

"El verdadero hito llegará cuando la IA no solo evite errar en el conocimiento establecido, sino cuando sea capaz de ofrecer respuestas a incógnitas que aún no hemos logrado resolver", destacaron los investigadores.



LA SOMETIERON A PREGUNTAS COMPLICADAS

Científicos comprobaron que la IA aún no logra el razonamiento humano complejo

La investigación buscaba determinar si los modelos actuales poseen una verdadera capacidad de resolución creativa o se limitan a una sofisticada gestión de datos.

AGENCIAS
diario@ladiscusion.cl
FOTOS: AGENCIAS

Un equipo internacional de científicos comprobó que la inteligencia artificial (IA) sigue cometiendo errores a la hora de responder preguntas que requieren razonamiento conceptual y no solo búsquedas en internet, lo que revela que aún no iguala el pensamiento humano en tareas académicas complejas.

Según detalló este jueves en un comunicado la Universidad de Granada (UGR), de España, la investigación nació con el objetivo de determinar si los actuales grandes modelos de lenguaje (LLM) poseen una amplia y verdadera capacidad de resolución creativa o si, por el contrario, su funcionamiento se limita a una sofisticada gestión de los datos que esos modelos rescatan de la nube.

Para ello, la comunidad científica diseñó una batería de preguntas de alta complejidad técnica y conceptual como parte de su investigación, llamada 'Humanity's Last Exam' ('Último Examen de la Humanidad', en español), publicada en la revista Nature.

Así, un equipo formado por 1.100 científicos de todos los ámbitos de la ciencia y las humanidades -entre los que

se encuentra la española María Cruz Boscá, de la UGR- sometió a la IA a un "gran examen" para valorar su réplica a preguntas de respuesta inequívoca y verificable pero que requieren más que una búsqueda en internet.

Cada pregunta tenía una solución conocida que es inequívoca y verificable, pero que no puede responderse rápida y fácilmente mediante una búsqueda en internet.

Como detalló el artículo 'Una base de referencia de preguntas académicas de nivel experto para evaluar las capacidades de la IA', el resultado es que, a día de hoy, incluso los modelos más avanzados de IA tropiezan con conceptos científicos profundos y heredan errores de los manuales clásicos, lo que pone de relieve una marcada diferencia entre las capacidades actuales de los LLM y las de los expertos humanos en las diferentes preguntas académicas planteadas, referentes a diversos campos científicos.

Razonamiento

Según explicó la profesora Boscá, que puso a prueba los límites del razonamiento artificial en el ámbito de la física cuántica, su investigación comprobó que las inteligencias artificiales no lograron elegir las respuestas correctas cuando tenían que realizar una profunda comprensión conceptual.