INTELIGENCIAS POCO INTELIGENTES

Por: Esteban Magnani. 09/04/2025

Las IA no piensan: solo procesan estadísticas y datos producidos por humanos, con sus errores y prejuicios. El peligro de considerar como verdad absoluta todo lo que digan.

Cualquier persona que haya cocinado alguna vez sabe que por mucho esfuerzo que ponga en seguir una receta, si los ingredientes son de mala calidad, el plato no quedará bien. Lo que ocurre con la Inteligencia Artificial Generativa (IAG), como puede ser ChatGPT, no es tan distinto.

En este tipo de tecnologías, los ingredientes serían los datos de entrenamiento y la receta ocuparía el lugar del software que los procesa en busca de patrones que permiten conocer las posibilidades estadísticas de que, por ejemplo, a una palabra la siga otra. Por mucho que el software busque patrones en los datos acumulados, si estos son de mala calidad o insuficientemente variados, los resultados van a ser similares por más que su apariencia sea convincente.

Por eso, creer que la IAG dará respuestas verdaderas o neutrales es falso: «Nosotros somos personas atravesadas por una serie de contextos, de subjetividades, de condicionantes al momento de vincularnos, elaborar discursos o pensarnos: lo mismo pasa con los algoritmos que generan predicciones», explica Micaela Sánchez Malcolm, docente universitaria (UBA) especializada en temas de IA. «Estos modelos de lenguaje, por ejemplo, están basados en contextos, en sociedades que tienen sesgos, que tienen limitaciones, que tienen diferentes perspectivas y que se trasladan, con diferentes matices, a los modelos de lenguaje que procesan los datos producidos por personas. Por eso no se podría hacer una IAG neutral». Sin embargo, como aclara la especialista –quien también fue secretaria de Innovación Pública de la Nación–, «es muy común atribuir un aura de neutralidad a las IA».

Esta idea de que la IAG encuentra verdades que los humanos no pueden hallar es la que se repite en cada vez más titulares en los que se asegura que «la IA dice que…». Esto no tiene mucho sentido ya que, por empezar, las IAG son muchas y de

distintos tipos: las hay que pueden escribir, dibujar o hacer canciones a partir de algunas instrucciones. Incluso si se le repite una pregunta varias veces a la misma IA dará respuestas distintas. Por eso las IAG pueden servir para legitimar una conclusión que se buscaba originariamente.

«Hay una idea de que un sistema de inteligencia artificial entrenado con cientos de miles de papers académicos de las revistas científicas puede resolver casos que ningún equipo médico podría por más alto nivel que tenga», explica Carolina Martínez Elebi, docente (UBA) y miembro del Observatorio de Impactos Sociales de la IA de Untref. «En realidad, cualquier persona que va a un médico y después a otro ve que dentro de la misma profesión cada uno tiene su librito. ¿Cómo harían entonces las IAG para llegar a algo definitivo? ¿Con qué criterio este sistema de entrenamiento automático decide cuál de todos esos libritos es más relevante?».

Datos

«Hay un discurso hegemónico vinculado con la IA según el cual habría un nivel de procesamiento de información, de datos, que permitiría llegar a una idea de "verdad"», explica Sánchez Malcolm. «La cuestión es qué representatividad tienen esos datos. Si tenemos que analizar estudios médicos, no es lo mismo evaluar nuestra constitución que la de personas que viven en zonas nórdicas y miden 20 o 25 centímetros más que nosotros. Los estudios que se hagan tomando esos datos darán resultados basados en gente que no es a la que se está analizando».

Es importante comprender que la IAG no accede a la realidad más que procesando datos generados por humanos y esos datos contienen sesgos, errores, exageraciones o pueden no ser relevantes en otros contextos. Por ejemplo, si en la mayoría de los textos de entrenamiento aparece que las enfermeras son mujeres y los médicos, varones, una IAG tenderá a producir contenidos que repitan eso mismo. Es decir, que la IAG tenderá a reforzar lo que ya ocurre, obturando o, al menos, no facilitando el cambio o una mejora.

Un caso muy famoso en ese sentido ocurrió en Amazon: allí se utilizaba una IA para automatizar la preselección de nuevos trabajadores. Como esa IA había sido entrenado con los CV de los empleados contratados en los últimos años y la mayoría de ellos habían sido varones, tendía a descartar a las mujeres. Cuando la empresa comprendió lo que estaba ocurriendo, dio de baja el sistema; ¿pero qué

podría haber ocurrido si nadie lo notaba y se aceptaban los resultados confiando ciegamente en la estadística?

Entrenamiento

Hay quienes sostienen que los problemas de sesgos se pueden resolver utilizando datos en mayor cantidad y variedad. Para Martínez Elebi eso conlleva nuevos riesgos: «Diversificar los tipos de datos que se tienen a la hora de entrenar esos sistemas es un arma de doble filo. Por ejemplo, si entrenamos una IA con imágenes de personas, podemos buscar que sean de la mayor diversidad posible en términos de género y edad. Hoy por hoy, por ejemplo, hay muchas más imágenes de adultos jóvenes; entonces una IA entrenada con eso luego no reconoce el rostro de una persona mayor. Por eso, lo que más ayuda es contar con mayor volumen y diversidad. Pero el arma de doble filo es que, sobre todo pensando por ejemplo en datos biométricos o sensibles, terminás teniendo mucha mayor probabilidad de que se vulneren derechos como la privacidad: hay mucha mayor probabilidad de que esos datos se crucen y efectivamente se pierda el anonimato».

Por otra parte, detectar la fuente de los sesgos en los resultados de una IAG para corregirlos es muy complicado ya que estas funcionan como cajas negras en las que interactúan numerosas variables. De esa manera resulta imposible saber exactamente cómo se llegó a un resultado, incluso si se conocen los datos de entrenamiento.

¿Cómo se resuelve este problema? «Lo que se obtenga de una IA tiene que ser chequeado, validado con fuentes de referencia, es decir, que tiene que ser validado por un criterio humano no solamente de una persona sino de equipos multidisciplinares respecto de qué está procesando, qué se está consultando y qué resultados están obteniendo», explica Sánchez Malcom. El problema es que en la medida en que más tareas se deleguen en IA cada vez habrá menos gente en condiciones de entender lo que hacen y detectar errores.

Resulta fundamental insistir en que las IAG no son inteligentes, sino que usan estadísticas. No vaya a ser que el plato que nos sirven, aunque tenga una apariencia suculenta, no contenga en realidad nada nutritivo.

LEER EL ARTÍCULO ORIGINAL PULSANDO AQUÍ

Fotografía: Hamartia

Fecha de creación
2025/04/09