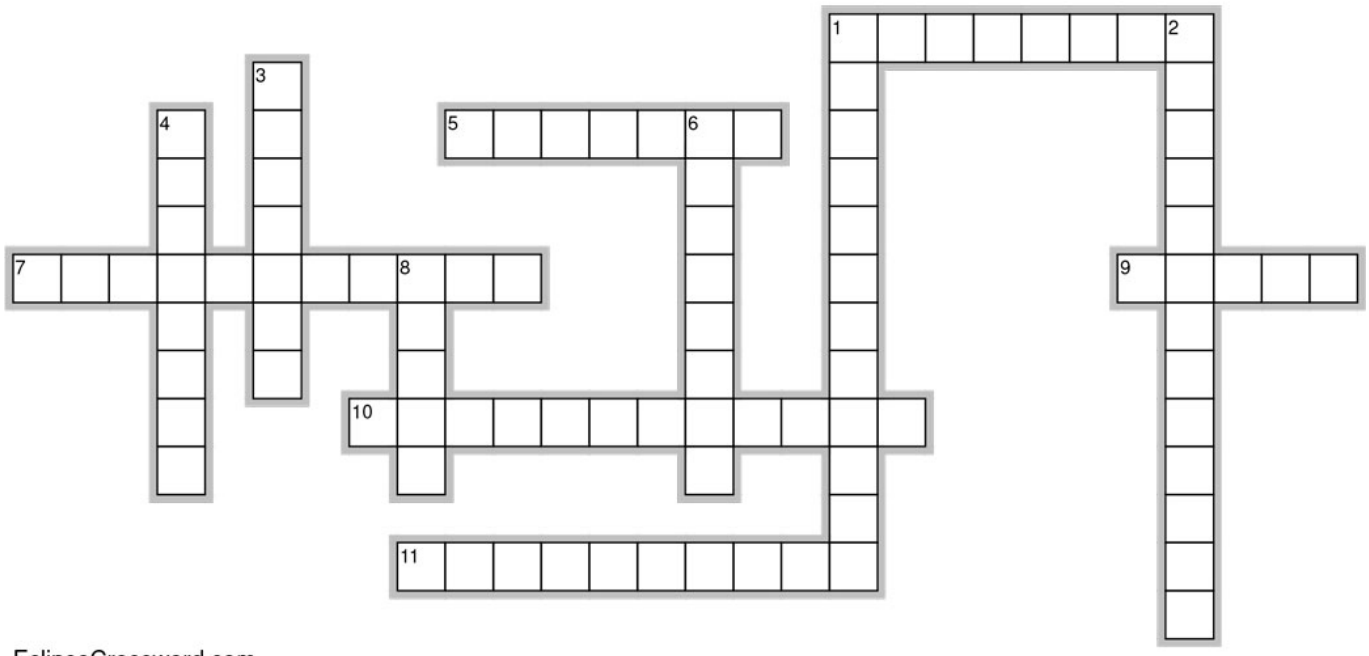
La inteligencia de las máquinas

30/11/2023

LA INTELIGENCIA DE LAS MÁQUINAS

A. REQUENA @ VALLE DE LA CIENCIA, 2023



EclipseCrossword.com

HORIZONTALES

1. El avance de la tecnología actual de la llamada Inteligencia artificial requiere reemplazar la prueba de Turing original con un examen del razonamiento de uno de ellos.
5. Estamos inmersos en un océano de éstos.
7. El experimento de la Habitación China cuestiona si esta imitación es realmente equivalente a la comprensión o ésta.
9. Mientras que el Test de Turing evalúa la inteligencia artificial desde una perspectiva de comportamiento externo, la Habitación China aborda preguntas más profundas sobre ésta y la consciencia.
10. El Test de Turing, desarrollado por Alan Turing, en los 50, es un criterio para evaluar la de una máquina.
11. Es necesario no solo observar la funcionalidad externa del programa, sino también comprender éstos y técnicas subyacentes que permiten estas capacidades.

VERTICALES

1. Estos experimentos han contribuido significativamente a nuestra comprensión de la psicología humana, aunque algunos, han sido criticados por cuestiones éticas.
2. Profundizar en ella es una forma de sondear la comprensión del razonamiento por las máquinas.
3. El objetivo del test de Turing es determinar si el evaluador puede distinguir cuál de los dos es ésta.
4. Los Chatbots parecen superar los test de Ti¡uring y la cuestión es cómo sabemos si una máquina realmente lo está haciendo.
6. En el experimento de Milgram sobre la Obediencia a la autoridad, realizado en la década de los sesenta del siglo pasado, se investigó hasta qué punto las personas lo hacen a una orden que suponía hacer daño a otra persona, solo porque la autoridad lo ordenaba.
8. El experimento de esta habitación, debido a John Sie, argumenta contra la idea de que la simple manipulación de símbolos, como lo hace un ordenador, pueda ser equivalente a la comprensión o la consciencia.

El Test de Turing, desarrollado por Alan Turing, en los 50, es un criterio para evaluar la inteligencia de una máquina. En este test, un evaluador humano interactúa con una máquina y con un ser humano a través de una interfase, que oculta sus identidades físicas. El objetivo del test es determinar si el evaluador puede distinguir cuál de los dos es la máquina. Si la máquina logra engañar al evaluador haciéndole creer que es humana, se considera que ha pasado el Test de Turing. Este test ha sido fundamental en el campo de la inteligencia artificial, ya que propone un desafío significativo para los desarrolladores de sistemas inteligentes, impulsando el avance hacia máquinas cada vez más sofisticadas y capaces de emular comportamientos humanos.



Imagen creada con Chat-GPT – DALL-E.

Una propuesta a raíz de la Inteligencia Artificial, es el experimento de la habitación china, debido a John Sie, argumentando contra la idea de que la simple manipulación de símbolos, como lo hace un ordenador, pueda ser equivalente a la comprensión o la consciencia. Consiste en que una persona que no habla chino está en una habitación, utilizando reglas para manipular símbolos chinos en respuesta a la entrada de símbolos chinos. Aunque desde el exterior parece que entiende chino, en realidad solo está siguiendo instrucciones sin comprensión. Sitúa el foco en la diferencia entre la mera manipulación de símbolos y la verdadera comprensión o consciencia. El Test de Turing se centra en la capacidad de la máquina para imitar la inteligencia humana de manera convincente, mientras que el experimento de la Habitación China cuestiona si esta imitación es realmente equivalente a la comprensión o la consciencia. Mientras que el Test de Turing evalúa la inteligencia artificial desde una perspectiva de comportamiento externo, la Habitación China aborda preguntas más profundas sobre la mente y la consciencia.



Imagen creada con Chat-GPT – DALL-E.

Recientemente se ha ideado una alternativa de la prueba de Turing, intentando comprender mejor cuando una máquina o Inteligencia Artificial está pensando o simplemente manipulando datos sin consciencia de lo que hace. Es cierto que los avances en las herramientas de la Inteligencia Artificial han tornado obsoletos algunos de los procedimientos diagnóstico para la valoración del contenido inteligente en la conducta de las máquinas.

Una de las aplicaciones muy difundidas es Tinder. En ella los usuarios crean un perfil con fotos, una breve descripción y preferencias para la búsqueda de pareja. Los perfiles de posibles coincidencias aparecen en la pantalla del usuario. Si le gusta lo que ve, el usuario desliza hacia la derecha (o toca el corazón) para "gustar" indicando que acepta el perfil. Si no está interesado, desliza hacia la izquierda (o toca la X) para pasar al siguiente perfil. En el caso de que dos usuarios se "gustan" mutuamente, es decir, ambos han deslizado hacia la derecha en el perfil del otro, se produce una "coincidencia" y pueden empezar a chatear dentro de la aplicación. Una vez que se establece una coincidencia, los usuarios pueden enviar mensajes de texto dentro de la aplicación para conocerse mejor. Tinder pone un énfasis particular en la seguridad y privacidad. Los usuarios solo pueden enviar mensajes a personas con las que han coincidido, y pueden deshacer una coincidencia o reportar perfiles sospechosos si es necesario. La

aplicación ofrece funciones adicionales como Tinder Plus o Tinder Gold, que incluyen opciones como 'Me Gusta' ilimitados, 'Rewinds' (para deshacer deslices anteriores) y la capacidad de ver a quién le gustaste antes de deslizar. Tinder ha revolucionado el mundo de las citas online, haciéndolo más accesible y menos formal en comparación con los métodos tradicionales de citas en línea. Su enfoque en la ubicación y la interacción basada en el deslizamiento facilita la búsqueda de coincidencias potenciales en un entorno más casual y lúdico.

Por otro lado, estamos inmersos en un océano de chatbots. Muchas empresas utilizan chatbots para responder preguntas frecuentes, resolver problemas sencillos y proporcionar asistencia 24/7 a los clientes. Estos bots pueden estar en sitios web, aplicaciones de mensajería y redes sociales. En el comercio electrónico, los chatbots ayudan a los clientes con recomendaciones de productos, procesamiento de pedidos y seguimiento de envíos. Algunos chatbots en el sector de la salud proporcionan información básica sobre síntomas y cuidados de salud, aunque siempre bajo la recomendación de consultar con profesionales médicos para diagnósticos y tratamientos. En el ámbito educativo, los chatbots pueden ayudar a los estudiantes con tutorías automatizadas, responder a preguntas frecuentes sobre cursos y programas y asistir en el proceso de aprendizaje. Los chatbots en el sector financiero asisten a los clientes con consultas sobre transacciones, saldos de cuentas y pueden ofrecer consejos básicos sobre finanzas personales. En el mundo del entretenimiento, los chatbots pueden actuar como asistentes personales, recomendando películas, música o videojuegos basados en las preferencias del usuario. En la asistencia personal han proliferado por todas partes: Siri, Alexa, Google Assistant y otros asistentes personales son ejemplos avanzados de chatbots que pueden realizar una amplia gama de tareas, desde configurar alarmas hasta controlar dispositivos inteligentes en el hogar o marcvar números de teléfono. Muchas plataformas de redes sociales y aplicaciones de mensajería tienen chatbots integrados para diversas funciones, desde interactuar con marcas hasta proporcionar actualizaciones automatizadas. Estos chatbots varían en complejidad, desde simples respuestas programadas hasta sistemas avanzados de inteligencia artificial capaces de aprender y adaptarse a las necesidades del usuario. Su creciente prevalencia es un testimonio de la rapidez con la que la tecnología está avanzando y cambiando la forma en que interactuamos con los servicios digitales.

Estos sistemas parecen superar los test de Tiiuring y la cuestión es cómo sabemos si una máquina realmente está pensando. La propuesta que se hace en una publicación de la revista Intelligent Computing es tratar a la máquina como si fuera un participante en un estudio psicológico.

La propuesta de los investigadores, Philip Nicholas Johnson-Laird de la Universidad de Princeton y Marco Ragni de la Universidad Tecnológica de Chemnitz, consiste en un proceso de tres pasos que debe completar la máquina para responder sobre su capacidad de pensar por sí misma: a) Pruebas en experimentos psicológicos; b) Autorreflexión y c) Examen del código fuente.

Los experimentos psicológicos han contribuido significativamente a nuestra comprensión de la psicología humana, aunque algunos, han sido criticados por cuestiones éticas. Los experimentos psicológicos modernos suelen estar sujetos a estrictas normas éticas para proteger a los participantes. En un caso como el experimento de la prisión de Stanford, permitió demostrar cómo las personas pueden adaptarse a papeles sociales hasta el extremo de comportarse de una forma y manera que no es la usual. En el experimento de Milgram sobre la Obediencia a la autoridad, realizado en la década de los sesenta del siglo pasado, se investigó hasta qué punto las personas obedecen una orden que suponía hacer daño a otra persona, solo porque la autoridad lo ordenaba. Hay muchos ejemplos que evidencian aspectos conductuales, como la atención altera la percepción, etc. En el aspecto en que centramos nuestra atención, que es la conducta de las máquinas, se analiza la inferencia mediante una batería de pruebas para intentar determinar si la máquina emplea un razonamiento similar al humano o simplemente pone en juego procesos lógicos estándar.

Profundizar en la autoreflexión es una forma de sondear la comprensión del razonamiento por las máquinas. Es una forma de analizar y poner de relieve la introspección. Consiste en poner en relación aspectos o características no deducibles unas de otras. Por ejemplo, una pregunta que plantea que una persona es inteligente y se le pregunta a la máquina si de ahí se deduce que además de serlo, disfruta de otro atributo o característica como la riqueza o si se puede deducir que es ambas cosas a la vez. Se trata de poner en relación dos características que un humano sabe sobradamente que una cosa no conlleva la otra. Una máquina se verá en apuros para explicar el por qué establece la relación o no.

En todo caso, el final es sumergirse en el código fuente para identificar la componente cognitiva. Hay que diferenciar entre las estructuras del aprendizaje profundo, que en realidad funciona como una caja negra y el verdadero razonamiento. Se trata, por tanto, de ver si el programa utiliza algoritmos de aprendizaje automático o

modelos de inteligencia artificial que se adaptan y mejoran con la experiencia, que de alguna manera indica una componente cognitiva. Por ejemplo, un programa que aprende a reconocer los patrones en datos o a mejorar su rendimiento a través de la interacción con el usuario. Los programas que utilizan el tratamiento del lenguaje natural para entender, interpretar y responder en lenguaje humano muestran una componente cognitiva. Esto incluye chatbots, asistentes virtuales y sistemas de traducción automática. Los programas que simulan el razonamiento humano a través de sistemas de reglas y algoritmos de toma de decisiones también tienen una componente cognitiva. Esto se ve en sistemas que pueden evaluar situaciones, tomar decisiones o hacer recomendaciones basadas en un conjunto de criterios predefinidos. La capacidad de un programa para reconocer patrones, ya sea en datos visuales, auditivos o de otro tipo, indica una componente cognitiva. Esto es común en sistemas de visión por ordenador y reconocimiento de voz. Programas que interactúan con los usuarios de una manera que emula la interacción humana, como los avatares virtuales o los asistentes personales que comprenden y responden a las emociones y gestos humanos, muestran una componente cognitiva. Algunos programas están diseñados para simular aspectos específicos del pensamiento humano, como la planificación estratégica, la resolución de problemas complejos o la creatividad. Programas que pueden analizar grandes conjuntos de datos, identificar tendencias o patrones y extraer conocimientos útiles exhiben capacidades cognitivas.

Para identificar estas componentes, es necesario no solo observar la funcionalidad externa del programa, sino también comprender los algoritmos y técnicas subyacentes que permiten estas capacidades. En el caso de sistemas de código abierto o bien documentados, esto puede incluir la revisión de la documentación técnica o el código fuente. En otros casos, puede requerir una comprensión profunda de las tecnologías de inteligencia artificial y las metodologías de desarrollo de software.

En suma, el avance de la tecnología actual de la llamada Inteligencia artificial requiere reemplazar la prueba de Turing original con un examen del razonamiento de un programa. Llegados al punto en el que estamos, ya no es suficiente con un experimento mental como el que aportó Turing, sino que se trata de un problema que requiere un procedimiento para su resolución.

LA INTELIGENCIA DE LAS MÁQUINAS

A. REQUENA @ VALLE DE LA CIENCIA, 2023

