

El primer ensayo de **terapia** con **IA generativa**



arroja
resultados
prometedores
para
**tratar la
depresión**

**MIT
Technology
Review**

Publicado por Opinno

JAMES O'DONNELL

01/04/2025

El modelo respaldado por evidencia arrojó resultados impresionantes, pero no valida la ola de bots terapéuticos con IA que inundan el mercado.

El primer ensayo clínico de un *bot* de terapia que utiliza IA generativa demuestra que fue tan efectivo como la terapia humana para los participantes con depresión, ansiedad o riesgo de desarrollar trastornos alimentarios. Aun así, no da luz verde a las docenas de empresas que promocionan estas tecnologías que operan en un área gris regulatoria.

Un equipo liderado por investigadores psiquiátricos y psicólogos de la Escuela de Medicina Geisel en Dartmouth College (Nuevo Hampshire, EE UU) construyó la herramienta, llamada Therabot. Los resultados se publicaron el 27 de marzo en NEJM AI, una revista del New England Journal of Medicine. Muchas empresas tecnológicas han creado herramientas de IA para la terapia, con la promesa de que las personas puedan hablar con un *bot* con más frecuencia y a menor costo que con un terapeuta capacitado. Aseguran que este enfoque es seguro y efectivo.

Muchos psicólogos y psiquiatras han compartido esta visión, señalando que menos de la mitad de las personas con un trastorno mental reciben terapia, y quienes lo hacen pueden recibir solo 45 minutos por semana. Los investigadores han intentado construir tecnología para que más personas puedan acceder a la terapia, pero se han visto frenados por dos cosas.

Primero, un *bot* de terapia que diga algo incorrecto podría causar un daño real. Por eso, muchos investigadores han desarrollado estos programas con respuestas predefinidas y aprobadas, como ocurrió con Eliza, un software de los años 60 que simulaba ser un psicoterapeuta. Sin embargo, este enfoque hace que la conversación resulte menos atractiva y los usuarios pierdan interés.

En segundo lugar, los elementos clave de una buena relación terapéutica, como la colaboración y la construcción de objetivos comunes, son difíciles de reproducir en un programa informático.

En 2019, mientras los primeros modelos de lenguaje a gran escala, como el GPT de OpenAI, comenzaban a desarrollarse, un grupo de investigadores de Dartmouth vio en la IA generativa una posible solución a estos desafíos. Su objetivo era crear un modelo capaz de ofrecer respuestas basadas en evidencia. Primero, entrenaron la IA con conversaciones generales sobre salud mental extraídas de foros en internet. Luego, recurrieron a miles de horas de transcripciones de sesiones reales con psicoterapeutas.

«Obteníamos muchos ‘ajá’, ‘continúa’ y luego ‘Tus problemas provienen de tu relación con tu madre’», explicó Nick Jacobson, profesor asociado de ciencia de datos biomédicos y psiquiatría en Dartmouth, y autor principal del estudio, en una entrevista. «Eran más bien clichés sobre la psicoterapia, en lugar de lo que realmente buscábamos.»

Insatisfechos, se pusieron a trabajar en la creación de sus propios conjuntos de datos basados en prácticas respaldadas por evidencia, que finalmente fueron los que integraron en el modelo.

En contraste, muchos *bots* de terapia con IA disponibles en el mercado son solo ligeras variaciones de modelos base como Llama de Meta,



entrenados principalmente con conversaciones extraídas de internet. Esto supone un problema, especialmente en temas como los trastornos alimentarios. «Si dijeras que quieres perder peso», dice Jacobson, «estos bots te apoyarían sin dudar, incluso si ya tienes un peso bajo para empezar». Un terapeuta humano no haría eso.

Para probar el *bot*, los investigadores realizaron un ensayo clínico de ocho semanas con 210 participantes que tenían síntomas de depresión o trastorno de ansiedad generalizada o estaban en alto riesgo de trastornos alimentarios. Aproximadamente la mitad tuvo acceso a Therabot, y un grupo de control no. Los participantes respondieron a las indicaciones de la IA e iniciaron conversaciones, con un promedio de 10 mensajes por día.

Los participantes con depresión experimentaron una reducción del 51% en los síntomas, el mejor resultado del estudio. Aquellos con ansiedad experimentaron una reducción del 31%, y aquellos en riesgo de trastornos alimentarios vieron una reducción del 19% en las preocupaciones sobre la imagen corporal y el peso. Estas mediciones se basan en autoevaluaciones a

muchos psicólogos y psiquiatras han compartido esta visión, señalando que menos de la mitad de las personas con un trastorno mental reciben terapia, y quienes lo hacen pueden recibir solo 45 minutos por semana.

través de encuestas, un método que no es perfecto, pero sigue siendo una de las mejores herramientas que tienen los investigadores.

Estos resultados, dice Jacobson, son aproximadamente lo que se encuentra en ensayos controlados aleatorios de psicoterapia con 16 horas de tratamiento proporcionado por humanos, pero el ensayo de Therabot lo logró en aproximadamente la mitad del tiempo. «He estado trabajando en terapias digitales durante mucho tiempo, y nunca he visto niveles de compromiso que sean prolongados y sostenidos a este nivel», dice.

Jean-Christophe Bélisle-Pipon, profesor asistente de ética de la salud en la Universidad Simon Fraser (Columbia Británica, Canadá), quien ha escrito sobre *bots* de terapia con IA, pero no participó en la investigación, dice que los resultados son impresionantes, pero señala que, al igual que cualquier otro ensayo clínico, este no necesariamente representa cómo actuaría el tratamiento en el mundo real.

«Aún estamos lejos de dar luz verde para su implementación clínica a gran escala», escribió en un correo electrónico.

un equipo liderado por investigadores psiquiátricos y psicólogos de la Escuela de Medicina Geisel en Dartmouth College (Nuevo Hampshire, EE UU) construyó la herramienta, llamada Therabot.

Uno de los principales desafíos es la supervisión que requeriría una implementación a gran escala. Al inicio del ensayo, Jacobson supervisó personalmente todos los mensajes de los participantes (que dieron su consentimiento) para detectar respuestas problemáticas del *bot*. Si los *bots* de terapia necesitaran este nivel de control, su alcance sería mucho más limitado.

Cuando le pregunté a Jacobson si los resultados validaban la creciente industria de la terapia con IA, su respuesta fue clara: «Todo lo contrario». Advirtió que la mayoría de estas plataformas no entrenan sus modelos en prácticas basadas en evidencia, como la terapia cognitivo-conductual, ni cuentan con equipos de investigadores capacitados que supervisen las interacciones. «Me preocupa mucho la velocidad con la que avanzamos sin evaluar realmente esto», agregó.

Cuando estos sitios se promocionan como plataformas legítimas de terapia en un contexto clínico, Jacobson señala que deberían estar bajo la regulación de la Administración de Alimentos y Medicamentos (FDA). Hasta ahora, la FDA no ha tomado medidas contra la mayoría de ellos, pero si lo hiciera, «sospecho que prácticamente ninguno—probablemente ninguno—de los que operan en este espacio obtendría la aprobación para respaldar sus afirmaciones sobre beneficios terapéuticos».

Por su parte, Bélisle-Pipon advierte que, si estas terapias digitales no son aprobadas e integradas en los sistemas de salud y seguros, su alcance será muy limitado. Como consecuencia, las personas que podrían beneficiarse de ellas podrían recurrir a herramientas de IA no diseñadas para ese propósito. De hecho, una nueva investigación de OpenAI sugiere que las interacciones con sus modelos de IA pueden tener un impacto real en el bienestar emocional.

«Es muy probable que muchas personas sigan utilizando *chatbots* más asequibles y no terapéuticos—como ChatGPT o Character.AI—para necesidades cotidianas, desde generar ideas de recetas hasta gestionar su salud mental», escribió.

Corrección: El estudio fue publicado en *NEJM AI*, una revista del *New England Journal of Medicine*, y no en el *New England Journal of Medicine*, como se indicó originalmente. También

se actualizaron las atribuciones de citas, que corresponden a Nick Jacobson y no a su coautor Michael Heinz. </>



 Editor principal de energía de *MIT Technology Review*. Especialista en energías renovables y en uso de la tecnología para combatir el cambio climático.

El artículo original «El primer ensayo de terapia con IA generativa arroja resultados prometedores para tratar la depresión» pertenece a la edición digital de *MIT Technology Review*.

Los contenidos bajo el sello *MIT Technology Review* están protegidos enteramente por copyright. Ningún material puede ser reimpresso parcial o totalmente sin autorización.

Si quisiera syndicar el contenido de la revista *MIT Technology Review*, por favor contáctenos.

E-mail: redaccion@technologyreview.com

Tel: +34 911 284 864