

Módulo 2. Fundamentos de la IA y su historia. IA generativa y LLMs



☰ Introducción

☰ 1. Historia de la IA

☰ 2. IA generativa

☰ Referencias

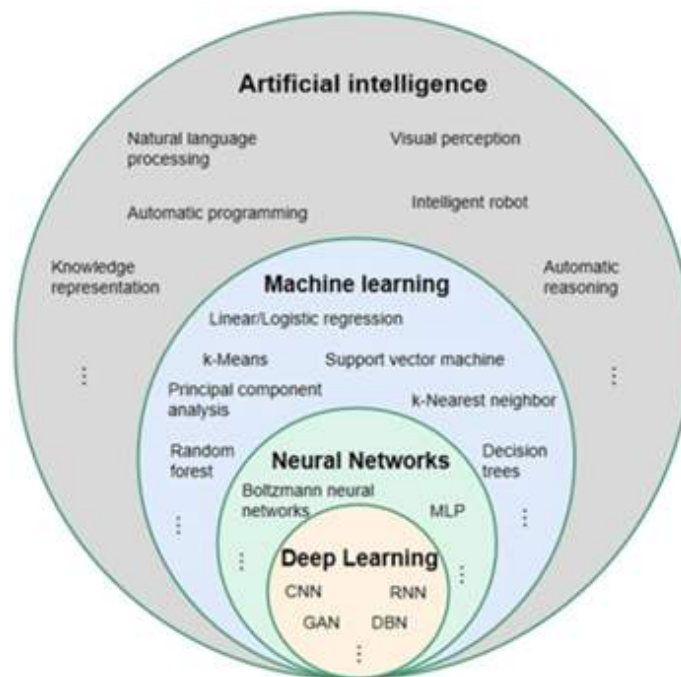
Introducción

El término inteligencia artificial ha desplazado en el uso común a expresiones más precisas empleadas en la literatura científica, como *machine learning*. El propio uso de la palabra «inteligencia» puede resultar engañoso. La inteligencia artificial (IA) consiste en la simulación de procesos propios de la inteligencia humana por parte de máquinas, basada en datos, algoritmos y alta capacidad de cómputo. Surgida formalmente en 1956, ha evolucionado desde sistemas lógicos y redes neuronales iniciales hasta el aprendizaje profundo y los modelos generativos actuales.

En cuanto a sus fundamentos, la IA se apoya en tres pilares principales: **datos, algoritmos y capacidad de cómputo**, ya que requiere grandes volúmenes de información para identificar patrones y mejorar su desempeño. El aprendizaje automático (*machine learning*) permite que las máquinas aprendan y mejoren sin ser programadas explícitamente para cada tarea. Las redes neuronales y el *deep learning*, inspirados en modelos propuestos por McCulloch y Pitts en 1943,

posibilitan el procesamiento de información compleja. Asimismo, disciplinas como el álgebra booleana, la probabilidad y la estadística constituyen la base matemática del razonamiento automatizado.

Figura 1. Relación jerárquica entre inteligencia artificial, machine learning, redes neuronales y deep learning



Fuente: Lior, 2024, <https://goo.su/FFjojS>

No resulta sorprendente que, cuando OpenAI publicó la primera versión de GPT (generative pretrained transformer), el modelo atrajera rápidamente una gran atención, al representar un avance significativo en la materialización de

este objetivo tecnológico. GPT-3 marcó un hito en el ámbito de la IA debido a su tamaño sin precedentes, con 175 000 millones de parámetros, lo que le permitió abordar una amplia gama de tareas de lenguaje natural sin requerir ajustes extensivos. El modelo fue entrenado con grandes volúmenes de datos, lo que posibilitó la generación de texto coherente y la interacción conversacional. Además, incorporó la capacidad de aprendizaje few-shot, lo que incrementó su versatilidad y favoreció su aplicación en entornos comerciales, como chatbots y asistentes virtuales.

En la actualidad, la IA se encuentra cada vez más integrada en múltiples aspectos de la vida cotidiana, desde las redes sociales hasta los procesos laborales. A medida que la tecnología continúa evolucionando, su influencia tiende a ampliarse. Comprender su desarrollo histórico permite analizar con mayor claridad las posibles direcciones futuras. Hoy se aplica en ámbitos como la agricultura —mediante drones y robótica—, el reconocimiento de voz y la generación de contenido, al tiempo que plantea desafíos éticos vinculados con sesgos, responsabilidad y uso adecuado de la tecnología.

¿Dónde está Chatgpt, Copilot, DeepSeek, etc?

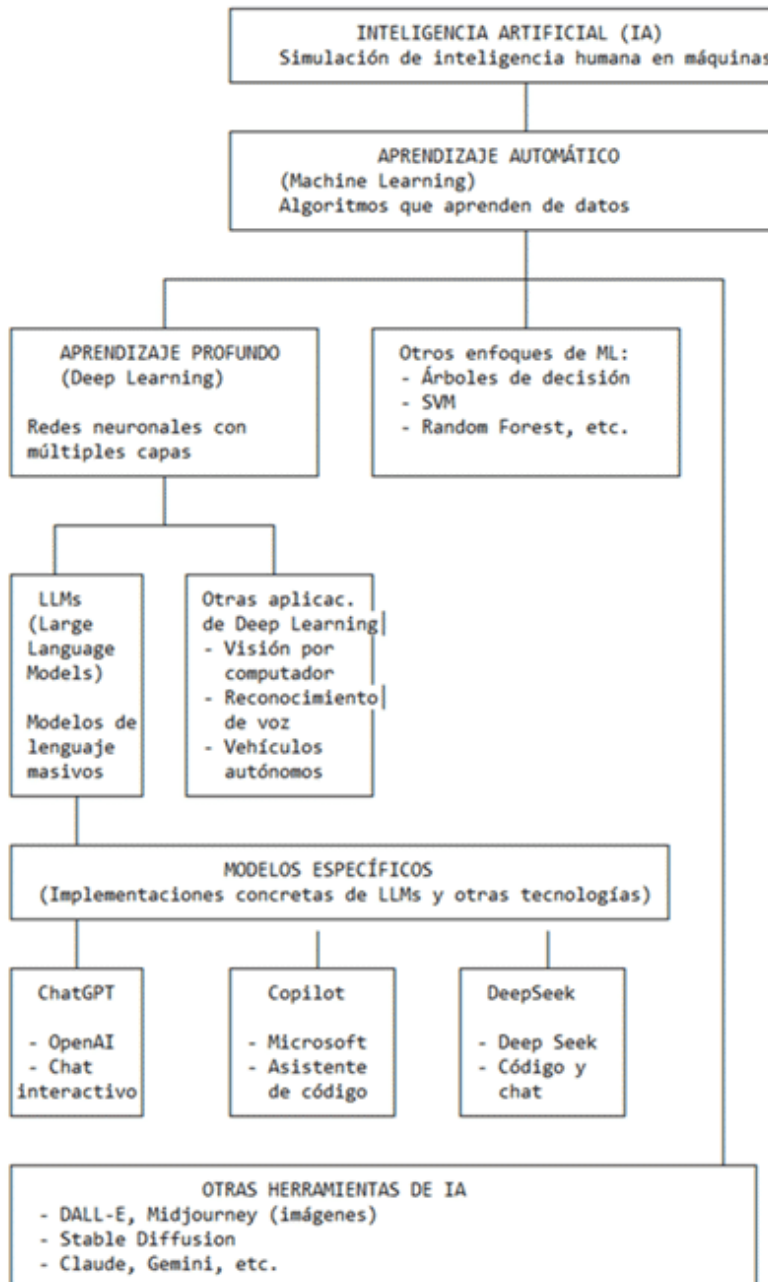
En la figura 1 quedó en claro dónde se ubican los conceptos de inteligencia artificial, *machine learning*, entre otros. Siguiendo esa misma línea, surge la pregunta acerca de dónde se localizan modelos como ChatGPT, Copilot o DeepSeek.

La figura 2 responde esta inquietud. Para facilitar su comprensión, puede establecerse una analogía con un vehículo:

- si la inteligencia artificial es el automóvil,
- el *machine learning* es el motor,
- el *deep learning* es el sistema de navegación GPS,
- y los LLM (*large language models*, modelos de lenguaje masivos) representan el conductor experto.

De este modo, herramientas como ChatGPT, Copilot o DeepSeek pueden entenderse como implementaciones concretas de modelos basados en LLM dentro del ecosistema más amplio de la inteligencia artificial.

Figura 2. Ubicación de ChatGPT, Copilot y DeepSeek dentro del ecosistema de inteligencia artificial



Fuente: elaboración propia

CONTINUAR

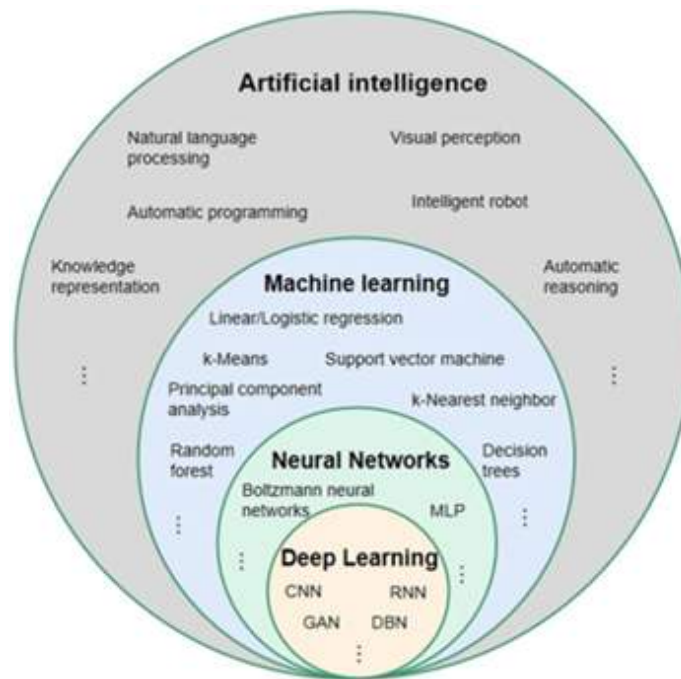
1. Historia de la IA

Según la nacionalidad del autor, puede observarse cierta tendencia a atribuir acontecimientos que justifiquen términos o desarrollos con el fin de señalar qué país o investigador fue el punto de partida de lo que hoy denominamos inteligencia artificial.

Con argumentos sustentados en la tradición oral, experiencias personales y aportes de distintos autores, en el apunte 4.1 de este curso avanzado se ejemplificó cómo nuestros antepasados, a partir de la experiencia acumulada, realizaban predicciones acertadas sobre el estado meteorológico del día siguiente. Este ejemplo se vincula directamente con el funcionamiento de los algoritmos de *machine learning*, que operan sobre un *dataset* para generar predicciones.

Como forma de reconocimiento a esa sabiduría empírica, en las conferencias se expone este caso acompañado por la figura 3.

Figura 3. Representación conceptual del uso temprano de predicción y su analogía con la IA en Argentina



Fuente: elaboración propia

1900–1950

1914. El ingeniero español Leonardo Torres y Quevedo presentó en la Exposición Universal de París la primera máquina capaz de jugar ajedrez de manera automatizada, denominada *El Ajedrecista*. Este dispositivo utilizaba electroimanes y funcionaba de forma completamente autónoma. *El Ajedrecista* estaba diseñado para jugar automáticamente un final sencillo de ajedrez —rey y torre contra rey—. Una vez puesto en funcionamiento, no requería intervención humana: ejecutaba jugadas

legales de manera autónoma y, si el oponente realizaba un movimiento ilegal, la máquina señalaba el error. Además, cuando alcanzaba una posición ganadora, era capaz de dar jaque mate al adversario de forma consistente.

1921. Se estrena en Londres la obra *R. U. R. (Robots Universales Rossum)*, de Karel Čapek. Es la primera vez que se utiliza la palabra «robot» en inglés. El término proviene del checo «robota», asociado al trabajo obligatorio o forzado de los campesinos en el sistema feudal. Tras el éxito de la obra, «robot» adquiere reconocimiento internacional y pasa a emplearse para designar seres artificiales creados para realizar tareas. Aunque los robots de Čapek eran orgánicos, el término terminó asociándose con máquinas mecánicas humanoides destinadas a trabajos repetitivos y no cualificados.

1939. John Vincent Atanasoff, profesor del Iowa State College, junto con su estudiante Clifford Berry, desarrolla la Atanasoff-Berry Computer (ABC) con una subvención de 650 dólares de la Universidad Estatal de Iowa. Utiliza sistema binario (0 y 1) para representar datos y emplea circuitos electrónicos en lugar de sistemas mecánicos o electromecánicos, permitiendo cálculos más rápidos y fiables. Aunque nunca fue plenamente operativa ni utilizada de forma generalizada, introdujo conceptos fundamentales: separación entre memoria y unidad de procesamiento, uso de condensadores para almacenar datos y capacidad para manejar hasta treinta ecuaciones simultáneamente.

1943. Warren S. McCulloch y Walter Pitts publican *A Logical Calculus of the Ideas Immanent in Nervous Activity (Un cálculo lógico de las ideas inmanentes a la actividad nerviosa)* en *Bulletin of Mathematical Biophysics*. Este trabajo constituye uno de los aportes más relevantes en la historia de la neurociencia y de la inteligencia artificial. El artículo establece

las bases de la idea de que el cerebro puede entenderse como un sistema computacional e introduce el concepto de redes neuronales artificiales, fundamento de desarrollos posteriores como las redes neuronales modernas y el *deep learning*.

1950. Se publica en la revista *Mind* el artículo *Computing Machinery and Intelligence*, del matemático británico Alan Turing. Este texto fundacional en el campo de la IA aborda la pregunta «¿Pueden pensar las máquinas?». Para evitar responderla de forma directa, Turing reformula el problema de manera operativa: determinar si una máquina puede exhibir un comportamiento indistinguible del de un ser humano. Propone entonces el denominado «juego de imitación», hoy conocido como prueba de Turing, que evalúa la capacidad de una máquina para imitar de forma convincente la conversación y el comportamiento humanos.

1950–1980

1951. Marvin Minsky y Dean Edmonds construyen la primera red neuronal artificial, denominada SNARC (*stochastic neural analog reinforcement calculator*). Este dispositivo constituye uno de los primeros intentos de modelar los procesos de aprendizaje del cerebro humano, específicamente mediante aprendizaje por refuerzo. La SNARC fue diseñada para simular el comportamiento de una rata al recorrer un laberinto, imitando la forma en que los animales aprenden a partir de recompensas y castigos, es decir, ajustando su comportamiento con base en la retroalimentación recibida.

Se trataba de una computadora analógica que utilizaba 3000 tubos de vacío y pesos sinápticos para simular 40 unidades similares a neuronas.

1952. Allen Newell, matemático e informático, y Herbert A. Simon, politólogo, desarrollan programas influyentes como *Logic Theorist* y *General Problem Solver*, considerados entre los primeros sistemas capaces de imitar la resolución humana de problemas mediante métodos computacionales.

1955. El término «inteligencia artificial» se acuña por primera vez en la propuesta de taller titulada *A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence* (una propuesta para el proyecto de investigación de verano de Dartmouth sobre inteligencia artificial), presentada por John McCarthy, Marvin Minsky, Nathaniel Rochester y Claude Shannon. El taller se llevó a cabo en julio y agosto de 1956 y se considera, de manera general, la fecha de nacimiento oficial del campo de la inteligencia artificial.

1957. Frank Rosenblatt desarrolla el perceptrón, una de las primeras redes neuronales artificiales capaces de reconocer patrones mediante una estructura de dos capas. El perceptrón introduce el concepto de clasificador binario que aprende a partir de los datos ajustando las ponderaciones de sus entradas mediante algoritmos de aprendizaje. Aunque estaba limitado a problemas linealmente separables, sentó las bases para el desarrollo posterior de redes neuronales y del *machine learning*.

1958. John McCarthy desarrolla el lenguaje de programación *Lisp*, acrónimo de LISt Processing. Este lenguaje surge de sus trabajos sobre la formalización de algoritmos y lógica matemática, con el objetivo de crear

un entorno capaz de manipular información simbólica. Lisp se convierte rápidamente en el lenguaje predominante en la investigación en inteligencia artificial durante varias décadas.

1959:

1. Arthur Samuel desarrolla un programa de damas capaz de mejorar su rendimiento con el tiempo, demostrando que una computadora puede aprender a partir de la experiencia y superar incluso a su programador. Con este trabajo acuña el término «*machine learning*» y establece las bases del aprendizaje automático basado en la experiencia.
2. Oliver Selfridge publica el artículo *Pandemonium: A Paradigm for Learning*. En él propone un modelo compuesto por múltiples «demonios» o unidades de procesamiento que compiten para reconocer patrones en datos no preprogramados, simulando un proceso de aprendizaje no supervisado. Su propuesta influye en desarrollos posteriores en reconocimiento de patrones y visión artificial.
3. John McCarthy introduce el concepto de «advice taker» en su artículo *Programs with Common Sense*. El sistema propuesto busca resolver problemas mediante la manipulación de enunciados en lógica formal, con el objetivo de sentar las bases del razonamiento automatizado. Este planteamiento impulsa las primeras investigaciones sobre representación del conocimiento y razonamiento en inteligencia artificial.

1965:

1. Hubert Dreyfus publica *Alchemy and Artificial Intelligence*, obra en la que sostiene que la mente humana funciona de manera fundamentalmente distinta a las computadoras. Señala límites al

progreso de la IA debido a la dificultad de replicar la intuición y la comprensión humanas. Su postura influye en los debates filosóficos y prácticos sobre los alcances de la inteligencia artificial.

2. I. J. Good escribe *Speculations Concerning the First Ultraintelligent Machine*, donde plantea que, una vez creada una máquina ultrainteligente, esta podría diseñar sistemas aún más inteligentes. Afirma que ello podría convertirla en el último invento de la humanidad, siempre que permanezca bajo control. Sus ideas anticipan debates contemporáneos sobre superinteligencia y riesgos asociados a la IA.

3. Joseph Weizenbaum desarrolla ELIZA, un programa capaz de simular conversación humana a partir de entradas en lenguaje natural. Aunque su intención era evidenciar la superficialidad del intercambio entre humanos y máquinas, le sorprendió la tendencia de los usuarios a atribuirle cualidades emocionales, lo que abrió interrogantes éticos sobre la interacción entre personas y sistemas artificiales.

4. Edward Feigenbaum, Bruce Buchanan, Joshua Lederberg y Carl Djerassi desarrollan DENDRAL en la Universidad de Stanford. Este sistema experto automatiza la toma de decisiones en química orgánica mediante la simulación de la formulación de hipótesis. Su éxito demuestra que los sistemas computacionales pueden realizar tareas altamente especializadas con un nivel comparable al de expertos humanos.

1966. Desarrollado en el SRI a finales de la década de 1960, Shakey se convierte en el primer robot móvil capaz de razonar sobre sus propias acciones, al combinar percepción, planificación y resolución de problemas.

1969:

1. Arthur Bryson y Yu-Chi Ho presentan la retropropagación, método destinado a optimizar sistemas dinámicos de múltiples etapas. Aunque fue concebido inicialmente para sistemas de control, el algoritmo se vuelve determinante en el entrenamiento de redes neuronales. Su impacto se consolida décadas más tarde, especialmente en los años 2000 y 2010, cuando el aumento de la capacidad de cómputo impulsa el desarrollo del *deep learning*.
2. Marvin Minsky y Seymour Papert publican *Perceptrons: An Introduction to Computational Geometry* («Perceptrones: una introducción a la geometría computacional»), obra en la que analizan de manera crítica las limitaciones de las redes neuronales de una sola capa. Con frecuencia se señala este trabajo como uno de los factores que redujeron el interés en las redes neuronales durante ese período. En la edición de 1988, los autores sostienen que el estancamiento ya se había producido por la falta de fundamentos teóricos sólidos, pese a la intensa experimentación realizada a mediados de la década de 1960.

1970. Terry Winograd desarrolla SHRDLU, un programa pionero en comprensión del lenguaje natural. SHRDLU interactúa con usuarios mediante lenguaje sencillo para manipular objetos en un entorno virtual compuesto por bloques. Este desarrollo demuestra el potencial de las computadoras para interpretar y ejecutar instrucciones complejas, aunque su funcionamiento se limita a entornos estructurados. El sistema evidencia tanto las posibilidades como las dificultades de alcanzar una comprensión lingüística más amplia en IA.

1972. Desarrollado en la Universidad de Stanford, MYCIN se convierte en uno de los primeros sistemas expertos diseñados para asistir a médicos en

el diagnóstico de infecciones bacterianas y en la recomendación de tratamientos antibióticos. Basado en reglas, simula el proceso de toma de decisiones de especialistas humanos y sienta las bases para el desarrollo de sistemas de IA médica. No obstante, debido a consideraciones éticas y legales, nunca se implementa en la práctica clínica.

1973. James Lighthill presenta un informe crítico al British Science Research Council sobre el progreso de la investigación en inteligencia artificial. En él concluye que el campo no había cumplido sus promesas iniciales ni producido avances significativos, lo que derivó en una drástica reducción de la financiación gubernamental en el Reino Unido. Este informe contribuye al inicio del denominado primer invierno de la IA, período caracterizado por la disminución del interés y de la inversión en la investigación en este ámbito.

1980–2000

1980. WABOT-2, robot humanoide desarrollado en la Universidad de Waseda (Japón), comienza a construirse en 1980 y se completa hacia 1984. Este proyecto sucede a WABOT-1, creado en 1973. Mientras que WABOT-1 se centraba en la movilidad y la comunicación básicas, WABOT-2 fue diseñado como un robot músico. Podía leer partituras mediante cámaras, conversar con personas, interpretar música en un órgano electrónico e incluso acompañar a un cantante humano. El desarrollo de WABOT-2 representa un avance en la creación de robots humanoides y sistemas de IA capaces

de ejecutar tareas complejas asociadas a habilidades humanas, incluida la expresión artística.

1982. Japón lanza el Proyecto de Sistemas Informáticos de Quinta Generación (FGCS), con el objetivo de desarrollar computadoras capaces de realizar razonamiento lógico y resolución de problemas, e impulsar la investigación en inteligencia artificial. El proyecto busca construir máquinas aptas para tareas como el procesamiento del lenguaje natural y los sistemas expertos. Aunque se interrumpe en 1992, sus resultados contribuyen al desarrollo de la programación lógica concurrente.

1984. En la reunión anual de la Asociación para el Avance de la Inteligencia Artificial (AAAI), Roger Schank y Marvin Minsky advierten sobre un posible «invierno de la IA». Señalan que las expectativas sobredimensionadas podrían derivar en una reducción de la inversión y de la investigación, situación similar a la ocurrida a mediados de la década de 1970. Tres años después, la disminución del interés y de la financiación confirma ese pronóstico, dando lugar al denominado segundo invierno de la IA.

1986. David Rumelhart, Geoffrey Hinton y Ronald Williams publican *Learning Representations by Back-Propagating Errors (Aprendizaje de representaciones mediante errores de retropropagación)*, artículo en el que describen el algoritmo de retropropagación aplicado a redes neuronales multicapa. Este método permite ajustar las ponderaciones internas propagando el error hacia atrás a través de la red, lo que mejora la capacidad de aprendizaje de patrones complejos. El trabajo se apoya en investigaciones previas de 1969 de Arthur Bryson y Yu-Chi Ho y contribuye a reactivar el interés en las redes neuronales. Este avance sienta las bases del *deep learning* moderno y posibilita su desarrollo práctico en las décadas de 2000 y 2010.

Este avance hace que las redes neuronales sean viables para aplicaciones prácticas y abre la puerta para la revolución del deep learning de las décadas de 2000 y 2010.

1987. Durante su discurso de apertura en Educom, el director ejecutivo de Apple, John Sculley, presenta el video *Knowledge Navigator*, en el que imagina un futuro donde agentes digitales inteligentes asisten a los usuarios en el acceso a grandes volúmenes de información a través de sistemas en red. En la presentación se muestra a un profesor interactuando con un asistente activado por voz capaz de recuperar datos, responder preguntas y mostrar información de lo que hoy se reconoce como Internet. El video anticipa elementos que posteriormente se materializarían en tecnologías como asistentes virtuales basados en IA, bases de conocimiento en red y el actual entorno digital interconectado. Puedes ver el video a continuación:

Fuente: Knownav [knownav]. *Knowledge Navigator* [video]. YouTube. https://www.youtube.com/watch?v=QRH8eimU_20

1988:

1. Judea Pearl publica *Probabilistic Reasoning in Intelligent Systems (Razonamiento probabilístico en sistemas inteligentes)*, obra que transforma la manera en que la inteligencia artificial aborda la incertidumbre. En ella introduce las redes bayesianas como formalismo para representar modelos probabilísticos complejos y desarrollar algoritmos de inferencia. Sus aportes permiten que los sistemas de IA tomen decisiones fundamentadas en entornos inciertos e influyen en múltiples disciplinas. En 2011 recibe el Premio Turing por su contribución a la base representativa y computacional del razonamiento probabilístico moderno.

2. Rollo Carpenter desarrolla Jabberwacky, uno de los primeros chatbots diseñados para simular conversaciones humanas de manera entretenida y espontánea. A diferencia de los sistemas basados exclusivamente en reglas, Jabberwacky aprende de las interacciones con los usuarios, lo que le permite generar diálogos más naturales y anticipa desarrollos posteriores en IA conversacional.
3. Investigadores del Centro de Investigación T. J. Watson de IBM publican *A Statistical Approach to Language Translation (Un enfoque estadístico de la traducción lingüística)*, marcando el paso de los métodos basados en reglas a los enfoques probabilísticos en traducción automática. El proyecto Candide utiliza 2,2 millones de pares de frases en inglés y francés, procedentes en gran medida de las actas del Parlamento canadiense. Este enfoque se centra en el análisis estadístico de grandes volúmenes de datos en lugar de intentar modelar explícitamente la comprensión lingüística, y sienta las bases para avances posteriores en procesamiento del lenguaje natural.
4. Marvin Minsky y Seymour Papert publican una edición ampliada de *Perceptrons*. En el nuevo prólogo, titulado *A View from 1988 (Una visión desde 1988)*, reflexionan sobre el lento progreso en la inteligencia artificial y señalan la necesidad de una comprensión teórica más profunda. Reafirman sus críticas iniciales a las redes neuronales de una sola capa y reconocen enfoques emergentes que posteriormente contribuirían al desarrollo del *deep learning*.

1989. Yann LeCun y un equipo de investigadores de AT&T Bell Labs aplican con éxito el algoritmo de retropropagación a una red neuronal multicapa para el reconocimiento de códigos postales escritos a mano. Este trabajo constituye una de las primeras aplicaciones prácticas del *deep learning*

mediante redes neuronales convolucionales. A pesar de las limitaciones del hardware de la época, el entrenamiento de la red requería aproximadamente tres días, lo que representaba una mejora significativa respecto de intentos anteriores. El desempeño alcanzado en el reconocimiento de dígitos manuscritos, tarea relevante para la automatización de servicios postales, evidenció el potencial de las redes neuronales en problemas de visión artificial y sentó las bases para el desarrollo del *deep learning* en las décadas posteriores.

1993. El autor de ciencia ficción y matemático Vernor Vinge publica el ensayo *The Coming Technological Singularity (La próxima singularidad tecnológica)*, en el que predice la creación de una inteligencia sobrehumana en un plazo aproximado de treinta años, capaz de transformar de manera profunda la civilización humana. Vinge sostiene que los avances tecnológicos, especialmente en inteligencia artificial, conducirán a una explosión de inteligencia —máquinas que superen la capacidad humana— y al fin de la era humana tal como se conoce. El ensayo populariza el concepto de «singularidad tecnológica», entendido como el momento en que la IA superaría el control humano, lo que genera debate en los ámbitos de la IA, la ética y el futurismo.

1995. Richard Wallace desarrolla el chatbot A.L.I.C.E. (*Artificial Linguistic Internet Computer Entity, Entidad informática lingüística artificial de Internet*), a partir de las bases establecidas por ELIZA de Joseph Weizenbaum. A diferencia de ELIZA, que se apoyaba en respuestas predefinidas para simular conversación, A.L.I.C.E. aprovecha la emergente World Wide Web para recopilar y procesar grandes volúmenes de datos en lenguaje natural, lo que le permite sostener diálogos más complejos. Utiliza una técnica de concordancia de patrones denominada AIML (*Artificial Intelligence Markup Language*), lo que la hace más adaptable que sus

predecesoras. El trabajo de Wallace contribuye al desarrollo posterior de la IA conversacional e influye en asistentes virtuales y chatbots modernos.

1997:

1. Sepp Hochreiter y Jürgen Schmidhuber presentan la memoria a largo y corto plazo (LSTM), un tipo de red neuronal recurrente (RNN) diseñada para superar las limitaciones de las RNN tradicionales, en particular su dificultad para capturar dependencias a largo plazo en los datos. Las redes LSTM se aplican en reconocimiento de escritura manuscrita, reconocimiento de voz, procesamiento del lenguaje natural y predicción de series temporales.
2. Deep Blue, desarrollado por IBM, derrota al campeón mundial de ajedrez Garry Kasparov en un enfrentamiento de seis partidas. Es la primera vez que un programa informático vence a un campeón mundial bajo condiciones oficiales de torneo. La capacidad del sistema para calcular millones de posiciones por segundo, junto con avances en teoría de juegos y heurísticas de búsqueda, le permite superar a Kasparov y consolidar un hito en la historia de la inteligencia artificial.

1998:

1. Dave Hampton y Caleb Chung crean Furby, una de las primeras mascotas robóticas domésticas de gran éxito comercial. Furby responde al tacto, al sonido y a la luz, y simula un proceso de aprendizaje lingüístico: comienza comunicándose en su idioma propio, denominado furbish, y progresivamente incorpora más inglés a medida que interactúa con los usuarios. Su capacidad para imitar el aprendizaje y generar vínculo con las personas lo convierte en un

antecedente de los robots sociales orientados al entretenimiento y al consumo masivo.

2. Yann LeCun, Yoshua Bengio y sus colaboradores publican trabajos influyentes sobre la aplicación de redes neuronales al reconocimiento de escritura manuscrita. La investigación se centra en el uso de redes neuronales convolucionales para optimizar el algoritmo de retropropagación y mejorar el entrenamiento de redes profundas. Al perfeccionar este proceso y demostrar la eficacia de las CNN en el reconocimiento de imágenes y patrones, sientan las bases de las técnicas modernas de *deep learning* utilizadas actualmente en múltiples aplicaciones de inteligencia artificial.

2000–2020

2000. Cynthia Breazeal, investigadora del MIT, desarrolla Kismet, un robot diseñado para interactuar con seres humanos mediante señales emocionales y sociales. Kismet cuenta con cámaras, micrófonos y rasgos faciales expresivos que le permiten percibir y responder a emociones como la felicidad, la tristeza y la sorpresa. Este desarrollo representa un avance en la robótica social, orientada a estudiar cómo los robots pueden interactuar con las personas de manera más natural.

2006. Geoffrey Hinton publica *Learning Multiple Layers of Representation* («Aprendizaje de múltiples capas de representación»), trabajo en el que sintetiza avances relevantes en *deep learning* y describe métodos más eficaces para entrenar redes neuronales multicapa. Su investigación se centra en entrenar redes con múltiples capas para generar

representaciones jerárquicas de datos sensoriales, lo que marca la transición conceptual hacia lo que actualmente se denomina *deep learning*.

2007. Fei-Fei Li y su equipo, entonces en la Universidad de Princeton, inician el proyecto ImageNet, destinado a crear una de las bases de datos más extensas de imágenes anotadas. ImageNet proporciona millones de imágenes etiquetadas en miles de categorías y favorece el desarrollo de sistemas de reconocimiento visual. La magnitud y calidad del conjunto de datos impulsan avances significativos en visión artificial y en el entrenamiento de modelos de *deep learning* para la clasificación de objetos.

2009:

1. Rajat Raina, Anand Madhavan y Andrew Ng publican *Large-Scale Deep Unsupervised Learning Using Graphics Processors* («Aprendizaje profundo no supervisado a gran escala mediante procesadores gráficos»), donde sostienen que las unidades de procesamiento gráfico (GPU) superan ampliamente a las CPU multinúcleo tradicionales en tareas de *deep learning*. Demuestran que la mayor capacidad de cómputo de las GPU permite entrenar modelos más grandes y complejos de manera más eficiente, lo que impulsa la adopción de esta tecnología y favorece los avances de la década de 2010 en campos como la visión artificial y el procesamiento del lenguaje natural.
2. Investigadores del Laboratorio de Información Inteligente de la Universidad Northwestern desarrollan Stats Monkey, un sistema capaz de generar automáticamente crónicas deportivas sin intervención humana. A partir de estadísticas de los partidos, el programa produce

narraciones coherentes sobre encuentros de béisbol, incluyendo resúmenes, desempeño de los jugadores y análisis del juego.

2011:

1. Watson, desarrollado por IBM, alcanza notoriedad al competir en el programa televisivo *Jeopardy!* frente a dos de sus campeones más destacados, Ken Jennings y Brad Rutter, y vencerlos. Diseñado como un sistema avanzado de respuesta a preguntas en lenguaje natural, Watson demuestra una notable capacidad para procesar, interpretar y analizar grandes volúmenes de información en tiempo real. Su desempeño evidencia avances significativos en la comprensión y generación de lenguaje por parte de la inteligencia artificial.
2. Apple lanza Siri, un asistente virtual integrado en el sistema operativo iOS. Siri incorpora una interfaz basada en lenguaje natural que permite interactuar con el dispositivo mediante comandos de voz. Puede enviar mensajes, establecer recordatorios, ofrecer recomendaciones y responder preguntas. El sistema emplea *machine learning* para adaptarse a las preferencias y patrones de voz de cada usuario, lo que mejora la experiencia personalizada y marca un avance en la accesibilidad de los asistentes basados en inteligencia artificial.

2012:

1. Jeff Dean y Andrew Ng llevan a cabo un experimento con una red neuronal de gran escala entrenada con 10 millones de imágenes sin etiquetar extraídas de videos de YouTube. Durante el proceso, la red aprende de manera autónoma a identificar patrones en los datos y, de forma inesperada, una de sus neuronas se especializa en detectar

imágenes de gatos. Este resultado constituye una demostración del aprendizaje no supervisado y evidencia la capacidad de las redes profundas para extraer características relevantes a partir de grandes volúmenes de datos sin intervención humana directa.

2. Investigadores de la Universidad de Toronto, liderados por Geoffrey Hinton, desarrollan una red neuronal convolucional denominada AlexNet, que obtiene un resultado destacado en el ImageNet Large Scale Visual Recognition Challenge. El modelo alcanza una tasa de error del 16 %, frente al 25 % registrado por el mejor sistema del año anterior. Este logro representa un punto de inflexión en la visión artificial, al demostrar que las redes neuronales convolucionales pueden superar a los métodos tradicionales de clasificación de imágenes cuando se entrenan con conjuntos de datos masivos.

2016:

1. AlphaGo, desarrollado por Google DeepMind, derrota por 4-1 a Lee Sedol, uno de los mejores jugadores de Go del mundo. El Go, considerado durante años un desafío para la inteligencia artificial debido a la enorme cantidad de combinaciones posibles, representaba un reto estratégico de alta complejidad. La victoria de AlphaGo constituye un hito en la historia de la IA y demuestra la eficacia de las técnicas de *deep learning* para abordar problemas estratégicos que hasta entonces se consideraban fuera de su alcance.
2. Hanson Robotics presenta a Sophia, un robot humanoide avanzado capaz de reconocer rostros, establecer contacto visual y mantener conversaciones mediante la integración de reconocimiento de imágenes y procesamiento del lenguaje natural.

2017. Investigadores del laboratorio de Investigación en Inteligencia Artificial de Facebook (FAIR) entrenan a dos chatbots para negociar entre sí. Aunque estaban programados para comunicarse en inglés, durante el proceso comienzan a apartarse del lenguaje humano estructurado y desarrollan una forma abreviada propia para intercambiar información con mayor eficiencia. Este comportamiento no había sido previsto, ya que los sistemas optimizan su comunicación sin intervención directa. El experimento se detiene con el fin de mantener el intercambio dentro de un lenguaje comprensible para los humanos, pero el episodio pone de manifiesto la capacidad de los sistemas de inteligencia artificial para generar dinámicas emergentes no anticipadas.

2020:

1. OpenAI presenta GPT-3, un modelo de lenguaje con 175 000 millones de parámetros, que se convierte en uno de los sistemas más grandes y sofisticados desarrollados hasta ese momento. GPT-3 demuestra capacidad para generar texto coherente, mantener conversaciones, escribir código, traducir idiomas y producir contenidos creativos a partir de instrucciones en lenguaje natural. Como uno de los primeros modelos de lenguaje de gran tamaño (LLM), su escala le permite realizar múltiples tareas lingüísticas con escaso ajuste específico, evidenciando el potencial de la IA para comprender y producir lenguaje de manera altamente consistente.
2. AlphaFold 2, desarrollado por DeepMind, representa un avance significativo en biología al predecir con alta precisión la estructura tridimensional de proteínas a partir de sus secuencias de aminoácidos. Este logro aborda un problema científico planteado durante décadas: comprender el plegamiento proteico. La precisión alcanzada tiene implicancias en la investigación biomédica y en el desarrollo de

fármacos, al facilitar el estudio de mecanismos moleculares asociados a enfermedades y la identificación de posibles estrategias terapéuticas.

2021 - Presente

2021:

1. Google presenta MUM (*Multitask Unified Model*), un modelo de inteligencia artificial diseñado para mejorar la experiencia de búsqueda mediante la comprensión y generación de lenguaje en 75 idiomas. MUM puede realizar múltiples tareas de forma simultánea y procesar información en distintos formatos, como texto, imágenes y videos. Esta capacidad multimodal le permite abordar consultas complejas que requieren integrar diversas fuentes de información y ofrecer respuestas contextualizadas.
2. Tesla lanza la versión beta de Full Self-Driving (FSD), un sistema avanzado de asistencia a la conducción orientado a la automatización vehicular. El sistema utiliza *deep learning* y redes neuronales para operar en entornos complejos, como calles urbanas, autopistas e intersecciones. Permite que el vehículo realice maniobras como dirigir, acelerar y frenar de manera automatizada en determinadas condiciones. Aunque representa un avance hacia la conducción autónoma, su implementación generalizada continúa sujeta a desafíos regulatorios y de seguridad.

2021-2023. OpenAI lanza DALL-E, seguido de DALL-E 2 y DALL-E 3, modelos de IA generativa capaces de producir imágenes detalladas a partir de

descripciones textuales. Estos sistemas emplean técnicas avanzadas de *deep learning* y arquitecturas basadas en transformadores para generar imágenes complejas, realistas o de carácter artístico a partir de instrucciones en lenguaje natural. Las versiones posteriores amplían las posibilidades de creación visual asistida por IA y permiten transformar ideas en representaciones gráficas sin requerir conocimientos técnicos especializados en diseño.

2024

Febrero

1. Google lanza Gemini 1.5 en versión beta limitada, un modelo de lenguaje capaz de manejar contextos de hasta un millón de *tokens*. El sistema puede procesar y comprender grandes volúmenes de información en una única instrucción, lo que incrementa su capacidad para mantener coherencia en conversaciones y tareas complejas basadas en textos extensos. Gemini 1.5 representa un avance en procesamiento del lenguaje natural al ampliar la memoria contextual sobre entradas largas.
2. OpenAI anuncia públicamente Sora, un modelo de texto a video capaz de generar secuencias de hasta un minuto de duración a partir de descripciones textuales. Esta innovación extiende la generación de contenidos por IA más allá de las imágenes estáticas y permite crear clips dinámicos a partir de instrucciones en lenguaje natural.
3. Stability AI presenta Stable Diffusion 3, su nuevo modelo de generación de imágenes a partir de texto. Al igual que otros sistemas generativos recientes, utiliza arquitecturas avanzadas para producir contenidos visuales detallados y creativos a partir de descripciones escritas.

Mayo

1. Google DeepMind presenta una extensión de AlphaFold orientada a la identificación de cáncer y enfermedades genéticas. Esta herramienta amplía las aplicaciones biomédicas del modelo y ofrece nuevas posibilidades en diagnóstico genético y medicina personalizada.
2. IBM presenta la familia Granite de modelos de IA generativa dentro de su plataforma watsonx. Con tamaños que oscilan entre 3 000 y 34 000 millones de parámetros, los modelos Granite están diseñados para tareas como generación de código, previsión de series temporales y procesamiento documental. Son de código abierto, se distribuyen bajo licencia Apache 2.0 y se caracterizan por su eficiencia y posibilidad de personalización en entornos empresariales.

Junio. Apple anuncia Apple Intelligence, una integración de ChatGPT en los nuevos iPhone y en Siri. Esta incorporación permite realizar tareas más complejas, mantener interacciones más naturales y comprender instrucciones con mayor nivel de matiz contextual.

Septiembre. NotebookLM presenta DeepDive, una herramienta de IA multimodal capaz de transformar materiales de origen en presentaciones de audio estructuradas con formato de pódcast. El sistema analiza y resume información procedente de diversos formatos, como páginas web, texto, audio y video, lo que amplía las posibilidades de generación automatizada de contenidos en ámbitos educativos y de producción multimedia.

2025

En el Foro Económico Mundial de Davos (Suiza), la inteligencia artificial ocupa un lugar central en la agenda. Durante el encuentro resuenan las

posiciones de quienes suscribieron el documento *Rome Call for AI Ethics* («Llamamiento de Roma por la ética de la IA»), promovido por el papa Francisco en 2020. Entre los puntos debatidos se destacan los posibles efectos adversos derivados de la profundización de la polarización social y de la ampliación de la brecha de habilidades.

El director ejecutivo de Anthropic advierte que la IA podría igualar o incluso superar capacidades humanas en un plazo no mayor a cinco años, mientras que Demis Hassabis, director ejecutivo de Google DeepMind, adopta una postura de optimismo prudente.

Satya Nadella, director ejecutivo de Microsoft, sostiene que el verdadero impacto de la IA radica en su difusión masiva, capaz de modificar la curva de productividad global y generar bienestar, siempre que los beneficios se distribuyan de manera equitativa. Por su parte, Jensen Huang, fundador de Nvidia, afirma que la IA dejó de ser una curiosidad tecnológica gracias a tres avances: modelos más fiables, el impulso de los modelos abiertos y el desarrollo de la denominada «inteligencia física».

En consonancia con lo expresado por Ursula von der Leyen acerca de la soberanía tecnológica, se destaca la importancia de fortalecer capacidades propias en cada país. En Argentina,

ello se refleja en iniciativas como la patente de un software de IA desarrollado para el Departamento Forense de la Gendarmería Nacional y en proyectos en el ámbito militar vinculados con investigaciones de posgrado.

La magnitud del cambio que introduce la IA puede observarse en que una persona sin conocimientos de programación es capaz de desarrollar formularios web con persistencia en bases de datos en pocos minutos, incorporando además mecanismos básicos de ciberseguridad. Quienes trabajan en el ámbito de la inteligencia artificial enfrentan el desafío de sostener una posición ética firme, orientada al bien común y al desarrollo responsable de la tecnología (Agüero, 2026).

Figura 3. Foro Económico Mundial, Reunión Anual en Davos (19–23 de enero de 2026)



Fuente: Agüero, 2026, <https://goo.su/6l0fm>

Las tendencias actuales de la inteligencia artificial apuntan a nuevas evoluciones de la IA generativa, que opera sobre modelos fundacionales más pequeños y eficientes, así como al desarrollo de la IA agentiva, en la que modelos específicos colaboran entre sí para resolver con mayor rapidez las solicitudes de los usuarios.

En el futuro, los vehículos autónomos circularán por las carreteras, la IA multimodal integrará audio, video, texto e imágenes en una misma plataforma, y los asistentes de IA acompañarán a las personas en la gestión de su vida personal y profesional.

CONTINUAR

2. IA generativa

La IA generativa es la inteligencia artificial (IA) que puede crear contenidos originales (como texto, imágenes, vídeo, audio o código de software) en respuesta a las instrucciones o peticiones de un usuario.

La IA generativa es un tipo de inteligencia artificial capaz de crear contenidos originales —como texto, imágenes, video, audio o código de software— a partir de instrucciones proporcionadas por el usuario.

Se fundamenta en modelos avanzados de *machine learning*, particularmente en algoritmos de *deep learning*, que buscan emular ciertos procesos de aprendizaje y toma de decisiones asociados al cerebro humano. Estos modelos identifican y codifican patrones y relaciones en grandes volúmenes de datos y utilizan esa información para interpretar solicitudes en

lenguaje natural y generar nuevos contenidos coherentes y pertinentes.

Los LLM (*large language models*) constituyen un subconjunto específico de la IA generativa orientado exclusivamente al procesamiento y generación de lenguaje humano.

Con el fin de precisar la relación entre la IA generativa y los modelos de lenguaje de gran tamaño (LLM), se presenta a continuación una comparación que permite distinguir su alcance, tipo de salida, objetivo y ejemplos representativos. Mientras la IA generativa constituye una categoría amplia orientada a la creación de contenidos en múltiples formatos, los LLM representan un subconjunto especializado en el procesamiento y generación de lenguaje natural.

Tabla 1. Diferencias entre IA generativa y modelos de lenguaje (LLM)

Característica	IA Generativa	LLMs (Modelos de Lenguaje)
Alcance	Categoría general (el "paraguas").	Categoría específica (un «tipo de»).

Salida (output)	Multimodal: imágenes, música, código, video.	Principalmente texto y código.
Objetivo	Crear cualquier tipo de contenido original.	Comprender y generar lenguaje natural.
Ejemplos	Midjourney, DALL-E, Sora.	GPT-4, Llama, Claude, Gemini.

Fuente: elaboración propia

¿Cómo se relacionan?

La relación entre la IA generativa y los modelos de lenguaje de gran tamaño (LLM) puede entenderse en términos de inclusión y especialización. Todos los LLM forman parte de la IA generativa, ya que generan contenido nuevo a partir de datos y de instrucciones en lenguaje natural. Sin embargo, no toda la IA generativa se limita a los LLM, dado que también existen modelos diseñados para producir imágenes, audio, video o combinaciones multimodales.

En la práctica, los LLM actúan como modelos fundacionales para numerosas aplicaciones actuales, debido a su capacidad para comprender contexto, razonar sobre instrucciones y generar respuestas coherentes. A partir de estos modelos base se construyen asistentes virtuales, sistemas de automatización, herramientas de generación de código y múltiples soluciones empresariales.

Asimismo, muchos sistemas contemporáneos adoptan un enfoque multimodal, en el que un LLM se integra con otros modelos especializados. De este modo, el sistema no solo puede generar texto, sino también procesar imágenes, interpretar audio o combinar distintos tipos de información en una misma interacción.

¿Cómo funciona la IA generativa?

En términos generales, la IA generativa opera a través de tres fases principales que permiten construir, especializar y perfeccionar los modelos.

En primer lugar, se lleva a cabo el **entrenamiento**, etapa en la que se crea un modelo fundacional a partir del procesamiento de grandes volúmenes de datos. Durante este proceso, el

modelo aprende patrones, estructuras y relaciones estadísticas que le permiten generar contenido coherente.

En segundo lugar, se realiza el **ajuste** (*fine-tuning*), mediante el cual el modelo fundacional se adapta a una aplicación específica. Esta fase permite especializar el comportamiento del sistema en determinados dominios, tareas o contextos, mejorando su precisión y pertinencia.

Finalmente, se desarrolla la etapa de **generación, evaluación y reajuste**, en la que el sistema produce resultados, estos se analizan y, en función de su desempeño, se realizan ajustes adicionales. Este proceso iterativo permite mejorar progresivamente la calidad, coherencia y confiabilidad del contenido generado.

Formación

La IA generativa se inicia a partir de un **modelo fundacional**, es decir, un modelo de *deep learning* que sirve como base para múltiples aplicaciones. En la actualidad, los modelos fundacionales más difundidos son los modelos de lenguaje de gran tamaño (LLM), orientados a la generación de texto. No obstante, también existen modelos fundacionales destinados a la generación de imágenes, video, sonido y música, así como

modelos multimodales capaces de trabajar simultáneamente con distintos tipos de contenido.

La construcción de un modelo fundacional requiere entrenar un algoritmo de *deep learning* con volúmenes masivos de datos sin procesar, no estructurados y sin etiquetar, como terabytes de información obtenida de internet u otras fuentes extensas. Durante el entrenamiento, el algoritmo ejecuta millones de tareas basadas en la predicción del siguiente elemento en una secuencia —por ejemplo, la siguiente palabra en una oración, el siguiente píxel en una imagen o el siguiente comando en una línea de código— y ajusta sus parámetros de manera continua para reducir la diferencia entre sus predicciones y los resultados reales.

El producto final de este proceso es una red neuronal compuesta por millones o miles de millones de parámetros, que representan entidades, patrones y relaciones aprendidas a partir de los datos, y que permiten generar contenido de manera autónoma ante nuevas instrucciones.

Este proceso implica un elevado costo computacional, ya que demanda miles de unidades de procesamiento gráfico (GPU) funcionando en paralelo durante semanas, lo que representa inversiones de millones de dólares. Por esta razón, los proyectos de modelos fundacionales de código abierto, como

Llama 2 de Meta, permiten a desarrolladores y organizaciones reutilizar modelos previamente entrenados, evitando asumir directamente esos costos iniciales.

Ajuste

De manera metafórica, un modelo fundacional puede entenderse como un generalista: posee conocimientos amplios sobre distintos tipos de contenido, pero no siempre genera resultados específicos con el nivel de precisión o fidelidad requerido. Para alcanzar ese grado de especialización, es necesario ajustarlo a una tarea concreta de generación.

Fine tuning

El *fine-tuning* consiste en entrenar nuevamente el modelo con datos etiquetados propios de la aplicación deseada. Estos datos incluyen ejemplos de instrucciones o preguntas que el sistema podría recibir, junto con las respuestas correctas en el formato esperado. Por ejemplo, si se desarrolla un chatbot de atención al cliente, se elaboran numerosos documentos que contienen consultas frecuentes debidamente etiquetadas y sus respuestas correspondientes. El modelo se entrena con ese conjunto específico para mejorar su desempeño en ese dominio.

Este proceso exige un trabajo intensivo de preparación y etiquetado de datos. Con frecuencia, las organizaciones externalizan esta tarea a equipos especializados en anotación y curación de información.

Aprendizaje por refuerzo a partir de la retroalimentación humana (RLHF)

En el aprendizaje por refuerzo a partir de la retroalimentación humana, los usuarios evalúan los contenidos generados por el modelo, y esas evaluaciones se utilizan para ajustarlo y mejorar su precisión o relevancia. En muchos casos, el RLHF implica que personas comparen distintas respuestas producidas ante una misma instrucción y las clasifiquen según su calidad. Sin embargo, el proceso también puede ser más simple: basta con que los usuarios interactúen con un chatbot o asistente virtual y corrijan sus respuestas, aportando así información que permita optimizar el comportamiento del sistema.

Generación, evaluación, más ajuste

Los desarrolladores y los usuarios evalúan de manera permanente los resultados producidos por las aplicaciones de IA generativa y realizan ajustes adicionales para mejorar su

precisión y relevancia. Estas modificaciones pueden efectuarse con relativa frecuencia, incluso semanalmente. En cambio, el modelo fundacional subyacente suele actualizarse con menor periodicidad, por ejemplo, una vez al año o cada dieciocho meses.

Otra estrategia para optimizar el desempeño de una aplicación es la generación aumentada por recuperación (retrieval-augmented generation, RAG). Este enfoque permite complementar el modelo fundacional con fuentes externas a los datos utilizados en su entrenamiento original, incorporando información adicional que amplía y actualiza su base de conocimiento.

La RAG posibilita que la aplicación acceda a información más reciente y específica. Además, las fuentes externas empleadas en este proceso pueden ser identificables para el usuario, a diferencia del conocimiento interno del modelo fundacional, cuya trazabilidad no siempre resulta transparente.

Arquitecturas de modelos de IA generativa y su evolución

Los modelos de IA generativa —entendidos como modelos de deep learning capaces de crear contenido de manera autónoma a partir de una instrucción— han experimentado una evolución significativa durante la última década. A lo largo de este período, diversas arquitecturas han marcado hitos en el desarrollo del campo.

Entre ellas se encuentran los autocodificadores variacionales (variational autoencoders, VAE), que impulsaron avances en reconocimiento de imágenes, procesamiento del lenguaje natural y detección de anomalías.

Posteriormente, las redes generativas adversativas (generative adversarial networks, GAN) y los modelos de difusión incrementaron la calidad y precisión de los resultados, posibilitando algunas de las primeras aplicaciones capaces de generar imágenes fotorrealistas.

En la actualidad, los transformadores constituyen la arquitectura dominante. Esta estructura de deep learning sustenta los modelos fundacionales y las principales soluciones de IA generativa contemporáneas, al permitir el procesamiento eficiente de grandes volúmenes de datos y la generación coherente de contenido en múltiples formatos.

CONTINUAR

Referencias

Agüero, W. (2026). IA en la mesa del Foro Económico Mundial, Davos (Suiza). *LinkedIn*.

<https://www.linkedin.com/feed/update/urn:li:activity:7422314353359663104/>

Knownav [knownav]. *Knowledge Navigator* [video]. YouTube.

https://www.youtube.com/watch?v=QRH8eimU_20

Lior, A. (2024). The difference between Artificial Intelligence (AI), Machine Learning (ML), Deep Learning (DL). *LinkedIn*.

https://www.linkedin.com/posts/pxlpainter_the-difference-between-artificial-intelligence-activity-7042868957246078976-OIB7/?originalSubdomain=cr

Referencias bibliográficas de consulta

IBM, (s.f.). ¿Qué es la IA generativa? <https://www.ibm.com/es-es/think/topics/generative-ai>

Innovación Educativa, (s.f.). *Cómo funciona la IA generativa.*
<https://innovacioneducativa.upm.es/sites/default/files/saga/presentacion-taller1-como-funciona-ia-generativa.pdf>

CONTINUAR