

A pesar del éxito actual de los *Large Language Models (LLM)*, la evolución de la inteligencia artificial podría tocar techo si no se combina la aproximación actual mediante el aprendizaje automático con otras, como puede ser la inteligencia artificial simbólica.

El éxito reciente del chatbot ChatGPT ha puesto en el centro del debate público la inteligencia artificial, una tecnología que ha conocido un avance considerable a lo largo de la década pasada de forma que ya está presente en numerosos aspectos de nuestras vidas. Aunque ya tratamos con sistemas inteligentes de forma cotidiana, muchas veces sin darnos cuenta (por ejemplo, cuando las plataformas nos recomiendan contenidos audiovisuales personalizados o cuando utilizamos el texto predictivo al escribir un mensaje en el móvil), el producto de la empresa Open.AI nos ha enfrentado a un nuevo fenómeno asaz vistoso, como es contemplar la capacidad creativa de las maquinas.

Este tipo de inteligencia artificial se denomina “generativa”, pues da lugar a algún tipo de creación, ya sea un

texto, una imagen o sonido. A pesar de lo novedoso que parecen estas habilidades, ya existe una larga tradición de algoritmos que escriben novelas o poesía, o que pintan cuadros o diseñan imágenes. En 2016 saltaba la noticia de que un programa de inteligencia artificial japonés había escrito una novela que casi gana un premio literario[1], y, dos años antes, un estudiante del MIT programó un sistema para que escribiese sonetos como salidos de la pluma del mismísimo William Shakespeare[2]. ¿A qué se debe entonces el revuelo actual en torno a estas máquinas creadoras?

Es probable que la popularidad que han cosechado los llamados *Large Language Models (LLM)* -como ChatGPT, Turing NLG de Microsoft, Bard de Google o Gopher de Deep Mind- resida en que han abandonado el ámbito de los expertos y están en la mayoría de los casos en manos de la gente. De esta forma, todo el mundo ha podido probar en primera persona cómo funcionan estos chatbots y las posibilidades que ofrecen.

Ahora bien, a pesar de lo espectacular de los resultados que ofrecen, existen voces que señalan las limitaciones de esta rama de la inteligencia artificial, y lo lejos que está de emular las funciones del cerebro humano. El pasado agosto, el físico Michio Kaku denunciaba el sensacionalismo que se ha generado en torno a estos robots conversacionales, que, a su juicio, no hacen más que seleccionar y ordenar contenidos de internet, sin saber discernir la verdad de la ficción, ni distinguir los datos reales de la desinformación. Básicamente, los LLM no serían más que versiones avanzadas de los buscadores de la web, como Google, que han estado funcionando durante los últimos veinte años[3].

Existen voces que señalan las limitaciones de esta rama de la inteligencia artificial y lo lejos que está de emular las funciones del cerebro humano

Uno de los mayores críticos de la inteligencia artificial actual es sin duda el profesor y empresario Gary Marcus, quien considera que la rama más utilizada hoy en día, el aprendizaje automático y las redes neuronales -responsables de los avances producidos en los últimos veinte años-, presenta carencias relativas al razonamiento, no resulta fiable y se aleja de la forma de pensar que tenemos los humanos. A su juicio, las carencias que presenta el aprendizaje profundo de las redes neuronales podrían ser salvadas gracias a la denominada inteligencia artificial simbólica, una aproximación clásica a esta tecnología que hoy en día ha sido desplazada a un segundo plano.

IA simbólica vs. *machine learning*

La oleada actual de inteligencia artificial en la que nos vemos inmersos está basada en el aprendizaje automático (*machine learning*) y el aprendizaje profundo (*deep learning*) o redes neuronales, y hace uso del *big data*, es decir, se nutre de inmensas cantidades de información para poder construir modelos. Con frecuencia se compara el deep learning con la neurología, pues la inteligencia artificial permite construir sistemas que imitan el funcionamiento del cerebro en la forma en que este clasifica la información recibida, por ejemplo, identificando objetos en función de los rasgos que contienen. Son programas que elaboran sus diagnósticos o predicciones en base a la probabilidad, y llevan a cabo una forma de aprendizaje que contempla la retroalimentación: en sucesivas vueltas reciben el grado de acierto de sus dictámenes y realizan

las correcciones oportunas.

Por otra parte, la inteligencia artificial simbólica se centra en la manipulación y el procesado de símbolos y conceptos en vez de grandes cantidades de datos numéricos. Este tipo de algoritmos funciona manejando símbolos que representan objetos o ideas del mundo y sus relaciones. La principal aproximación de esta rama es el uso de programación basada en la lógica, en la que las normas y los axiomas son utilizados para realizar inferencias y deducciones.

Mientras que la inteligencia artificial simbólica es más adecuada para trabajar en entornos de conocimiento bien definido y estructurado, el aprendizaje automático es más útil cuando existen grandes volúmenes de datos y patrones complejos.

La inteligencia artificial simbólica es más adecuada para trabajar en entornos de conocimiento bien definido

La tesis que defiende Gary Marcus es que la inteligencia artificial actual basada en el aprendizaje automático y el aprendizaje profundo solamente podrá seguir evolucionando si se combina con el enfoque simbólico, en la forma de una inteligencia artificial "híbrida". De no ser así, predice que se podría producir otro "invierno de la inteligencia artificial", es decir, un estancamiento y un desencanto generalizado en torno a esta tecnología, como el que tuvo lugar a partir de los años 70 del siglo pasado.

Principios de una inteligencia artificial fiable

Marcus establece los dieciséis principios que debería cumplir una inteligencia artificial fiable y robusta -más cercana al funcionamiento del cerebro humano-, y que, en su opinión, los desarrollos actuales no presentan, por lo menos, en su totalidad:

1. Explicación: los sistemas inteligentes deberían poder recopilar y presentar su línea de razonamiento detrás de cada respuesta que ofrecen.
2. Deducción: la inteligencia artificial debería poder realizar el mismo tipo de deducciones que hacen las personas, por ejemplo, "si los países tienen fronteras y Andorra es un país, entonces Andorra tiene fronteras".
3. Inducción: como complemento a lo anterior, utilizarla cuando ciertas conclusiones no pueden obtenerse de la lógica deductiva.
4. Analogía: establecer paralelismos entre cosas que en principio no están relacionadas.
5. Razonamiento abductivo: determinar la explicación más probable de una situación o problema de entre todas las posibles.
6. Teoría de la mente: una inteligencia artificial debería saber adaptar su discurso en función de los conocimientos de su interlocutor, como hacemos los humanos, dado que no hablamos igual con un niño que con un adulto, o con alguien experto en un tema que quien no lo es.
7. Fluencia de cuantificación: utilizar variables lógicas cuantificadas para evitar ambigüedades. Por ejemplo, la frase "todo sueco tiene un rey" no significa que haya tantos reyes como suecos en Suecia.
8. Fluencia modal: el lenguaje humano utiliza operadores modales del tipo de "Juan piensa que…", "él teme

que…” o “es posible que…”. Los algoritmos deben aprender a interpretarlos y manejarlos igual que hacemos las personas.

9. Anulabilidad: gran parte de la información que recibimos y de las creencias que tenemos son solamente verdaderas por defecto, y es posible que nos llegue nueva información que las altere. La inteligencia artificial debe ser capaz de asimilar nueva información y revisar y reconsiderar conclusiones a las que haya llegado en el pasado.
10. Argumentos a favor y en contra: muchas cuestiones complejas del mundo real no tienen una solución clara, y con frecuencia utilizamos argumentos a favor y en contra a la hora de tomar una decisión. Los algoritmos deben aprender a utilizar el razonamiento de pros y contras para tratar problemas sin una única solución.
11. Contexto: la comunicación humana depende en gran medida del contexto en el que se produce, y a menudo éste contiene elementos implícitos que conocen los interlocutores. Un sistema inteligente, como los robots conversacionales, debe ser capaz de reconocer en qué contexto se está produciendo la interlocución.
12. Meta conocimiento y meta razonamiento: la inteligencia artificial, al igual que el ser humano, debe ser capaz de razonar acerca de su propio conocimiento, incluyendo la historia y procedencia de cada hecho o regla del mismo, y, de esta forma, poseer un modelo realista y preciso de lo que sabe y no sabe, y de su grado de habilidad en las distintas tareas.
13. Explícitamente ética: una inteligencia artificial fiable debería seguir un conjunto de principios inviolables, como no mentir o no causar daños de ningún tipo.
14. Velocidad suficiente: los algoritmos deben ser capaces de dar respuesta con la rapidez suficiente que requiera cada tipo de problema planteado.
15. Suficientemente lingual y corpóreo: algunas aplicaciones deberán poder entender y hablar el lenguaje natural y entender varios idiomas si resulta necesario, y, también, según la funcionalidad que cumplan, deben poder ver para reconocer objetos, medir la luminosidad, moverse, detectar olores, etc.
16. Conocimientos amplios y profundos: al igual que ocurre con las personas, la inteligencia artificial debería tener un vasto conocimiento fundamental del mundo que le rodea, para poder contrastar, evaluar y poner en contexto cada nueva pieza de información que recibe.

La inteligencia artificial híbrida

La solución a estas carencias que apunta Marcus podría estar en lo que denomina “modelos híbridos” de inteligencia artificial, que combinan lo mejor de las dos aproximaciones: el aprendizaje basado en datos de las redes neuronales con las poderosas capacidades de abstracción implícitas en la manipulación simbólica. El problema es que a menudo estas dos visiones han sido concebidas como antagónicas, pero no tiene por qué ser así. Mientras que el aprendizaje profundo extrae patrones de grandes conjuntos de datos no estructurados recogidos del mundo real, los métodos basados en reglas pueden realizar la manipulación simbólica de los datos recogidos, creando abstracciones de estos. Sin ir más lejos, el buscador de Google es un modelo híbrido.

Mientras que el aprendizaje profundo extrae patrones de grandes conjuntos de datos no estructurados recogidos del mundo real, los métodos basados en reglas pueden realizar la

manipulación simbólica de los datos recogidos

Esta tercera vía de innovación está siendo explorada por IBM a través de lo que denomina *inteligencia artificial neuro simbólica*, un planteamiento experimental que combina el uso de redes neuronales convolucionales dedicadas a la clasificación de imágenes con la capacidad de establecer relaciones entre elementos simbólicos, como secuencias de palabras, de forma que el sistema aprende de sus errores de una forma mucho más rápida que mediante el entrenamiento clásico de una red neuronal.

Para Gary Marcus el futuro de la inteligencia artificial no puede basarse en modelos basados en un exceso de datos y en una carencia de conocimiento y raciocinio. Necesitamos sistemas que manifiesten una elevada comprensión del mundo, lo que conlleva centrarnos en cómo se representa, adquiere y razona el conocimiento abstracto.

Imagen de Gerd Altmann en Pixabay