CIENCIA

launion.com.mx

@uniondemorelos

Inteligencia artificial para todos: ¿Cómo funcionan los modelos que conversan contigo?

VALERIA MAEDA GUTIÉRREZ Y JUAN JOSÉ OROPEZA VALDEZ

La Dra. Maeda-Gutiérrez es Docente-Investigadora en la Unidad Académica de Ingeniería Eléctrica de la Universidad Autónoma de Zacatecas. Su trabajo se centra en el desarrollo y aplicación de modelos de Inteligencia Artificial, Procesamiento de Imágenes y Aprendizaje Profundo en el ámbito biomédico, con énfasis en el estudio de complicaciones microvasculares asociadas a la Diabetes Tipo 2. Es miembro del Sistema Nacional de Investigadoras e Investigadores (SNII) nivel C en el área de Medicina y Ciencias de la Salud.

El Dr. Oropeza-Valdez es Investigador postdoctoral en el Centro de Ciencias de la Complejidad (C3) de la Universidad Nacional Autónoma de México, donde trabaja en el Laboratorio de Biología de Sistemas Humanos. Realiza investigación que utiliza Inteligencia Artificial para analizar la microbiota y datos metabolómicos, con el objetivo de desentrañar las complejas interacciones entre los sistemas del huésped y las comunidades microbianas, desarrollando así estrategias diagnósticas y terapéuticas innovadoras en medicina personalizada.

Esta publicación fue revisada por el comité editorial de la Academia de Ciencias de Morelos.

Introducción

Imagina que estás hablando con alguien que puede explicarte qué es un agujero negro, ayudarte con la tarea de matemáticas o redactar un informe sobre avances en medicina en segundos. No es un genio encerrado en una lámpara, sino un modelo de inteligencia artificial como ChatGPT. Estos modelos funcionan como programas que reciben una entrada (el *prompt*) y siguen instrucciones aprendidas para producir una salida. Un buen *prompt* aporta datos, contexto y una pregunta clara. ¿Cómo es posible que una máquina pueda conversar, escribir historias o responder preguntas como si fuera humana? La respuesta está en algo que suena complicado pero que podemos entender juntos: los *Modelos Grandes de Lenguaje* (en inglés, Large Language Models o LLMs). Estos son modelos matemáticos entrenados que han leído millones de textos y aprendido a imitar cómo hablamos los seres humanos. Aunque carecen de emociones o conciencia, nos sorprenden por su capacidad para generar respuestas coherentes, creativas y a veces hasta científicas. En las siguientes líneas descubrirás con nosotros cómo funciona esta tecnología, que ya utilizas sin darte cuenta: cuando tu teléfono predice qué palabra escribirás en WhatsApp, siempre que Gmail autocompleta tus correos electrónicos, o cuando preguntas algo rápido a asistentes como Siri o Alexa. Entender la Inteligencia Artificial (IA) no es un lujo reservado a especialistas, es una forma de prepararnos para un futuro en donde convivir con estas herramientas será parte de lo cotidiano.

La inteligencia artificial tiene muchas caras

Cuando escuchamos hablar de la IA, lo primero que nos viene a la mente puede ser una máquina, un robot o un programa que traduce idiomas. Todas esas ideas están relacionadas entre sí, pero lo que pocos saben es que la IA no es una tecnología, sino un conjunto de áreas distintas, como caminos que se abren en diferentes direcciones.

Una de las áreas más importantes es el *Aprendizaje Automático* (en inglés, *Machine Learning* o *ML*), que permite a las máquinas (computadoras) aprender a partir de datos. Hay un conjunto de pruebas que se hacen sobre todos los datos hasta que logran encontrar las relaciones o patrones entre estos. Por ejemplo, si un sistema analiza cientos de partidos de fútbol, puede aprender a predecir cuándo un jugador tiene mayor probabilidad de anotar un gol o incluso sugerir estrategias para lograrlo según el comportamiento previo de los equipos. Dentro del Aprendizaje Automático, existe una técnica avanzada llamada *Aprendizaje Profundo* (en inglés, *Deep Learning* o *DL*), que emplea redes de "neuronas artificiales" inspiradas (de forma muy simplificada) en el cerebro humano. Gracias a esto, las máquinas pueden resolver tareas más complejas. Algunas aplicaciones en las que hay mucho interés son la detección de células anormales en estudios médicos o conducir un automóvil sin conductor. Dos aplicaciones muy conocidas del aprendizaje profundo son la *Visión por Computadora* (en inglés, *Computer Vision*) y el *Procesamiento de Lenguaje Natural* (en inglés, *Natural Language Processing* o NLP). La visión por computadora permite a las máquinas "ver" e interpretar imágenes, como cuando un celular se desbloquea al mostrarle tu rostro o cuando un sistema identifica una fractura usando una radiografía. Por otro lado, el procesamiento de lenguaje natural permite que los programas comprendan lo que decimos o escribimos [1-3]. Esta tecnología está detrás de los traductores automáticos, los asistentes virtuales como *Siri* o *Alexa* y también en herramientas como *ChatGPT*. Cada una de estas áreas cubre una brecha distinta entre lo que los humanos hacemos con naturalidad y lo que las máquinas han aprendido a hacer. Algunas les permiten ver el mundo a través de imágenes, otras escuchar y reconocer sonidos y otras tomar decisiones basadas en datos. Hay un área que destaca porque se enfoca en algo profundamente humano: el lenguaje, nuestra principal forma de pensar, aprender y comunicarnos. Lograr que una máquina entienda nuestras palabras y responda de manera coherente era un sueño de la ciencia ficción hasta hace apenas unas décadas. Hoy, gracias a los modelos grandes de lenguaje, ese sueño se está volviendo realidad, cambiando la manera en que interactuamos con la tecnología: hablamos con asistentes virtuales, obtenemos respuestas rápidas a preguntas complejas e incluso conversamos con sistemas capaces de explicar conceptos científicos en segundos.

¿Qué es un modelo grande de lenguaje?

Son programas de IA diseñados para trabajar exclusivamente con texto. Su función principal parece sencilla: predecir la siguiente palabra en una oración. Sin embargo, para lograrlo, necesitan aprender cantidades enormes de información y emplear técnicas muy avanzadas. Imagina el autocompletado de tu celular: cuando empiezas a escribir un mensaje, te sugiere cuál podría ser la siguiente palabra. Ahora, imagina un sistema que ha "leído" millones de libros, artículos científicos, noticias, conversaciones y páginas web, y que puede usar todo ese conocimiento para continuar la frase de manera coherente, responder preguntas o redactar un texto completo sobre casi cualquier tema. Eso es un LLM: un "superautocompletador" capaz de generar texto con fluidez y naturalidad.

Estos modelos no entienden lo que dicen como lo hacen los humanos. No tienen emociones, opiniones ni conciencia propia, solo *procesan texto basándose en probabilidades*. Lo que realmente hacen es analizar patrones en miles de millones de palabras, aprendiendo cómo suelen relacionarse entre sí y usando ese conocimiento para construir frases coherentes. Por ejemplo, si les damos la frase: "*El fútbol es un deporte muy...*", el modelo recuerda haber visto millones de textos en los que después de esas palabras aparecen términos como "popular" o "emocionante". Calcula cuál de estas palabras tiene mayor probabilidad de aparecer y la elige para continuar la frase. Si le damos un contexto distinto como "*El fútbol es un deporte peligroso cuando...*", ajustará su predicción para ofrecer una continuación más coherente, como "*los jugadores no usan protección adecuada*". La fuerza de estos modelos no está en que comprendan las ideas, sino en que pueden manejar contextos muy amplios al mismo tiempo, encontrando relaciones entre palabras y frases que para nosotros son naturales, pero que para una máquina son solo patrones matemáticos. Gracias a este enfoque, pueden responder preguntas, redactar historias, contar chistes o incluso mantener conversaciones, pero siempre a partir de cálculos estadísticos, no de comprensión real.

Por eso, aunque sus respuestas puedan sonar inteligentes o creativas, un LLM no tiene experiencias propias ni comprende el mundo como nosotros. Todo lo que dice proviene de la información con la que fue entrenado y de las reglas matemáticas que aprendió para ordenar palabras.

¿Cómo aprenden a hablar estas inteligencias artificiales?

Para que un modelo como ChatGPT pueda mantener una conversación, no basta con programarlo palabra por palabra. Necesita aprender de grandes cantidades de información, casi como lo hacemos las personas cuando empezamos a hablar: escuchamos, leemos, practicamos y poco a poco mejoramos nuestra manera de expresarnos. En el caso de las IA, este aprendizaje ocurre en dos etapas principales: el preentrenamiento y el ajuste fino.

El pre-entrenamiento: leer todas las bibliotecas del mundo

En esta primera etapa, los modelos son expuestos a cantidades inmensas de texto; se les alimenta con miles de millones de palabras para que aprendan la estructura del lenguaje, el significado de las palabras según el contexto y cómo suelen relacionarse entre sí. Podemos imaginarlo así: es como un estudiante que pasa años leyendo todo lo que se encuentra en una biblioteca gigantesca, desde cuentos infantiles hasta investigaciones científicas. Con esta lectura masiva, aprende gramática, estilo, expresiones y cómo suelen construirse las ideas en diferentes situaciones. No obstante, en esta fase, *el modelo solo aprende patrones generales*, sin distinguir qué información es confiable, correcta o apropiada para cada situación [2,4-5].

El ajuste fino: clases particulares para expresarse mejor

Una vez que el modelo ha leído todo lo posible, necesita mejorar su forma de responder. Aquí entra la segunda etapa, conocida como *ajuste fino* (en inglés, *fine-tuning*). Durante este proceso, los creadores del modelo le dan ejemplos más precisos y usan retroalimentación humana para enseñarle qué tipo de respuestas son más útiles, claras y seguras. Podría asemejarse a un estudiante que, tras pasar años leyendo por su cuenta, decide asistir a clases particulares con un profesor experimentado. Durante estas sesiones, el profesor revisa cómo responde el estudiante, le señala dónde se equivoca, le enseña a expresarse con mayor claridad y le muestra cómo dar respuestas útiles y respetuosas. Con el tiempo, el estudiante aprende a ordenar mejor sus ideas, a elegir las palabras adecuadas según la situación y a evitar comentarios equivocados o confusos. Del mismo modo, en la etapa de ajuste fino, los modelos reciben ejemplos de cómo responder a distintas preguntas y, gracias a la corrección humana, mejoran su capacidad para mantener conversaciones coherentes y útiles para las personas [2,4-5].

Con estas dos etapas, los modelos grandes de lenguaje logran construir oraciones, mantener conversaciones y adaptarse a distintos temas, desde deportes hasta biomedicina. También desarrollan la capacidad de organizar ideas y responder a las personas. Sin embargo, su desempeño está limitado por los textos con los que fueron entrenados y la calidad de la guía humana recibida durante el ajuste fino. Si la información original es incompleta, contradictoria o contiene errores, el modelo puede repetirlos en sus respuestas. Es como un estudiante que aprende de apuntes equivocados: aunque hable con seguridad, sus explicaciones no siempre serán correctas.

¿Cómo generan texto cuando conversamos con ellas?

Una vez que los modelos han sido entrenados y han aprendido patrones de lenguaje, llega la parte impresionante: la generación de texto en tiempo real. Cada vez que escribimos una pregunta o iniciamos una conversación con ChatGPT, el modelo no busca una respuesta almacenada ni copia un párrafo exacto de algún libro. Lo que hace es predecir palabra por palabra cuál es la continuación más probable, basándose en todo lo que aprendió. Este proceso se llama funcionamiento autorregresivo, y se parece a cuando una persona improvisa una respuesta. Por ejemplo, cuando te hacen una pregunta inesperada, no tienes toda la respuesta lista en tu mente. Más bien, vas construyendo tu respuesta palabra por palabra según recuerdas información relevante. De manera similar, estos modelos improvisan sus respuestas a partir de los patrones aprendidos, palabra por palabra, en tiempo real. La gran diferencia es que en un modelo como ChatGPT no solo tiene en cuenta las últimas dos o tres palabras, sino todo el contexto de la conversación, pudiendo analizar frases largas, ideas previas y la estructura completa de lo que estás preguntando. Esto le permite generar textos adaptados a tu solicitud. Supongamos que escribes: "El corazón humano es un órgano que…", el modelo calcula qué palabra podría aparecer a continuación basándose en todas las veces que ha visto oraciones similares durante su entrenamiento. Podría continuar con "bombea", "late" o "impulsa sangre", eligiendo aquella opción que tiene mayor probabilidad de encajar con el resto de la frase. Después de esa palabra, vuelve a calcular cuál es la siguiente palabra más probable y así sucesivamente, hasta completar la respuesta. Cabe resaltar que todo esto ocurre en milésimas de segundo. Este mecanismo es tan rápido y preciso que parece que el modelo "piensa" en lo que dice, pero en realidad solo está siguiendo un cálculo estadístico muy sofisticado. No tiene un plan previo ni una idea clara de toda la respuesta desde el inicio: simplemente va construyendo el texto paso a paso, evaluando cada posible palabra y escogiendo la que mejor se ajusta al contexto y al estilo de la conversación. Para hacer posible este proceso, el modelo utiliza cientos de miles de millones de parámetros (por ejemplo, ≈175 mil millones en GPT-3), que funcionan como pequeñas conexiones internas en una red enorme, cada una influyendo en cómo las palabras se relacionan entre sí. Durante el entrenamiento, estas conexiones se ajustaron millones de veces hasta encontrar la combinación que permite generar textos comprensibles y naturales. Podríamos decir que, cuando conversamos con él, el modelo activa esas conexiones y las coordina para formar respuestas sin necesidad de memorizar frases exactas. Por ejemplo, el modelo podría funcionar como un médico virtual, que ha leído millones de libros de medicina, artículos científicos y reportes clínicos. Cuando le preguntas algo, no busca una respuesta almacenada, sino que va construyendo la frase palabra por palabra (Figura 1), asignando distintas probabilidades a las posibles continuaciones de la oración. Si le preguntas: "¿Cuál es la función del corazón humano?", el sistema puede responder "bombea sangre para llevar oxígeno a todo el cuerpo", seleccionando cada palabra según su probabilidad condicional dada el contexto, no porque 'quiera' lograr una explicación clara. Así funcionan los modelos grandes de lenguaje: no repiten textos, lo que realizan son cálculos en tiempo real para dar la respuesta más probable [1,4-5].

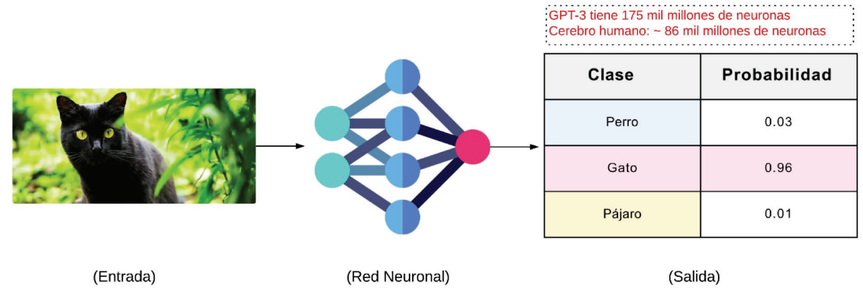

FIGURA 2. CLASIFICACIÓN de imágenes mediante una red neuronal de visión por computadora. La fotografía de un gato pasa por una red neuronal convolucional que genera probabilidades para cada clase ("Perro", "Gato", "Pájaro"). Se incluye, a modo de referencia de escala, el número aproximado de parámetros que emplea GPT-3 (≈175 mil millones) y el número de neuronas de un cerebro humano (≈86 mil millones).

Usos cotidianos

Aunque parezca una tecnología muy reciente, la realidad es que estos sistemas están presentes en distintas actividades y campos, ayudando a personas de todas las edades. En las compras en línea, los modelos pueden ayudarte a comprar productos, recomendarte opciones según tus gustos o redactar mensajes para vendedores. En la cocina, proponen recetas usando los ingredientes que tienes a mano, ajustan platillos para hacerlos más saludables o convierten medidas y proporciones al instante mientras preparas tus alimentos. Durante un viaje, sirven para planear itinerarios, buscar actividades en un destino turístico, traducir frases a otro idioma o incluso escribir mensajes para comunicarte mejor en el extranjero. En la organización diaria, son útiles para elaborar listas de compras, programar recordatorios, redactar felicitaciones personalizadas o incluso escribir invitaciones y mensajes para eventos familiares. En el ámbito de la creatividad digital, son útiles no solo para escribir publicaciones en redes sociales, sino también para generar descripciones de fotografías. En un ejemplo de cómo funciona esto (Figura 2), la fotografía de interés se ingresa a una red neuronal, que analiza la imagen y asigna la mayor probabilidad a la clase 'Gato' o al objeto que desees analizar. También son aliados en la educación, explicando temas difíciles en palabras sencillas, ofreciendo ejemplos adicionales para entender una lección o generando cuestionarios y resúmenes para estudiar. Y, por supuesto, contribuyen al entretenimiento, sirviendo para crear historias interactivas, diálogos para videojuegos o guiones para podcasts y videos. Incluso pueden ayudarte a componer canciones o inventar cuentos para niños, adaptando el estilo y el tono a lo que pidas [3]. Estos ejemplos muestran cómo esta tecnología, que parecía exclusiva de laboratorios y expertos en informática, ya es parte de nuestras actividades diarias, transformando la manera en que estudiamos, trabajamos y nos comunicamos.

Limitaciones y precauciones

A pesar de sus sorprendentes capacidades, los *modelos grandes de lenguaje* no son perfectos ni infalibles. Aunque pueden dar la impresión de "saberlo todo", es importante recordar que no tienen emociones, conciencia ni comprensión real de lo que dicen. Solo repiten patrones aprendidos a partir de los textos con los que fueron entrenados. Una de sus principales limitaciones son las llamadas "alucinaciones", que ocurren cuando el modelo inventa información que suena coherente pero no es verdadera. Por ejemplo, puede citar un dato inexistente, inventar una referencia científica o mezclar hechos reales con otros completamente falsos. Otra limitación son los *sesgos en los datos de entrenamiento*. Si los textos con los que aprendió contenían prejuicios, estereotipos o información poco equilibrada, es posible que el modelo refleje esos mismos problemas en sus respuestas. Además, no pueden verificar información en tiempo real, ya que no están conectados permanentemente a bases de datos actualizadas. Tampoco pueden sustituir el juicio humano, especialmente en áreas sensibles como la medicina, el derecho o la toma de decisiones críticas. Por estas razones, se recomienda usar estas herramientas como apoyo, para obtener ideas, aclarar conceptos o redactar textos más rápido, pero nunca como una fuente única o infalible de información. La supervisión humana, la verificación de datos y el pensamiento crítico siguen siendo esenciales cuando utilizamos la IA [5].

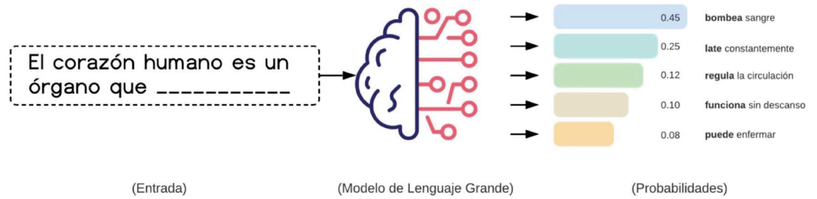

FIGURA 1. FUNCIONAMIENTO básico de un Modelo Grande de Lenguaje (LLM). A partir de una frase incompleta ("El corazón humano es un órgano que …"), el modelo calcula la probabilidad de varias continuaciones. La barra azul representa la opción más probable ("bombea sangre"), seguida de otras alternativas ("late", "regula", etc.). La salida no es una idea "comprendida", sino una distribución de probabilidades sobre palabras que podrían completar el enunciado.

Conclusión

La llegada de los *modelos grandes de lenguaje* marca un punto de inflexión en nuestra relación con la tecnología. Por primera vez, podemos interactuar con sistemas que generan texto de forma tan natural que parecen comprendernos. Esto abre la puerta a nuevas formas de aprender, comunicarnos y crear contenido, acercando el conocimiento y la información a más personas en cuestión de segundos. Pero más allá de la fascinación inicial, este avance nos invita a reflexionar sobre cómo queremos convivir con estas herramientas. Su uso responsable y ético será clave para evitar que la información errónea se propague, que los sesgos se amplifiquen o que dependamos demasiado de las respuestas automáticas. La IA no reemplaza nuestra capacidad de pensar, analizar y decidir; al contrario, nos reta a fortalecerla para aprovechar al máximo una tecnología que, bien utilizada, puede convertirse en una poderosa aliada para el futuro.

Referencias

[1] Tan, X., G. Cheng, and M. H. Ling. (2025). "Artificial intelligence in teaching and teacher professional development: A systematic review." *Computers and Education: Artificial Intelligence*, 8: 100355. doi: 10.1016/j.caeai.2024.100355

[2] Stöffelbauer, A. "How Large Language Models Work". Medium. Consultado el 31 de julio de 2025. [En línea]. Disponible en: https://medium.com/data-science-at-microsoft/how-large-language-models-work-91c362f5b78f.

[3] Kühl, N., M. Schemmer, M. Goutier, and G. Satzger. (2022). "Artificial intelligence and machine learning." *Electronic Markets*, 32(4):2235–2244. doi: 10.1007/s12525-022-00598-0

[4] Kasneci, E., K. Sessler, S. Küchemann y col. (2023). "ChatGPT for good? On opportunities and challenges of large language models for education." *Learning and Individual Differences*, 103: 102274. doi: 10.1016/j.lindif.2023.102274

[5] Zhao, W. X., K. Zhou, J. Li, T. Tang, y col. (2025). "A survey of large language models." https://doi.org/10.48550/arXiv.2303.18223

Esta columna se prepara y edita semana con semana, en conjunto con investigadores morelenses convencidos del valor del conocimiento científico para el desarrollo social y económico de Morelos.

ESTA PUBLICACIÓN FUE REVISADA POR EL COMITÉ EDITORIAL DE LA ACADEMIA DE CIENCIAS DE MORELOS

Academia de Ciencias de Morelos, A.C.

Para actividades recientes de la academia y artículos anteriores puede consultar: www.acmor.org

¿Comentarios y sugerencias?, ¿Preguntas sobre temas científicos? CONTÁCTANOS: coord.comite.editorial.acmor@gmail.com