LITORALPRESS

Fecha: 29-01-2025
Medio: Las Últimas Noticias
Supl. : Las Últimas Noticias
Tipo: Noticia general
Título: ¿Puede una IA sentir dolor? Investigadores logran sorprendentes resultados

Pág. : 13
Cm2: 727,7

Tiraje: 91.144
Lectoría: 224.906
Favorabilidad: No Definida

Experimento de Google DeepMind indagó en la sensibilidad de la tecnología

¿Puede una IA sentir dolor? Investigadores logran sorprendentes resultados

Rodrigo Castillo

Estudio fue realizado con nueve modelos extensos de lenguaje (LLM) sometidos a juegos de preguntas en los que debían optar entre sufrir y obtener recompensas.

Justo en momentos en que la carrera por el dominio en el mercado de la Inteligencia Artificial llega a un punto caliente, con la irrupción de la IA china DeepSeek, que ya está poniendo en apuros a colosos como Nvidia, Meta y Microsoft, una nueva investigación aporta sorprendentes datos para vislumbrar el real potencial de este tipo de sistemas. El estudio en cuestión intentó responder una pregunta inquietante: ¿puede la IA experimentar placer y dolor?

El experimento, realizado por especialistas de Google DeepMind y la London School of Economics and Political Science, se basó en una serie de juegos en los que nueve modelos extensos de lenguaje (LLM) fueron invitados a optar entre sufrir "dolor" y obtener recompensas.

Algunos de los resultados fueron sorprendentes: modelos como Claude 3.5 Sonnet, Command R+, GPT-4o y GPT-4o mostraron inclinación a renunciar al objetivo de maximizar puntos, cada vez que el "dolor" o el "placer" alcanzaban una cierta intensidad. El Gemini 1.5 Pro de Google, en tanto, prefirió evitar el dolor antes que conseguir puntos.

El hallazgo, que aún no ha pasado por la revisión de pares y fue difundido originalmente por la revista "Scientific American", se basó, por supuesto, en los conceptos de "placer" y "dolor" que los diversos modelos de IA han almacenado en sus bancos de datos sobre los seres humanos. El hecho de que esas máquinas carezcan de cuerpos orgánicos y sistemas nerviosos capaces de sufrir o gozar, sin embargo, no resta validez a la investigación. Al menos así lo piensa el español Jacinto Obispo, especialista de Apiux, consultora internacional en temas de tecnología.

"Yo creo que la IA va a ser capaz de imitar perfectamente y entender en qué circunstancias se producen el dolor y el placer, eso para mí está fuera de duda. Y podríamos llevarlo a un plano más complejo, metiendo la IA en un dispositivo físico de forma humanoide que podría llegar a sentir dolor físico real. De hecho, se me ocurren casos de uso en los que estos humanoides con sensibilidad podrían ser muy útiles, de gran beneficio para la humanidad", comenta el experto.

"Ellos podrían, por ejemplo, ayudar a predecir qué grado de dolor llegaría a sentir una persona si se sumergiera a tres mil metros de profundidad, o qué daños sufriría alguien en un accidente automovilístico, o si un aparato puede ser contraproducente para la salud de una persona", agrega Obispo, quien también imagina otros usos menos violentos para esa hipotética tecnología: "Estos humanoides podrían atender y acompañar a adultos mayores, haciendo que éstos se sientan bien, conversando y siendo empáticos con ellos".

Más escéptico se muestra Tomás Ossandón, neurocientífico de la Facultad de Medicina de la Universidad Católica, quien cree que la concepción del dolor que podría tener una IA está, aún, determinada por el aprendizaje que los mismos humanos les han proporcionado.

"Una IA podría llegar a emular la forma en que una persona se comportaría al alcanzar el límite de dolor que puede tolerar. Pero eso sería, más bien, una abstracción del dolor que no tendría nada que ver con la experiencia del dolor o el placer. De todas formas, creo que sería interesante que se pudiera generar una IA con esas capacidades, porque podrían llegar a tener una cierta agencia, ser capaces de salirse de sí mismas para buscar respuestas en otros lados", reflexiona.

"Yo creo que la única forma en que podríamos generar algo similar al dolor, en las máquinas, es hacer que éstas se den cuenta de que en función de ciertas respuestas van a tener menos energía, menos voltaje para funcionar. Para generar una abstracción en función de algún tipo de respuesta, tienes que tener alguna consecuencia dentro de tu funcionamiento, entonces habría que pensar de qué forma la integridad de una máquina puede verse afectada por sus decisiones", plantea.

Patricio Campos, especialista en ciberseguiridad e IA, y CEO de Resility, considera que no es probable que en el corto plazo se desarrolle la tecnología necesaria para que estos sistemas puedan llegar a sentir dolor. El tema, en todo caso, le genera inquietudes en el plano ético: si la IA llegara a convertirse en una entidad consciente, el hecho tendría que motivar un debate social complejo respecto de la condición ciudadana que tendría esa hipotética criatura.

"La aparición de una IA consciente no sólo sería un salto tecnológico, sino que nos obligaría a comprender bien de qué se trata la consciencia humana, porque ni siquiera los humanos tenemos muy claro cuál es el origen de nuestra consciencia. En el caso de que se lograra algo así, comienza una serie de implicaciones morales. ¿Qué pasa con los derechos de estas IA capaces de sentir? ¿Deberíamos tratarlas con el mismo respeto que a un ser vivo natural? ¿Cómo podemos evitar que estas entidades sean corrompidas, utilizadas para hacer daño, o abusadas?", se pregunta.