

DESAFÍOS PENDIENTES
PARA LOS SISTEMAS INTELIGENTES

Francisco José Serón Arbeloa



STVDIVM
GENERALE
CAESARAV-
GVSTANAE
CIVITATIS

PRENSAS DE LA UNIVERSIDAD DE ZARAGOZA

**DESAFÍOS PENDIENTES
PARA LOS SISTEMAS INTELIGENTES**

DESAFÍOS PENDIENTES
PARA LOS SISTEMAS INTELIGENTES

Francisco José Serón Arbeloa

PRENSAS DE LA UNIVERSIDAD DE ZARAGOZA

STVDIVM
GENERALE
CAESARAV
GVSTANAE
CIVITATIS

COLECCIÓN PARANINFO

PRIMA LECTIO

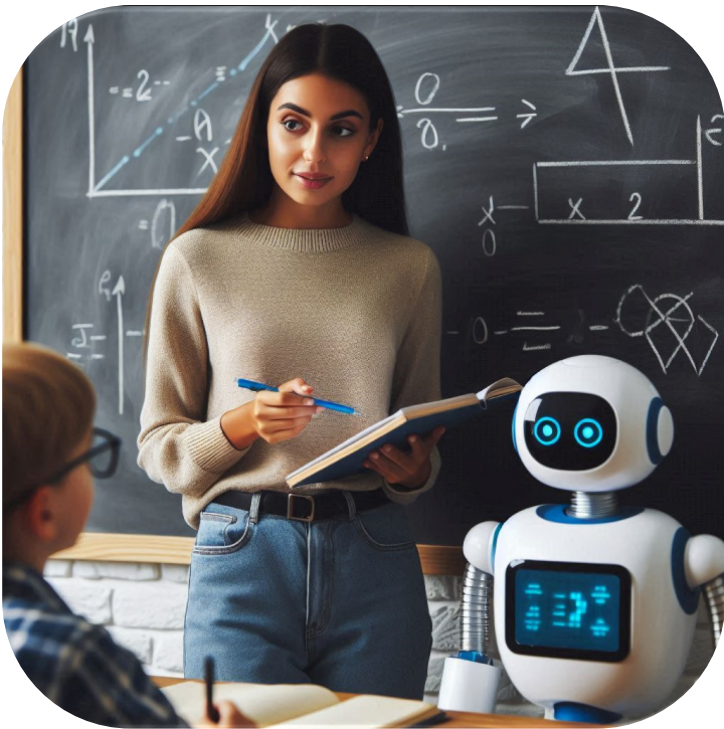
- © Francisco José Serón Arbeloa
© De la presente edición, Prensas de la Universidad de Zaragoza
(Vicerrectorado de Cultura y Proyección Social)
1.ª edición, 2024

Prensas de la Universidad de Zaragoza
Edificio de Ciencias Geológicas
c/ Pedro Cerbuna, 12 • 50009 Zaragoza, España
Tel.: 976 761 330
puz@unizar.es <http://puz.unizar.es>

Impreso en España
Imprime: Servicio de Publicaciones. Universidad de Zaragoza
ISBN: 978-84-1340-873-6
Depósito legal: Z 1326-2024

*Lo que un humano no entiende,
mal lo puede explicar a otro humano
y menos a una máquina.*

Frase de cualquier profesor con talento y de paso de los que construimos algoritmos



*Imagen creada mediante inteligencia artificial,
utilizando Microsoft Bing, Copilot Designer, con tecnología de Dall-E 3*

RESUMEN



*Imagen creada mediante inteligencia artificial,
utilizando Microsoft Bing, Copilot Designer, con tecnología de Dall-E 3*

El optimismo y el entusiasmo son cualidades esenciales que buscamos en el orador principal de cualquier conferencia relacionada con la revolución de los sistemas inteligentes.

Ahora todo el mundo habla de la inteligencia artificial y de la robótica, la gente siente curiosidad por saber cómo impactarán y seguirán impactando nuestras vidas y nuestro trabajo.

Personalmente estoy totalmente a favor de la innovación, cualquier tipo de tecnología transversal debemos entenderla, saber cuáles son sus límites y debemos proceder con cautela. En caso contrario repetiremos el mismo tipo de errores que se cometieron en la época de la Revolución Industrial o en la época de las redes sociales, cuyos desarrollos se realizaron sin restricciones.

Actualmente se espera que la inteligencia artificial y la robótica tengan mucho más impacto que las tecnologías innovadoras anteriores como, por ejemplo: la mecanización de la agricultura, las máquinas de vapor, la electricidad, la medicina moderna, las computadoras, Internet y los móviles. Pero debemos ser conscientes de que el potencial transformador de los sistemas inteligentes está limitado por sus problemas más difíciles que todavía están por resolver.

El objetivo final de este texto que tiene ante sí el lector, tan solo pretende analizar la siguiente frase «los seres humanos pretendemos crear sistemas artificiales con inteligencia parecida a la humana». En el capítulo final del texto, denominado «Inconclusión», daré mi opinión. Pero antes leamos sobre los desafíos que, en el momento actual, todavía tiene ante sí la construcción de los sistemas artificiales inteligentes.

Los sistemas basados en la inteligencia artificial y la robótica han experimentado un avance sin precedentes impulsados por la disponibilidad de grandes conjuntos

de datos, sensores, algoritmos sofisticados y mayor potencia de computación. Sin embargo, a pesar de estos avances, hoy en día estos sistemas aún no han logrado alcanzar muchas de las características de la inteligencia natural, como la que posee un ser humano.

Algunos de los principales problemas sin resolver que enfrenta el desarrollo de estos sistemas son:

La falta de sentido común

Los sistemas actuales pueden ser expertos en tareas específicas, pero carecen de la capacidad de comprender el mundo de la misma manera que lo hacemos los humanos. No tienen la capacidad de razonar utilizando el sentido común, comprender las normas sociales o interpretar las intenciones de los demás.

La ausencia de modelos del mundo

Los humanos poseemos modelos internos del mundo que nos permiten comprender cómo funciona, cómo interactúan los objetos y cómo se comportan las personas. Pero los sistemas actuales carecen de una comprensión adecuada del mundo que les rodea que les permita tomar decisiones informadas y actuar de manera coherente con cada contexto con el que se encuentran.

La interpretabilidad

A medida que los sistemas inteligentes se integran en nuestras vidas, surge una pregunta crucial: ¿Cómo podemos comprender los mecanismos internos de estos sistemas complejos y garantizar su uso responsable? ¿Cómo podemos desentrañar la caja negra de los algoritmos que tienen la capacidad de aprender para hacer que los sistemas inteligentes sean cada vez más transparentes e interpretables cuando toman decisiones?

El problema de la alineación

Por el momento los sistemas inteligentes son celebrados, en particular por el márketing de las grandes compañías, como si fueran por completo inofensivos. Pero nos enfrentamos al problema no resuelto de que estos sistemas deberían estar alineados con los valores y objetivos humanos para que puedan ser utilizados de manera segura y responsable. Sin embargo, es difícil garantizar que los sistemas inteligentes aprendan y tomen decisiones de manera consistente con nuestros valores.

Para alcanzar el objetivo planteado, se van a describir el conjunto de problemas que se acaban de citar, con la pretensión de avanzar en la adquisición de una comprensión más profunda de los desafíos que enfrentan la inteligencia artificial y la robótica, e intentar mostrar con ello, que en la actualidad los sistemas inteligentes, no tienen un conocimiento adecuado del mundo que los rodea y carecen del contexto real. La idea no es menospreciar a estas interesantes herramientas, ciertamente son útiles y podemos construir muchas aplicaciones a su alrededor, pero estos cuatro desafíos y otros que, por extensión no se detallan, son muy complejos y requieren de un esfuerzo multidisciplinar y holístico por parte de investigadores, científicos, tecnólogos y filósofos para encontrar soluciones responsables, seguras y beneficiosas, que nos acerquen un poco más hacia la construcción de sistemas que en su comportamiento se aproximen a la inteligencia humana. Es decir, naveguemos y veamos los desafíos que todavía nos depara el horizonte en este viaje que el ser humano empezó hace mucho, mucho tiempo.

SOBRE LA INTELIGENCIA



*Imagen creada mediante inteligencia artificial,
utilizando ideogram.ai*

ÍNDICE

Resumen	9
Sobre la inteligencia	13
¿Por qué estamos hoy aquí hablando de inteligencia artificial y de robótica?	17
Conocimiento explícito y conocimiento implícito	35
Algunos desafíos pendientes de alcanzar	41
Qué es el sentido común	47
¿Por qué el sentido común es tan importante para los sistemas de inteligentes?	53
Algunos proyectos famosos relacionados con la obtención del sentido común en máquinas	65
Modelos del mundo	75
Algunos proyectos relacionados con la obtención de modelos del mundo en máquinas	83
Interpretabilidad y alineación	85
Conclusiones	93
Inconclusión	99
Créditos	103
Agradecimientos	113

*Este libro se terminó de imprimir
en los talleres del Servicio de Publicaciones
de la Universidad de Zaragoza
el 31 de julio de 2024*

COLECCIÓN PARANINFO
PRIMA LECTIO



STVDIVM
GENERALE
CAESARAV-
GVSTANAE
CIVITATIS



Universidad Zaragoza