

ChatBot AI

OK I can answer your question in several ways:

- 1) The neural network that provides you with the math should be constantly monitored
- 2) Adjustments should be made to any feedback loop providing corrections
- 3) Statistical data may need to be updated
- 4) Analysis of that statistical data has to be checked / fact c
- 5) Inputs have to be recog understood

El desarrollo de la IA presenta grandes retos tecnológicos, de seguridad y éticos. Es una gran herramienta la servicio de la humanidad que hay que regular convenientemente.

CHATGPT Y EL TEST DE TURING

Recientemente hemos sabido que el ChatGPT —herramienta sencilla, capaz y precisa que aún casi todos los adjetivos para convertir a alguien inexperto en un perfecto profesional de casi cualquier campo— ha pasado el Test de Turing. Esta prueba, propuesta por el matemático y científico de la computación Alan Turing en 1950, para evaluar la capacidad de una máquina para exhibir un comportamiento inteligente equivalente o indistinguible del de un ser humano, ha sido objeto de debate y críticas a lo largo de los años (algunos argumentan que superar el test no garantiza automáticamente la presencia de una verdadera inteligencia), pero sigue siendo un concepto fundamental en el campo de la inteligencia artificial y la ética de la tecnología.

Texto de **PABLO BUSTOS TÁULER**,
coordinador DOIU Escuela Politécnica Superior, experto
en inteligencia artificial, Máster Univ. Francisco de Vitoria



ue la Inteligencia Artificial (IA) ha llegado para quedarse es un hecho. Este impacto es comparable al que supuso la llegada de internet, aquello cambió la manera de hacer las cosas a todos los niveles. Esto sugiere que aquellas personas que no sepan qué es la IA tienen la casi obligación de aprender a manejar la herramienta o, por lo menos, conocer qué es y de qué va. Luego, ya el día a día les obligará a utilizarla de una u otra manera. Quizá lo más conocido de la IA sea el ChatGPT. Esta herramienta es sencilla, capaz, precisa. Aúna casi todos los adjetivos para convertir a alguien inexperto en un perfecto profesional de casi cualquier campo. Resuelve casos de psicología, de jurisprudencia, realiza estrategias comerciales, incluso es capaz de pintar cuadros reproduciendo diferentes estilos. Y lo mejor de todo, es prácticamente infinita. Esto puede tener dos lecturas: bien utilizada puede aportar mucho y mal utilizada puede hacer mucho daño. Como casi todo en la vida. Recientemente hemos sabido que

el ChatGPT ha pasado las pruebas del Test de Turing y esto, ciertamente, es un avance. Pero para ponernos en contexto, detallaré qué es el test de Turing para entender mejor qué significa esto.

EL TEST DE TURING ES UNA PRUEBA PROPUESTA POR EL MATEMÁTICO Y CIENTÍFICO DE LA COMPUTACIÓN ALAN TURING en 1950 para evaluar la capacidad de una máquina para exhibir un comportamiento inteligente equivalente o indistinguible del de un ser humano. La prueba se centra en la capacidad de una máquina para mantener una conversación escrita con un humano, de manera que el humano no pueda determinar si está interactuando con otra persona o con una máquina. En la versión clásica del Test de Turing, un humano (el juez) realiza conversaciones escritas en lenguaje natural con dos participantes, uno de los cuales es una máquina y el otro es un humano. Si el juez no puede distinguir consistentemente entre las respuestas de la máquina y las del humano, entonces se considera que la máquina ha pasado el Test de Turing y ha demostrado una forma de inteligencia artificial. Es importante mencionar que el Test de Turing no es necesariamente un indicador definitivo de la inteligencia artificial, ya que se centra en la apariencia de inteligencia más que en la verdadera comprensión y cognición. También ha sido objeto de debate y críticas a lo largo de los años, ya que algunos argumentan que superar el Test de Turing no garantiza automáticamente la presencia de una verdadera inteligencia. A pesar de esto, el Test de Turing sigue siendo un concepto fundamental en el campo de la inteligencia artificial y la ética de la tecnología.

El matemático inglés Alan Turing consiguió descifrar los códigos de los mensajes que los alemanes enviaban en la Segunda Guerra mundial a través de la máquina Enigma.



POR LO TANTO, ¿QUÉ SUPONE QUE EL CHATGPT HAYA PASADO ESTA PRUEBA? ¿En qué nos estamos metiendo? ¿Avanza la IA demasiado rápido? El hecho de que el ChatGPT supere el Test de Turing significaría que ha demostrado la capacidad de generar respuestas y participar en conversaciones de una manera que es indistinguible de las respuestas y conversaciones humanas para un observador externo. En otras palabras, el observador no sería capaz de determinar si las respuestas provienen de un humano o de la máquina durante una serie de interacciones escritas. Esto implicaría que el ChatGPT ha alcanzado un nivel de sofisticación y habilidad en la generación de lenguaje natural que le permite imitar con éxito el comportamiento humano en una conversación. Sin embargo, es importante destacar que superar el Test de Turing no necesariamente significa que la máquina tenga una comprensión real o una inteligencia auténtica. Podría generar respuestas convincentes basadas en patrones y datos aprendidos, pero no necesariamente demostraría una verdadera comprensión o conciencia.

Pero además, la tecnología GPT-3 y sus variantes, incluido el ChatGPT, han sido sometidos a una variedad de pruebas y diagnósticos para evaluar su rendimiento y capacidad en diversos aspectos del procesamiento del lenguaje natural. Algunas de estas pruebas incluyen: Bleu y Rouge (estas métricas se utilizan para evaluar

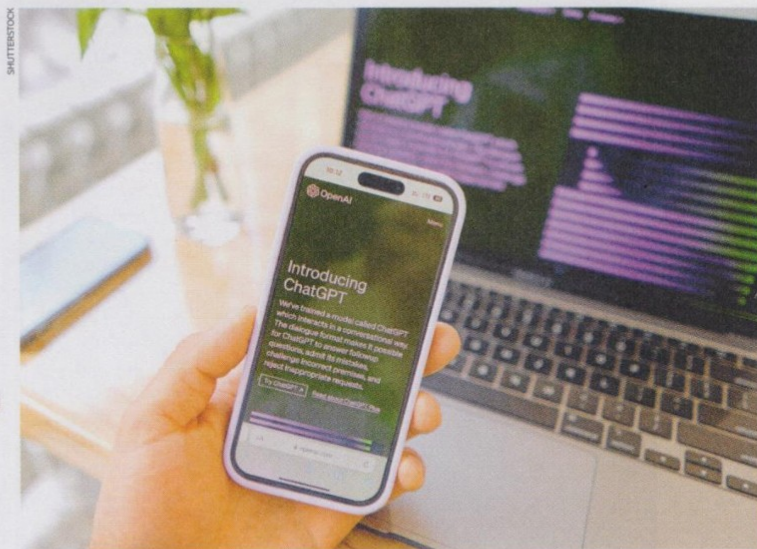
Que una máquina supere el Test de Turing no significa que esta tenga una comprensión real o una inteligencia auténtica

X realiza las preguntas



Aprobada si X no detectaba
quién le respondía

En el Test de Turing el juez conversa por escrito con dos participantes: un humano y una máquina. Si no puede distinguirlos se considera que la máquina ha superado el test.



Toda la tecnología GPT-3 ha sido sometida a diferentes pruebas y diagnósticos para evaluar su rendimiento y su capacidad en comprensión de lectura, calidad de las traducciones, etc.

la calidad de las traducciones automáticas y la generación de resúmenes, respectivamente, midiendo la similitud entre las respuestas generadas por el modelo y las respuestas de referencia humanas); SQuAD (GPT-3 que ha sido evaluado en el conjunto de datos SQuAD; Stanford Question Answering Dataset), que consiste en preguntas y respuestas basadas en pasajes de texto. El rendimiento se mide en función de la precisión de las respuestas generadas), GLUE y SuperGLUE (estos conjuntos de pruebas evalúan la capacidad de los modelos de lenguaje para realizar tareas de procesamiento del lenguaje natural, como comprensión de lectura, clasificación de sentencias y tareas de relación entre oraciones); Lectura Comprensiva (las pruebas de lectu-

ra comprensiva evalúan la capacidad del modelo para responder preguntas sobre un pasaje de texto dado). Ejemplos, incluye el conjunto de pruebas RACE (Reading Comprehension from Examinations) y el conjunto de datos MCTest y Generación Creativa (además de las pruebas centradas en la precisión, GPT-3 también se ha evaluado en su capacidad para generar contenido creativo, como poesía, historias cortas y diálogos).

ANTE LA PREGUNTA DE SI AVANZA DEMASIADO rápido la IA o de si es bueno o malo que esto sea así, es un tema complejo y multidimensional que depende de varios factores y perspectivas. Aquí detallo algunos puntos a considerar sobre el avance rápido de la IA:



El desarrollo de la IA en campos como la medicina o en la automatización de ciertas tareas supone grandes beneficios: el hallazgo de nuevos tratamientos médicos y el avance de la investigación científica y la liberación del hombre de ciertas tareas productivas.

- **Innovación y progreso tecnológico:** un avance rápido en la IA puede conducir a soluciones innovadoras en diversas áreas, como la atención médica, la energía, la educación y la investigación científica. Puede permitir el desarrollo de tecnologías que resuelvan problemas complejos de manera más eficiente y efectiva.
- **Automatización y eficiencia:** la IA avanzada puede automatizar tareas repetitivas y tediosas, lo que podría liberar a los humanos para enfocarse en tareas más creativas, estratégicas y de mayor valor.
- **Descubrimientos científicos:** la IA puede ayudar en la identificación de patrones y tendencias en grandes conjuntos de datos, lo que puede llevar a nuevos descubrimientos científicos y avances en campos como la medicina y la biología.
- **Desarrollo económico:** impulsando la innovación y creando nuevas oportunidades económicas en forma de empleos en tecnología, *startups* o nuevas industrias

La IA está llamada a cambiar nuestras vidas y la manera en que las cosas se hacen. Es un hito comparable a la llegada de internet allá por el año 83. Aquello revolucionó la sociedad como ya no recordamos y la inteligencia artificial aún lo hará más.

Y aunque hablamos maravillas de la IA, falta mucho por hacer. De hecho, nunca terminará de hacerse. Siempre habrá mejoras que implantar. Algunos ejemplos del recorrido que falta por hacer serían:

- **Detección de amenazas:** la IA se utiliza para analizar grandes cantidades de datos y patrones en tiempo real para identificar actividades sospechosas en las redes y sistemas. Los algoritmos de aprendizaje automático pueden detectar anomalías que podrían ser indicativas de ataques cibernéticos, como intrusiones o *malware*.
- **Análisis de *malware*:** los sistemas de IA pueden analizar el comportamiento de los programas y archivos en busca de indicadores de *malware*. Pueden identificar y clasificar nuevas amenazas de manera más rápida que los enfoques tradicionales de firmas de virus.
- **Autenticación biométrica:** la IA se utiliza en el reconocimiento facial y de voz, para garantizar que solo usuarios autorizados tengan acceso a sistemas y datos sensibles. Ya existen varios casos de suplantaciones de identidad por imitación de voces.
- **Seguridad de la red:** los sistemas de IA pueden analizar el tráfico de red y detectar patrones anómalos que podrían indicar un ataque DDoS (Denegación de Servicio Distribuido) u otros tipos de ataques de red.
- **Protección contra *phishing*:** la IA puede ayudar a identificar correos electrónicos de *phishing* mediante

La IA también plantea grandes desafíos, como que obtenga demasiada autonomía o que acentúe la brecha de desigualdad

el análisis de contenido y la detección de enlaces o archivos maliciosos.

- **Automatización de respuestas:** la IA puede automatizar la respuesta a incidentes de seguridad, lo que permite tomar medidas rápidas para contener y mitigar amenazas en tiempo real.

- **Predicción de amenazas:** los modelos de inteligencia artificial pueden analizar datos históricos y patrones de amenazas para predecir posibles futuros ataques y ayudar a las organizaciones a fortalecer su seguridad antes de que ocurran.

- **Gestión de contraseñas y acceso:** la IA puede ayudar a gestionar y fortalecer las políticas de contraseñas, así como identificar y corregir debilidades en la autenticación de usuarios.

- **Analítica de seguridad:** la IA puede ayudar a las organizaciones a analizar grandes cantidades de datos de seguridad y generar informes más detallados sobre las amenazas y vulnerabilidades.

- **Respuesta adaptativa:** ajustando las defensas de seguridad en tiempo real en función de las amenazas detectadas, lo que permite una respuesta más dinámica y efectiva.

- **Otras, como la seguridad cibernética, la privacidad, el desempleo tecnológico y la toma de decisiones éticas.** Por lo tanto, el desarrollo y la implementación de la IA deben ser cuidadosamente gestionados para asegurar que los beneficios superen los riesgos y se utilicen de manera ética y responsable.

La regulación adecuada y la consideración de los impactos sociales y éticos son esenciales a medida que la IA continúa avanzando. □



¿Cómo piensa la IA? Escanea este código QR y embarcate en un viaje al cerebro de una máquina. Descubre cómo los algoritmos de aprendizaje automático imitan la mente humana.

El desarrollo de la Inteligencia Artificial está cambiando nuestra vida diaria. En la imagen, supermercado en Portugal totalmente controlado por la IA, con cámaras en el techo que controlan lo que se retira de las estanterías y sensores de peso.

También modificará nuestra movilidad, haciendo que sea más eficiente y sostenible, como los aviones eléctricos verticales.

