



CIENTÍFICOS ESTADOUNIDENSES EVALUARON 21 SISTEMAS DE IA.

IA mejora diagnósticos médicos, pero sigue sin razonamiento crítico

Pese a que los chatbots de inteligencia artificial (IA) utilizados en atención médica suelen acertar el diagnóstico de los pacientes cuando disponen de información clínica completa, siguen mostrando deficiencias cuando tienen que diagnosticar con poca información, o en los casos de una variable de la enfermedad.

Según un estudio publi-

cado en JAMA Network Open, y realizado por la red de hospitales de Boston Mass General Brigham, Estados Unidos, la IA aún no está preparada para tomar decisiones médicas sin la supervisión constante de un profesional humano.

El equipo llegó a esta conclusión tras evaluar 21 de los modelos de lenguaje más avanzados, entre ellos

GPT-5, Grok 4, Claude, DeepSeek y Gemini. Se pidió a estas IA que actuaran como médicos en una serie de escenarios clínicos y comprobaron que a menudo fallan ante una lista comprobable de potenciales enfermedades o "diferenciales".

"Pese a las continuas mejoras, los modelos de lenguaje grandes estándar no están listos para un despliegue de

grado clínico sin supervisión", concluyó uno de los autores, Marc Succi.

"Los diagnósticos diferenciales son fundamentales para el razonamiento clínico y subyacen al 'arte de la medicina', que la IA no puede replicar actualmente", destacó. Por ahora, la IA solo sirve para "aumentar, no reemplazar, el razonamiento del médico".