

¿Y SI LAS TECNOLOGÍAS DE IA PUDIERAN RAZONAR COMO LOS HUMANOS?

POR DR. JOHN ATKINSON
ESPECIALISTA, CONSULTOR Y PROFESOR DE IA

La Inteligencia Artificial (IA) ya no es cosa del futuro, está aquí y toma decisiones que impactan directamente nuestra vida: desde si nos aprueban un crédito hasta cómo nos atiende un servicio al cliente. Por eso, es clave que estas decisiones no sean solo una repetición de lo que la IA ha visto en internet, sino que tengan una base lógica y que podamos entender.

El “cerebro” detrás de muchas de las tecnologías de IA más conocidas, como ChatGPT, se llama un Gran Modelo de Lenguaje (LLM). En pocas palabras, un LLM aprende a predecir la siguiente palabra en un texto basándose en enormes cantidades de información que ha “leído” desde variadas fuentes.

El problema es que estos LLM no “razonan” de verdad; solo parecen hacerlo, por lo que a menudo generan respuestas que suenan convincentes, pero que en realidad no tienen lógica o están equivocadas. Peor aún, estas respuestas pueden cambiar de un momento a otro, generando inconsistencias para la misma pregunta.

Las consecuencias de una IA sin lógica. Imagina que una herramienta de IA recomienda un producto, aprueba un crédito o rechaza una solicitud sin una justificación clara. Esto sucede porque no está planificando con reglas o consecuencias, solo está repitiendo patrones aprendidos. Esto puede llevar a errores costosos, decisiones injustas o a que los usuarios pierdan la confianza en las empresas.

En situaciones donde se necesita cumplir normas, justificar decisiones o manejar escenarios nuevos y complejos, esta falta de razonamiento consistente se puede convertir en un verdadero dolor de cabeza para las empresas.

Recientemente, una investigación de científicos de Apple se hizo viral al probar lo que muchos expertos en IA ya sabían: los LLM no pueden razonar. De hecho, esta discusión no es nueva; se remonta a los inicios de la IA en los años 50, cuando surgió el área del razonamiento automático. Desde entonces, se han desarrollado

herramientas de IA capaces de razonar, como programas que demuestran teoremas o resuelven problemas de planificación o diseñan y/o validan circuitos.

¿QUÉ ES EL RAZONAMIENTO AUTOMÁTICO Y POR QUÉ ES TAN IMPORTANTE?

Un sistema de razonamiento automático utiliza lógica matemática para realizar deducciones y probar que sus conclusiones son válidas y se derivan de la información inicial (premisas). A diferencia de las tecnologías utilizadas en los LLM, un sistema de razonamiento automático garantiza que sus conclusiones son lógicamente correctas. Para lograr esto, el razonamiento automático requiere de tres elementos:

1. Representación del conocimiento: El sistema traduce la información a un lenguaje lógico y formal usualmente basado en modelos matemáticos.

2. Motor de razonamiento: Aplica reglas lógicas para sacar conclusiones a partir de la información o premisas iniciales.

3. Generación de pruebas: El sistema crea “pruebas” o cadenas de inferencias lógicas para demostrar la validez de sus conclusiones. Esto es crucial, pues asegura que la “demostración” o el plan de la IA es válido y consistente.

EL FUTURO DE LA IA: COMBINANDO LO MEJOR DE AMBOS MUNDOS

Una máquina con razonamiento automático podría planificar antes de tomar una decisión, analizando reglas, contextos y consecuencias, tal como lo haría una persona responsable. Esto nos daría más seguridad, transparencia y justicia en un mundo cada vez más dominado por los algoritmos.

Para lograr esto, los sistemas de IA actuales necesitan ir más allá de la IA generativa y combinarla con tecnologías de razonamiento simbólico, que han existido por años. Luego, el futuro de la IA se

orientará a combinar los modelos de predicción (como los LLM) con los de razonamiento, dando lugar a los sistemas neuro-simbólicos, que ya se ha comenzado a trabajar hace algunos años.

Mientras esto no suceda, el supuesto razonamiento en los sistemas de IA ac-

tuales seguirá siendo una ilusión que, en la práctica, puede generar más problemas que beneficios debido a sus inconsistencias, errores y la dificultad para explicar sus resultados.