófico de varios capítulos, lo cual escapa del ámbito que aquí nos preocupa. Siendo pragmáticos y respetando los avances realizados hasta ahora, podemos definir la IA como la ciencia encargada de construir sistemas que i) piensan como humanos, ii) piensan racionalmente, iii) actúan como humanos y iv) actúan racionalmente.

Actuar como humanos

Para comprobar si una máquina actúa como un humano podemos aplicar el **test de Turing**. Esta prueba plantea una serie de preguntas por escrito que una máquina deberá responder también por escrito; si la máquina responde a las preguntas y la persona que plantea las preguntas no es capaz de identificar si dichas respuestas vienen de un humano o una máquina, entonces la máquina habrá pasado el test de Turing. Hacer que una máquina pase dicha prueba requeriría que la máquina tuviera que cumplir con los siguientes puntos:

- Procesamiento del lenguaje natural, para comunicarse con el humano en el mismo lenguaje que él.
- Representación del conocimiento, para almacenar lo que sabe o escucha.
- Razonamiento automático, para responder a las preguntas y extraer nuevas conclusiones.
- Aprendizaje automático, para adaptarse a nuevas circunstancias, y detectar y extrapolar patrones.
- Visión por computador y reconocimiento del habla, para percibir el mundo que le rodea.
- Robótica, para manipular objetos y mantenerse activo.

Estas seis disciplinas conforman aquellas principales que debe tener una IA. Aun así, la mayoría de investigadores de IA han mostrado poco interés en intentar que una IA cumpla con estas seis disciplinas y por tanto conseguir que pase el test de Turing, creyendo que es más importante estudiar los principios subyacentes de la inteligencia.



El ser humano no consiguió hacer que los aviones volaran hasta que dejó de imitar a los pájaros.

Pensar como humanos

Para decir que un programa piensa como un humano, primero debemos de saber cómo piensan los humanos. Podemos aprender cómo funciona el pensamiento humano de tres formas distintas:

- Introspectiva, intentando analizar nuestros propios pensamientos al vuelo.
- Experimentación psicológica, observando a una persona en acción.
- Fotografiando el cerebro, observando así al cerebro en acción.

Una vez que tuviéramos suficiente información sobre la mente humana, sería posible expresar dicha información y las reglas que la gobiernan como un programa informático. Si el comportamiento de las entradas y salidas del programa coincide con el comportamiento humano correspondiente, estaríamos ante una prueba de que algunos de los mecanismos que operan el programa también podrían estar presentes en los humanos. Las limitaciones de este planteamiento hicieron surgir el campo de la **ciencia cognitiva**, que trata de unir los modelos informáticos de la IA con las técnicas experimentales de la psicología para construir teorías sobre la mente humana que sean precisas y comprobables. Sin embargo, resulta preciso aclarar que no se debería de pensar que, dado que un programa realiza bien una tarea signifique que se trata de un buen modelo de la cognición humana.

Pensar racionalmente

El filósofo griego Aristóteles fue uno de los primeros en intentar codificar formalmente los pasos correctos que realizamos a la hora de pensar cuando planteamos un razonamiento. Sus **silogismos** proporcionaron patrones para estructurar argumentos que siempre arrojaban conclusiones correctas cuando se proporcionaban las premisas adecuadas. Así, desde el siglo XIX esto sería conocido como la **lógica** que conocemos hoy en día y que dio lugar a una notación común que permitiría definir los objetos de nuestro mundo y las relaciones que existen entre ellos.



El ejemplo canónico de silogismo que formuló Aristóteles denominado *modus ponens* fue que si *Sócrates es un hombre* y *todos los hombres son mortales*, entonces *Sócrates es mortal*.

Usando esta notación lógica, los programas informáticos podrían llegar a resolver cualquier problema descrito con dicha notación. Así, la llamada doctrina *logicista* dentro de la inteligencia artificial confía en poder basarse en esos programas para crear sistemas inteligentes. Sin embargo, el uso de la lógica tal y como la conocemos requiere un conocimiento del mundo que sea certero, algo que en la práctica es imposible de conseguir; por ejemplo, no conocemos todas las reglas que definen las guerras entre países del mismo modo que conocemos las reglas del ajedrez. De esta forma, no todo el comportamiento inteligente es fácil de modelar a través de la lógica.

Actuar racionalmente

Podemos identificar un comportamiento racional como aquel que hace lo correcto, o lo que es lo mismo, maximiza el objetivo (meta) con la información que hay disponible o, cuando hay incertidumbre, obtiene el mejor resultado posible. En el contexto de la IA, podemos modelar este comportamiento racional mediante un enfoque de agentes que actúen de manera autónoma, perciban su entorno, persistan durante un tiempo prolongado, se adapten al cambio y creen y alcancen metas.

Este enfoque de abordar la IA mediante agentes racionales ha conseguido prevalecer a lo largo de la historia. En los inicios, los agentes racionales eran contruidos sobre fundamentos lógicos definiendo planes para lograr metas concretas. Más tarde, se idearon métodos basados en la teoría de la probabilidad y el aprendizaje automático que permitieron la creación de agentes que podían tomar decisiones en situaciones de incertidumbre para lograr el mejor resultado posible. Así, la IA se ha centrado en el estudio y la construcción de agentes que hacen lo correcto; definir lo que es correcto viene dado por el objetivo que le proporcionamos al agente.

Aun así, este enfoque sigue presentando el problema de la incertidumbre en el mundo real, seguido por la complejidad de las metas a modelar donde la racionalidad perfecta no es posible; es imposible para un agente actuar apropiadamente cuando no hay suficiente tiempo para realizar todos los cálculos computacionales que necesitamos.

0.1.2. Tipos de inteligencia

Los conceptos anteriores servirían para definir una IA de tipo generalista, capaz de pensar como un humano. Este tipo de IA suele decirse que presenta una **inteligencia general o fuerte**. Como seres humanos, somos capaces de aprender de diversas experiencias e interacciones en el mundo y aplicar esa comprensión de un problema a otro. Por ejemplo, si de niño sentías dolor al tocar algo caliente, puedes extrapolar y saber que otras cosas calientes pueden hacerte daño. Sin embargo, la inteligencia general en los humanos es algo más que razonar algo como *las cosas calientes pueden ser dañinas*. La inteligencia general abarca la memoria, el razonamiento espacial a través de medios visuales, el uso del conocimiento y mucho más. Lograr la inteligencia general en una máquina parece una hazaña improbable a corto plazo, pero los avances en la computación cuántica, el procesamiento de datos y los algoritmos de IA podrían hacerlo realidad en el futuro.

Sería posible reutilizar algunos de los conceptos introducidos en las inteligencias fuertes para resolver problemas de dominios más concretos y acotados, dando lugar a una IA que presente una **inteligencia débil o limitada**. Por lo general, estos sistemas no pueden resolver un problema en un determinado contexto y aplicar esa misma comprensión en otro. Un sistema desarrollado para entender las interacciones de los clientes y el comportamiento de gasto, por ejemplo, no sería capaz de identificar gatos en una imagen. Normalmente, para que algo sea eficaz en la resolución de un problema, tiene que estar bastante especializado en el dominio del problema, lo que dificulta su adaptación a otros problemas.

Sin embargo, se pueden combinar varios de estos sistemas de IA débil para crear algo mayor que se asemeje a una IA fuerte. Un ejemplo sería un asistente de voz; este sistema puede entender el lenguaje natural, que por sí solo es un problema acotado, pero mediante la integración con otros sistemas de inteligencia débil, como las búsquedas en la web y las recomendaciones personalizadas de música, puede mostrar cualidades de una IA fuerte.

0.1.3. Historia de la IA

Una vez introducidas las bases que conforman la IA, podemos definir brevemente la evolución histórica que ha sufrido para situarnos en contexto:

- **1943:** surgen las redes neuronales artificiales en un artículo de investigación por McCulloch y Pitts, quienes idearon un modelo matemático sencillo y demostraron cómo podía utilizarse para calcular funciones lógicas arbitrarias. Gran parte de los primeros trabajos se centraron en comprender las propiedades matemáticas de estas redes, ya que los computadores de la época no eran lo suficientemente potentes como para realizar tareas de interés.
- **1956:** se lleva a cabo un seminario en la universidad Dartmouth College organizado por John McCarthy donde se acuña el término *Inteligencia Artificial* por primera vez.
- **1966:** a pesar del futuro prometedor que tenía la IA por delante (y las altas expectativas), algunas tareas como la traducción entre idiomas fue un fracaso total, lo que llevó a recortes en los programas de financiación, lo que llevaría al primer *invierno* de la IA. Algunos problemas que ocasionaron esto fueron las limitadas capacidades de los computadores de la época y la complejidad de los problemas que se querían resolver mediante IA, los cuales requerían además una ingente cantidad de datos. Aun así, todo ello repercutió en la aparición de nuevas tecnologías, como el lenguaje de programación Lisp, la técnica de liberación de memoria conocida como *recolección de basura* y un cambio clave de paradigma: separar lo que se quiere trabajar computacionalmente (modelado) de cómo trabajarlo (inferencia).
- **1969:** se publica un libro que exploraba las propiedades matemáticas de los *perceptrones* (modelos lineales) y donde se demostraba que ni siquiera eran capaces de resolver problemas simples como una puerta lógica XOR (OR exclusiva). Debido a esto, el interés en la investigación de redes neuronales se ve dramáticamente reducido.
- **1970-1980s:** surgen los sistemas basados en el conocimiento con el objetivo de permitir la codificación del conocimiento directamente en aquellos sistemas inteligentes para que tengan toda la información necesaria, reduciendo así los costes computacionales de rendimiento. Los investigadores se enfocaron en construir sistemas más acotados que se centraran en dominios concretos, dando lugar a los llamados *sistemas expertos*. Algunos ejemplos de estos sistemas serían DENDRAL (inferencia la estructura molecular a partir de la espectrometría de masas) y MYCIN (diagnóstico de infecciones en la sangre y recomendación de antibióticos), entre otros.
- **1986:** en esta época se re-descubre el algoritmo de propagación hacia atrás de errores o *backpropagation* en inglés, popularizándose así el entrenamiento de redes neuronales multi-capas y devolviendo de nuevo el interés investigador a las redes neuronales, perdido en 1969.

- **1987:** aparecen las primeras aplicaciones que tienen un impacto en la industria. Sin embargo, la tecnología existente se enfrentó a las limitaciones computacionales de la época y fracasó de nuevo a la hora de abarcar problemas más complejos que requerían modelar más cantidad de incertidumbre a partir de un considerable esfuerzo manual para crear las reglas necesarias, las cuales eran difíciles de mantener a largo plazo. Debido a un exceso de promesas y una falta de cumplimiento, el campo de la IA volvió a colapsar, dando lugar al segundo *invierno* de la IA.
- **1989:** Yann LeCun construye un sistema basado en redes neuronales para reconocer dígitos escritos a mano, convirtiéndose en uno de los primeros casos de éxito de aplicación de las redes neuronales, que además fue utilizado por el servicio de correo postal de los Estados Unidos (USPS) para reconocer códigos postales.
- **1990-2000s:** en estos años es donde aparecen los enfoques modernos de aplicación de la IA basados en probabilidad, estadística e incertidumbre. Se producen avances y mejoras técnicas en las tecnologías existentes, y surge el concepto de *big data* para referirnos a la producción de grandes conjuntos de datos (*datasets*) del orden de magnitud de, por ejemplo, trillones de palabras en un texto, billones de imágenes, y billones de horas de audio y vídeo, entre otros.
- **2010s-Actualidad:** en esta época se produce el punto de inflexión real en el ámbito de la IA. Con el auge de la computación debido a las capacidades técnicas de las GPUs y los enormes *datasets* existentes, como ImageNet (2009), la IA a través de las redes neuronales y el aprendizaje profundo (*deep learning*) había empezado a popularizarse.

Como se puede observar, la IA todavía es un campo de investigación joven que está empezando a producir resultados que pueden ser aprovechados por las empresas. Entre los campos de aplicación actuales podemos listar algunos como, por ejemplo, vehículos autónomos, reconocimiento del habla, traducciones automáticas, lucha contra el spam, sistemas de recomendaciones, videojuegos, reconocimiento de imágenes y medicina, entre otros.



El uso de la IA también presenta ciertos riesgos y compromisos éticos además de los posibles beneficios. Uno de estos riesgos obvios podría presentarse al intentar controlar sistemas inteligentes que evolucionen impredeciblemente a la comprensión humana.

0.2. Campos de aplicaciones

Como hemos visto, la IA es un campo de estudio joven y en constante evolución. Tanto es así, que continuamente se encuentran nuevos campos de aplicación donde la sociedad puede beneficiarse de la IA; donde haya datos y problemas que resolver, existirán aplicaciones de la IA potenciales. Así, las interacciones entre los seres humanos y las nuevas necesidades exigidas por las personas e industrias, la IA puede aplicarse de forma innovadora para resolver problemas del mundo real. Vamos a ver en esta sección algunas de las principales aplicaciones de la IA en la industria.

0.2.1. Agricultura: cultivo óptimo

La agricultura, aunque pueda parecer alejado del mundo de la IA, plantea problemas y retos que pueden resolverse aplicando tecnologías de IA: los agricultores tienen que ser capaces de cultivar productos de calidad para el consumo masivo de forma económica.

Muchos agricultores cultivan a escala comercial para que podamos comprar frutas y verduras en las tiendas con comodidad. Los cultivos crecen de forma diferente según el tipo de cultivo, los nutrientes del suelo, el contenido de agua del suelo, las bacterias del agua y las condiciones meteorológicas de la zona, entre otras cosas. El objetivo es cultivar la mayor cantidad posible de productos de alta calidad dentro de una temporada, ya que determinados cultivos suelen crecer bien sólo durante determinadas estaciones.

Los agricultores y otras organizaciones agrícolas llevan recopilando datos sobre sus explotaciones y cultivos desde hace años. Con esos datos, podemos aprovechar la tecnología para encontrar patrones y relaciones entre las variables del proceso de cultivo e identificar los factores que más contribuyen al éxito del crecimiento. Además, con los sensores digitales actuales, podemos registrar en tiempo real las condiciones meteorológicas, los atributos del suelo, las condiciones del agua y el crecimiento de los cultivos. Estos datos, combinados con algoritmos inteligentes, pueden proporcionar recomendaciones y ajustes en tiempo real para optimizar el crecimiento de los cultivos.

Por ejemplo, dadas algunas variables como la temperatura, el tipo de cultivo, las propiedades de la tierra y la cantidad de humedad, podemos crear algoritmos de IA que sean capaces de determinar qué sería lo más apropiado cultivar, cuándo hacerlo y qué necesidades de riego y fertilizante se requerirían.

0.2.2. Banca: detección de fraude

La banca ha evolucionado durante años para ofrecer diferentes opciones para guardar dinero, invertir dinero y realizar pagos. Una cosa que no ha cambiado es la creatividad de la sociedad para encontrar nuevas formas de engañar al sistema. Uno de los mayores problemas (no solo de la banca, sino de de la mayoría de instituciones financieras, como las aseguradoras) es el fraude.

El fraude ocurre cuando alguien hace algo ilegal o con malas intenciones para conseguir algo. Normalmente sucede cuando se aprovechan las lagunas existentes en algún proceso o se engaña a alguien para obtener información. Dado que el sector de los servicios financieros está muy conectado a través de Internet y de los dispositivos personales, hoy en día se realizan más transacciones electrónicas a través de una red informática que en persona, con dinero físico.

Gracias a la gran cantidad de datos disponible sobre las transacciones realizadas, podemos identificar, en tiempo real, patrones de transacciones específicos al comportamiento de gasto de un individuo que pueden estar fuera de lo normal. Estos datos ayudan a ahorrar enormes cantidades de dinero a las instituciones financieras y protegen a los consumidores desprevenidos de los robos.

0.2.3. Ciberseguridad: detección y tratamiento de ataques

Uno de los efectos secundarios interesantes del auge de Internet es la ciberseguridad. Enviamos y recibimos información sensible a través de Internet todo el tiempo: mensajes instantáneos, datos de tarjetas de crédito, correos electrónicos y otros datos confidenciales importantes que podrían ser mal utilizados si cayeran en manos equivocadas. Miles de servidores de todo el mundo reciben datos, los procesan y los almacenan. Los atacantes intentan comprometer estos sistemas para acceder a los datos, los dispositivos o incluso las instalaciones físicas.

Utilizando la IA, podemos identificar y bloquear posibles ataques a los servidores. Algunas grandes empresas de Internet almacenan datos sobre la forma en que determinadas personas interactúan con su servicio, incluidos los identificadores de sus dispositivos, la geolocalización y la forma de uso; cuando se detecta un comportamiento inusual, las medidas de seguridad limitan el acceso. Algunas empresas de Internet también pueden bloquear y redirigir el tráfico malicioso durante un ataque de denegación de servicio distribuido (DDoS), que consiste en sobrecargar un servicio con peticiones falsas en un intento de hacerlo caer o impedir el acceso de los usuarios auténticos. Estas peticiones no legítimas pueden ser identificadas y redirigidas para minimizar el impacto del ataque mediante el conocimiento del comportamiento de los usuarios, los sistemas y la red.

0.2.4. Atención sanitaria: diagnóstico de pacientes

La atención sanitaria ha sido una inquietud constante a lo largo de la historia de la humanidad. Necesitamos tener acceso al diagnóstico y al tratamiento de diferentes afecciones en distintos lugares y en distintos periodos de tiempo antes de que un problema se agrave o incluso sea mortal. Cuando nos fijamos en el diagnóstico de un paciente, podemos observar la gran cantidad de información registrada sobre el cuerpo humano, los problemas conocidos, la experiencia en el tratamiento de estos problemas y un sinfín de exploraciones del cuerpo. Tradicionalmente, los médicos debían analizar las imágenes de las exploraciones para detectar la presencia de tumores, pero este enfoque daba lugar a la detección de únicamente los tumores más grandes y avanzados. Los avances en el aprendizaje profundo han mejorado la detección de tumores en las imágenes generadas por las exploraciones. Ahora los médicos pueden detectar el cáncer con mayor antelación, lo que significa que el paciente puede recibir el tratamiento necesario a tiempo y tener una mayor esperanza de recuperación.

Además, la IA puede utilizarse para encontrar patrones en síntomas, dolencias, genes hereditarios, ubicaciones geográficas y similares. Podríamos saber que alguien tiene una alta probabilidad de desarrollar una afección específica y estar preparados para gestionar esa afección antes de que se desarrolle.

0.2.5. Logística: rutas y optimización

El sector de la logística es un enorme mercado formado por diferentes tipos de vehículos que entregan diversos tipos de mercancías en distintos lugares, con diferentes exigencias y plazos. Imaginemos la complejidad de la planificación de entregas de un gran portal de comercio electrónico. Tanto si se trata de bienes de consumo como de equipos de construcción, piezas de maquinaria o combustible, el sistema pretende ser lo más óptimo posible para garantizar que se satisfaga la demanda y se minimicen los costes.

Esta problemática es similar al problema del viajante de comercio: un vendedor tiene que visitar varios lugares para completar su trabajo, y el objetivo es encontrar la distancia más corta para realizar esta tarea. Los problemas de logística son similares, pero suelen ser inmensamente más complejos debido al entorno cambiante del mundo real. Gracias a la IA, podemos encontrar rutas óptimas entre distintas ubicaciones en términos de tiempo y distancia. Además, podemos encontrar las mejores rutas en función de los patrones de tráfico, los bloqueos de las obras e incluso los tipos de carreteras en función del vehículo utilizado. Además, podemos calcular la mejor manera de cargar cada vehículo de tal forma que se optimice cada entrega.

0.2.6. Telecomunicaciones: optimización de redes

La industria de las telecomunicaciones ha desempeñado un enorme papel en la conexión del mundo. Estas empresas instalan costosas infraestructuras compuestas por cables, torres y satélites para crear una red que pueda ser utilizada por consumidores y organizaciones para comunicarse a través de Internet o de redes privadas. El funcionamiento de estos equipos es caro, por lo que la optimización de una red permite un mayor número de conexiones, lo que permite que más personas accedan a conexiones de alta velocidad. La IA puede utilizarse para supervisar el comportamiento de una red y optimizar el enrutamiento en función de distintas variables como, por ejemplo, el tráfico de datos asociado a determinadas personas, zonas y redes locales específicas. Los datos de la red también pueden servir para saber dónde está la gente y quiénes son, lo que resulta útil para la planificación urbana.

0.2.7. Juegos: creación de agentes inteligentes

Desde que los ordenadores personales y domésticos empezaron a estar disponibles, los juegos han sido un reclamo para la venta de computadoras. Los juegos fueron muy populares en los primeros años de la historia de los ordenadores personales. Si pensamos en el pasado, podemos recordar las máquinas recreativas, los televisores y los ordenadores personales con capacidades de entretenimiento. Los juegos de ajedrez, backgammon y otros han sido dominados por las máquinas de IA. Si la complejidad de un juego es lo suficientemente baja, un ordenador puede encontrar potencialmente todas las posibilidades y tomar una decisión basada en ese conocimiento más rápido que un humano. Recientemente, un ordenador fue capaz de derrotar a campeones humanos en el juego de estrategia Go. El Go tiene reglas sencillas para el control del territorio, pero tiene una enorme complejidad en cuanto a las decisiones que hay que tomar para un escenario ganador. Un ordenador no puede generar todas las posibilidades para vencer a los mejores jugadores humanos porque el espacio de búsqueda es muy grande; en su lugar, se necesita un algoritmo más general que pueda "pensar" de forma abstracta, elaborar estrategias y planificar movimientos hacia un objetivo. Ese algoritmo ya se ha inventado y ha conseguido derrotar a campeones del mundo. También se ha adaptado a otras aplicaciones, como los juegos de la plataforma Atari y otros juegos multijugador. Este sistema se llama Alpha Go.

Varios grupos de investigación han desarrollado sistemas de IA capaces de jugar a juegos muy complejos mejor que los jugadores y equipos humanos. El objetivo de este trabajo es crear enfoques generales que puedan adaptarse a diferentes contextos. A primera vista, estos algoritmos de IA para jugar pueden parecer poco importantes, pero la consecuencia de desarrollar estos sistemas es que el enfoque pueda aplicarse eficazmente en otros problemas más importantes.



Figura 1: Imágenes de un aguacate generadas mediante la IA llamada DALL-E a partir de la siguiente descripción textual (adaptada del inglés): *un sillón con forma de aguacate*.

0.2.8. Arte: creatividad

En la pintura, cada artista tiene su propia manera de expresar el mundo que le rodea, al igual que en la música a la hora de componer una canción. En ambos casos, la calidad del arte no puede medirse cuantitativamente, sino cualitativamente (por lo que la gente disfruta de la obra). Los factores que intervienen son difíciles de entender y captar; el concepto se rige por la emoción.

Existen muchos proyectos de investigación que pretenden crear una IA que genere arte. El concepto implica la generalización; un algoritmo tendría que tener un conocimiento amplio y general del tema para crear algo que se ajuste a esos parámetros. Una IA de Van Gogh, por ejemplo, tendría que entender toda la obra de Van Gogh y extraer el estilo y la *sensación* para poder aplicar esos datos a otras imágenes.

En cualquier caso, el campo ha generado mucha expectación entre el público e interés por parte de los investigadores. En 2021 el laboratorio de investigación en IA, OpenAI, publicó una serie de imágenes generadas a partir de descripciones textuales mediante su algoritmo DALL-E ¹ (ver Figura 1).

0.3. Técnicas de la Inteligencia Artificial

Como se ha visto, los campos de aplicación de la IA son muy diversos. Es por ello que existen multitud de técnicas que podemos aplicar en función del problema al que nos enfrentemos; no existe un algoritmo *maestro* que nos permita abordar de manera generalista cualquier problema de IA, tal y como se ha introducido al comienzo del capítulo.

En esta sección se introducirán algunas de las técnicas más importantes utilizadas para *crear* IA. Cabe destacar que cada una de estas técnicas requerirían de varias páginas para explicarlas y detallarlas en profundidad, por lo que se explicarán los conceptos más importantes de manera superficial para que el lector tenga al menos una idea básica de las herramientas disponibles.

Algunas de estas técnicas para abordar problemas de IA se han clasificado históricamente en las denominadas *IA antigua* e *IA moderna*. Se entiende como IA antigua como aquella presente en los sistemas en los que las personas codificaban manualmente las reglas que gobernaban los algoritmos inteligentes, ya fuera a

¹ Link: <https://openai.com/blog/dall-e/>

través de un amplio conocimiento del dominio del problema o mediante prueba y error. Un ejemplo de esto sería el de una IA que utilizara un árbol de decisiones, sus reglas y opciones, creado manualmente por alguien. Por otra parte, la IA moderna pretende crear algoritmos y modelos que aprendan de los datos y creen sus propias reglas con la misma o mejor precisión que las creadas por los humanos. La diferencia principal entre ambas inteligencias es que la IA moderna puede encontrar patrones importantes en los datos que una persona nunca encontraría o que le llevaría mucho más tiempo encontrar.

Los algoritmos de búsqueda se consideran a menudo como IA antigua, pero es necesario entenderlos para aprender otras técnicas más complejas, tal y como veremos en esta sección.

0.3.1. Algoritmos de búsqueda

Los algoritmos de búsqueda son útiles para resolver problemas en los que se requieren varias acciones para alcanzar un objetivo, como encontrar un camino a través de un laberinto o determinar el mejor movimiento a realizar en un juego. Los algoritmos de búsqueda evalúan los estados futuros e intentan encontrar el camino óptimo hacia el objetivo más valioso. Normalmente, tenemos demasiadas soluciones posibles para forzar cada una de ellas. Incluso los espacios de búsqueda más pequeños pueden suponer miles de horas de cálculo para encontrar la mejor solución. Los algoritmos de búsqueda proporcionan formas inteligentes de evaluar el espacio de búsqueda y se utilizan en ámbitos tan diversos como, por ejemplo, motores de búsqueda, aplicaciones de mapas e incluso en videojuegos.

Para representar el espacio de búsqueda utilizado por un algoritmo de búsqueda se suelen utilizar estructuras de datos como grafos (utilizando matrices de adyacencia o listas de aristas) y árboles, principalmente estos últimos.

Los algoritmos de búsqueda pueden clasificarse en i) búsqueda no informada y ii) búsqueda informada.

Búsqueda no informada

En este tipo de algoritmos, el algoritmo de búsqueda no dispone de información adicional sobre el dominio del problema, más allá de la representación del problema (comúnmente en forma de *árbol*).

Un ejemplo típico de este tipo de búsqueda es el que se plantea a la hora de aprender sobre algo: algunas personas pueden estudiar una amplia cantidad de temas diferentes y aprender superficialmente lo básico de cada uno de ellos, mientras que otras personas pueden elegir un tema específico y explorar sus subtemas en profundidad. Esto es lo que implica la **búsqueda en anchura (BFS)** y la **búsqueda en profundidad (DFS)**, respectivamente: la búsqueda en anchura explora todas las opciones a una profundidad determinada antes de pasar a opciones más profundas en el *árbol*; la búsqueda en profundidad explora un camino específico desde el principio hasta que encuentra un objetivo a la máxima profundidad.

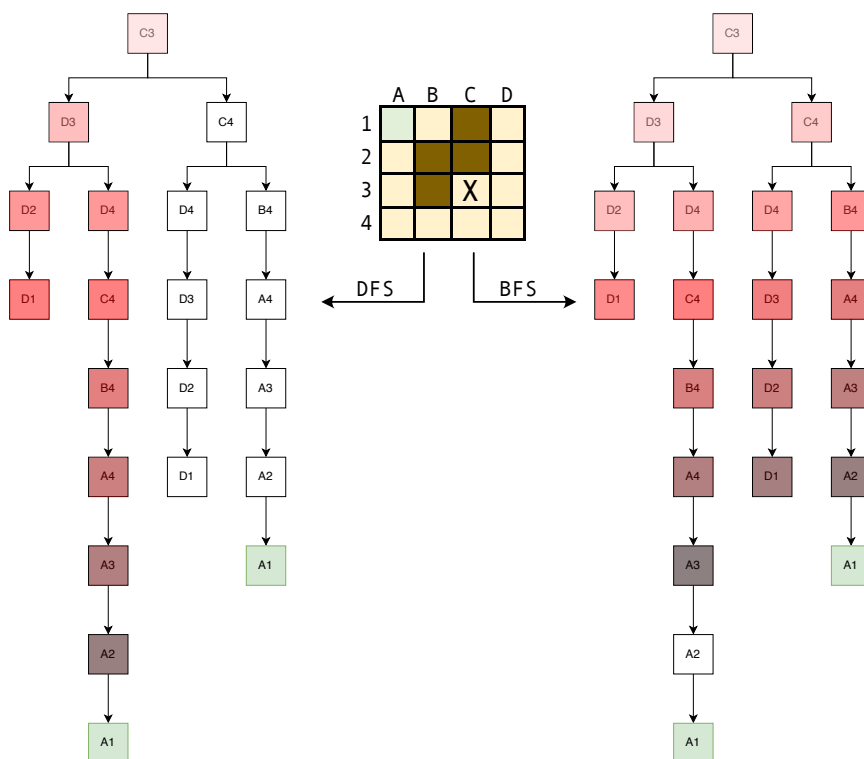


Figura 2: Búsqueda en profundidad (izquierda) y búsqueda en anchura (derecha) aplicadas a la resolución de un laberinto.

En la Figura 2 se muestra un ejemplo de ambos tipos de búsqueda aplicados a la resolución de un laberinto. El punto inicial del laberinto sería la casilla C3, y el objetivo es alcanzar la casilla A1. En el caso de la búsqueda en profundidad, la secuencia de movimientos trataría de seguir el camino $C3 > D3 > D2 > D1$; al no alcanzar la meta, intentaría de nuevo por la rama siguiente: $C3 > D3 > D4 > C4 > B4 > A4 > A3 > A2 > A1$, alcanzando en este caso la meta. En el caso de la búsqueda en anchura, la secuencia de pasos sería tal que así: $C3 > D3$, $C3 > C4$, $C3 > D3 > D2$, $C3 > D3 > D4$, $C3 > C4 > D4$, etc. Como vemos, en este caso aplicar la búsqueda en profundidad resulta más eficiente (se reduce el número de pasos) que aplicar una búsqueda en anchura.

Búsqueda informada

A diferencia de la búsqueda no informada, los algoritmos de búsqueda informada disponen de algún contexto del problema específico que están resolviendo. Las **heurísticas** son una forma de expresar este contexto, mediante una regla o conjunto de reglas usadas para evaluar un estado; pueden ser usadas para definir criterios que deban cumplir los estados o para medir el rendimiento de un estado concreto. En resumen, las heurísticas se utilizan cuando no existe un método para encontrar la solución óptima a un problema.

Podríamos modificar el ejemplo de la Figura 2 para añadir *costes* a algunos movimientos, concretamente, para desplazarse al este o al oeste por el laberinto existe un coste de 3 minutos, mientras que para ir al norte o al sur existe un coste de 1 minuto. Así, se podría definir una heurística para este nuevo ejemplo que sea: los mejores caminos minimizan el coste temporal del movimiento y minimizan el número total de movimientos para llegar a la meta. Esta heurística ayudaría a dirigir los nodos que se visitan porque estaríamos aplicando algún conocimiento del dominio para resolver el problema.

Normalmente, este tipo de búsqueda utiliza los algoritmos de búsqueda en anchura y búsqueda en profundidad, en combinación con algún tipo de inteligencia; la búsqueda es guiada por heurísticas, dado un conocimiento predefinido del problema en cuestión.

Los algoritmos más comunes que se suelen utilizar en este tipo de búsqueda son los de búsqueda voraz (también conocida como buscar primero el mejor) y el **algoritmo A*** (pronunciado como *A estrella*), siendo este último el más popular y útil en este tipo de búsquedas.



Existe otro tipo de problemas que se resuelven mediante la denominada búsqueda entre adversarios y que aplica el algoritmo de decisiones minimax para resolverlos y la poda alfa-beta como técnica de optimización. Este tipo de problemas surge principalmente en aquellos entornos que pueden cambiar por influencia de agentes externos, como por ejemplo, juegos en los que participan más de un jugador.

0.3.2. Algoritmos evolutivos

Los algoritmos evolutivos se inspiran en la teoría de la evolución definida por Charles Darwin. El concepto es que una población se reproduce para crear nuevos individuos y que a través de este proceso, la mezcla de genes y la mutación se producen individuos (descendientes) que rinden más que sus antepasados. Cabe destacar que la evolución no es un proceso lineal, sino que mediante la reproducción se producen una gran cantidad de variantes de una especie con distintas mezclas genéticas.

Aplicando este concepto a la informática, se aprovecha el aprendizaje de la evolución biológica para encontrar soluciones óptimas a problemas prácticos generando diversas soluciones y convergiendo en las más eficaces a lo largo de muchas generaciones.

Los **algoritmos genéticos** son una serie de algoritmos específicos de tipo evolutivo. Se utilizan para evaluar grandes espacios de búsqueda y *tratar* de encontrar una buena solución. Cabe destacar que los algoritmos genéticos no garantizan que se encuentre la mejor de las soluciones, pero sí una de las mejores. El flujo de trabajo de un algoritmo genético sigue el siguiente procedimiento:

1. Codificar el espacio de soluciones, por ejemplo, mediante una representación binaria de las soluciones.
2. Establecer los parámetros del algoritmo.
3. Crear un conjunto (población) de potenciales soluciones (individuos), por ejemplo, de manera aleatoria y validando la solución obtenida para comprobar que no se viola ninguna de las restricciones.
4. Medir la aptitud (*fitness*) de los individuos de la población; determinar la calidad de una solución concreta. Esta tarea se lleva a cabo utilizando una función de aptitud (similar a las heurísticas de los algoritmos de búsqueda) que puntúa las soluciones para determinar su calidad.
5. Selección de progenitores en función de su aptitud: selección de individuos que reproducirán la descendencia a partir de aquellos que obtuvieran la mayor puntuación (o menor, dependiendo del problema) en el paso anterior, seleccionando, por ejemplo, los N mejores.
6. Reproducción de individuos: generar descendencia a partir de los individuos seleccionados en el paso anterior combinando la información genética (cruce) y aplicando ligeros cambios a la descendencia (mutaciones).
7. Poblar la siguiente generación: seleccionar los individuos y los descendientes de la población que sobrevivirán a la siguiente generación.

Otro tipo de algoritmos evolutivos son las **inteligencias de enjambre** (*swarm*) donde grupos de individuos aparentemente *tontos* muestran un comportamiento inteligente. La *optimización por colonias de hormigas* y la *optimización por enjambres de partículas* son dos algoritmos populares de este tipo.

0.3.3. Aprendizaje automático (*machine learning*)

El aprendizaje automático adopta un enfoque estadístico para entrenar modelos que aprendan de los datos. El campo del aprendizaje automático cuenta con una variedad de algoritmos que pueden aprovecharse para mejorar la comprensión de las relaciones entre los datos, para tomar decisiones y para hacer predicciones basadas en esos datos.

Existen tres técnicas principales de aprendizaje automático:

- **Aprendizaje supervisado:** consiste en entrenar modelos con algoritmos cuando los datos de entrenamiento tienen resultados conocidos para una determinada pregunta como, por ejemplo, identificar el tipo de fruta si tenemos un conjunto de datos que incluye el peso, el color, la textura y el nombre de la fruta para cada caso.
- **Aprendizaje no supervisado:** trata de descubrir relaciones y estructuras ocultas dentro de los datos que nos guían a la hora de formular las preguntas relevantes sobre el conjunto de datos. Por ejemplo, puede encontrar patrones en las propiedades de frutas similares y agruparlas en consecuencia, lo que puede darnos una idea de las preguntas exactas que queremos hacer a los datos. Estos conceptos y algoritmos nos ayudan a crear una base que nos permita desarrollar algoritmos más avanzados en el futuro.
- **Aprendizaje por refuerzo:** se inspira en la psicología del comportamiento. En resumen, describe cómo se recompensa a un individuo si se realiza una acción positiva y se le penaliza si se realiza una acción negativa. Por ejemplo, cuando un niño obtiene buenos resultados en su expediente académico, suele ser recompensado, pero un mal rendimiento a veces da lugar a un castigo, lo que refuerza en el niño el comportamiento de obtener buenos resultados. El aprendizaje por refuerzo es útil para explorar cómo los programas informáticos o los robots interactúan con entornos dinámicos. Un ejemplo es un robot al que se le encarga abrir puertas; se le penaliza cuando no abre una puerta y se le recompensa cuando lo hace. Con el tiempo, tras muchos intentos, el robot *aprende* la secuencia de acciones necesarias para abrir una puerta.

De manera simplificada, podemos resumir los pasos implicados en el aprendizaje automático con los siguientes:

1. Recolección y comprensión de los datos.
2. Preparación de los datos: implica adaptar los datos al formato correcto y admitido por los algoritmos que vayamos a utilizar, y corregir, eliminar o adaptar aquellos datos incompletos o cualitativos que necesiten ser cambiados a valores cuantitativos.
3. Entrenamiento del modelo: para realizar predicciones utilizando el modelo, se debe utilizar el algoritmo correcto en función de la información que queramos obtener y los datos de los que dispongamos. Existen varios algoritmos para ello. Por ejemplo, la regresión lineal es uno de los algoritmos de aprendizaje automático más simples; busca relaciones entre dos variables y realiza predicciones de una de ellas (variable dependiente) en función de la otra (variable independiente), basándose en las ecuaciones de la recta ($y = c + mx$) y la inferencia de la pendiente (m) y el valor donde la recta intercepta al eje Y (c) mediante algún método de ajuste como, por ejemplo, el de los mínimos cuadrados. Conseguir esta línea de regresión significaría haber entrenado al modelo.

4. Probar el modelo: para ello, se mide la precisión del modelo con ejemplos ya conocidos de antemano y comprobando si el modelo arroja los resultados esperados. Cabe destacar, que no se puede probar el modelo con los mismos datos que se han usado para entrenarlo. Normalmente se suele reservar un 20 % de los datos disponibles para realizar las pruebas, y un 80 % para entrenar el modelo. La precisión del modelo puede probarse con algún método conocido. En el ejemplo de la regresión lineal, puede utilizarse el coeficiente de determinación R-cuadrado (R^2) para obtener un porcentaje de precisión del modelo para los datos de prueba.
5. Mejorar la precisión del modelo: en caso de que al medir la precisión del modelo durante las pruebas no se haya obtenido la precisión deseada, habría que volver a iterar sobre alguno de los pasos anteriores. Algunas de las recomendaciones que se suelen seguir en este caso serían i) recopilar más datos, ii) preparar o adaptar los datos usando otras técnicas, iii) seleccionar otras características de los datos, o iv) usar un algoritmo diferente para entrenar el modelo como, por ejemplo, árboles de decisión en caso de que nos enfrentemos a problemas de clasificación en lugar de regresión.

Estos pasos suelen realizarse en orden, aunque el último de ellos puede implicar volver a cualquiera de los anteriores para realizar posibles ajustes.

Redes neuronales artificiales (ANN)

Las redes neuronales artificiales (ANN, por sus siglas en inglés) son unas potentes herramientas dentro del ámbito del aprendizaje automático, que se utilizan de diversas maneras para conseguir distintos objetivos, como el reconocimiento de imágenes, el procesamiento del lenguaje natural o la resolución de problemas derivados de los juegos. Las ANNs aprenden de forma similar a otros algoritmos de aprendizaje automático: utilizando datos para ser entrenadas. Pero en este caso, se adaptan mejor a los datos no estructurados en los que es difícil entender cómo se relacionan las características entre sí.

Podemos entender las ANNs como otro modelo a utilizar en el proceso de aprendizaje automático que deberá ser entrenado y probado correctamente. Las ANNs vienen inspiradas por la biología, creando analogías entre elementos del cerebro humano y los componentes de las ANNs (hasta cierto límite). Así, nos encontramos ante el elemento base de una ANN: el *perceptrón*, que no es ni más ni menos que la representación lógica de una única neurona del cerebro y el sistema nervioso humanos. Como las neuronas, el perceptrón recibe entradas (como las *dendritas*), modifica esas entradas en base a algún criterio (como las *sinapsis*) y produce un resultado (como los *axones*).

Aunque el perceptrón es útil para resolver problemas lineares, no puede resolver problemas no lineares por sí solo; si no es posible clasificar el conjunto de datos del que partimos mediante una línea recta, el perceptrón fallará. Es por ello que las ANNs emplean el concepto de perceptrón a escala; utilizando múltiples perceptrones que trabajan juntos, es posible resolver problemas no lineares de varias dimensiones.

El **aprendizaje profundo** (del inglés, *deep learning*), que se deriva del aprendizaje automático, es una familia más amplia de técnicas y algoritmos que se utilizan para lograr una inteligencia limitada (débil), pero con el objetivo de alcanzar una inteligencia general (fuerte). Concretamente, el aprendizaje profundo es el nombre que reciben los algoritmos que utilizan ANNs en diferentes arquitecturas para lograr un objetivo. El aprendizaje profundo suele implicar que el enfoque intenta resolver un problema de forma más general, como el razonamiento espacial, o que se aplica a problemas que requieren más generalización, como la visión por computador y el reconocimiento del habla. El aprendizaje profundo también se basa en el aprendizaje supervisado, el aprendizaje no supervisado y el aprendizaje por refuerzo.

Los enfoques de aprendizaje profundo suelen emplear muchas capas de ANNs. Al aprovechar diferentes capas de componentes inteligentes, cada capa resuelve problemas especializados; juntas, las capas resuelven problemas complejos tendiendo a cumplir un objetivo mayor. Por ejemplo, identificar cualquier objeto en una imagen es un problema general, pero puede dividirse en la comprensión del color, el reconocimiento de las formas de los objetos y la identificación de las relaciones entre los objetos para lograr un objetivo.

0.4. Nuevas formas de interacción

Como se ha visto a lo largo de este capítulo, la IA puede lograrse a través de diferentes técnicas para lograr resolver problemas muy diversos. Sin embargo, es difícil distinguir entre lo que hay que saber para explotar sistemas de IA y los conocimientos necesarios para construir sistemas de IA.

Lo que hay que saber para gestionar proyectos de IA es lo mismo que se necesita saber con cualquier otro proyecto: cómo definir métricas y procesos que permitan comprender y supervisar adecuadamente la dirección y el éxito del proyecto. En general, se deben poder relacionar conceptos de IA con conceptos del negocio, siendo capaces de responder a las siguientes preguntas:

- ¿Qué puede hacer la IA y cómo puedo utilizarla en mi negocio?
- ¿Con qué tipo de proyecto de IA debería empezar primero?
- ¿Cómo voy a medir el éxito de la IA para ayudar a mi empresa?
- ¿Cómo debo gestionar un proyecto de IA?
- ¿Qué recursos son escasos y cuál es la mejor manera de asignarlos?

En el contexto de los negocios, una IA sería capaz de dirigir el negocio en las acciones concretas que deben realizarse para tener éxito. Para ello, es necesario recolectar datos para poder analizarlos posteriormente. Es en este paso donde utilizamos la IA para obtener información sobre las posibles opciones a realizar. Conocidas todas las opciones disponibles (aquellas obtenidas por la IA tras analizar los datos), es posible reaccionar para tomar alguna decisión sobre la próxima acción a realizar. Estos pasos se engloban dentro de un proceso iterativo para *sentir* (obtener datos), *analizar* (analizar los datos técnicamente) y *reaccionar* (llevar a cabo acciones concretas del dominio).

El último de los pasos del proceso anterior daría lugar a descubrir nuevas acciones del dominio gracias a la IA:

- *Sistema de apoyo a la toma de decisiones*: la inteligencia artificial ayuda a un empleado o directivo de su organización a tomar mejores decisiones. Los usos de estos sistemas van desde ayudar a la dirección a tomar decisiones que afectan a toda la organización hasta ayudar a los empleados en sus tareas cotidianas.
- *Producto mayor*: la IA podría ser sólo una parte de un producto más grande. Un producto de este tipo tiene capacidades que la IA puede hacer posibles, pero que no son capacidades propias de la IA. Un ejemplo de ello serían los robots de limpieza domésticos o los termostatos inteligentes. En el caso de un sistema totalmente autónomo, la IA guía el funcionamiento del sistema y toma sus decisiones sin necesidad de que intervenga el ser humano.
- *Automatización de procesos empresariales*: la IA automatiza algunos pasos del proceso empresarial. A veces se hace para reemplazar el trabajo humano; otras veces, se hace para procesar conjuntos de datos tan grandes que los humanos no pueden manejar.
- *IA como producto*: es posible empaquetar las herramientas de IA como un producto y venderlas a otras organizaciones. Un ejemplo sería un producto de IA capaz de reconocer señales de tráfico y que se vendería a los fabricantes de vehículos autónomos.

Para llevar a cabo dichas acciones es necesario conocer qué posibilidades ofrece la IA. Para ello, se introduce aquí la taxonomía presentada originalmente en los libros de Bill Schmarzo [Sch13, Sch15], cuyo objetivo principal es la de guiar la discusión entre el experto de IA y el experto de negocio:

- *Conocer los resultados más rápidamente*. En este caso, la IA serviría para descubrir un resultado más rápidamente, y eso tiene un valor comercial añadido en muchos escenarios. Por ejemplo, supongamos que diriges una planta de fabricación de coches y que estás ensamblando coches a partir de piezas que se fabrican en una zona de tu fábrica. Si sabes que alguna pieza del coche es defectuosa nada más fabricarla, puedes descartarla de inmediato y no instalarla nunca en el coche. Esto es mucho más preferible que enterarse de que la pieza es defectuosa después de haberla instalado en el coche y haberla enviado al cliente.

- *Predecir algún evento que ocurrirá en el futuro, basándose en las tendencias actuales.* Por ejemplo, estimando el precio del oro o los diamantes a largo plazo.
- *Utilizar datos estructurados.* A veces es posible encontrar la respuesta a alguna pregunta en una de las bases de datos relacionales ya existentes, especialmente en caso de contar con un gran volumen de datos. También hay métodos de IA que funcionan bien con datos que ya estén tabulados.
- *Utilizar datos no estructurados.* Los métodos de IA también pueden ayudar a procesar y comprender una gran cantidad de datos no estructurados, como texto, imágenes, vídeo y audio. En este caso, es posible utilizar métodos de IA para reconocer el contexto de una imagen, un vídeo o una grabación de audio.

Bibliografía

- [BKL09] Steven Bird, Ewan Klein, and Edward Loper. *Natural language processing with Python: analyzing text with the natural language toolkit*. O'Reilly Media, Inc., 2009.
- [BYRN99] R. Baeza-Yates and B. Ribeiro-Neto. *Modern Information Retrieval*. Addison Wesley, Boston, MA, USA, 1999.
- [HG10] Ralf Herbrich and Thore Graepel. *Handbook of Natural Language Processing, Second Edition*. Chapman and Hall/CRC, Machine Learning and Pattern Recognition Series, 2010.
- [Mil95] G. A. Miller. Wordnet: A lexical database for english. *Communications of the ACM*, 38:39–41, 1995.
- [Sar19] D. Sarkar. Python for natural language processing. In *Text Analytics with Python*. 2019.
- [Sch13] Bill Schmarzo. *Big Data: Understanding how data powers big business*. John Wiley & Sons, 2013.
- [Sch15] Bill Schmarzo. *Big Data MBA: Driving business strategies with data science*. John Wiley & Sons, 2015.