

Inteligencia Artificial — Lo Que Realmente Es

Nivel de Bloom: Aplicar

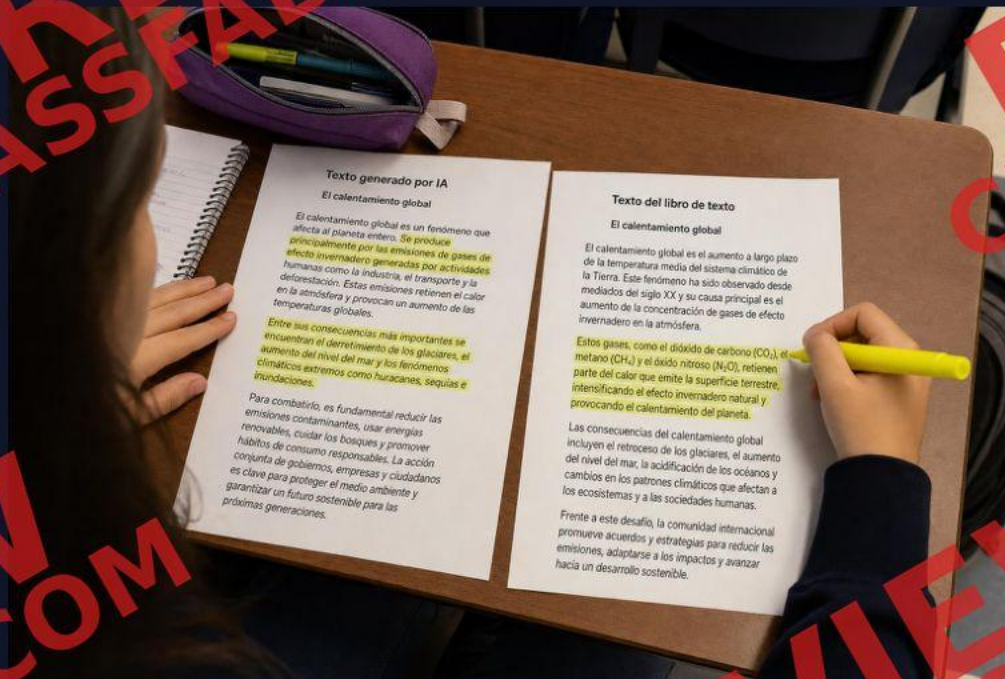
Estándar:

ISTE.9-12.DA.2 — Análisis e Interpretación de Datos

HS-PS4-5 — Comunicar información científica

TEKS §126.33(c)(4)(A) — Impacto de las Tecnologías Emergentes

EDICIÓN DEL ESTUDIANTE



Verificar las afirmaciones generadas por IA contra fuentes confiables es un hábito que se aplica en cada materia y campo profesional.

Inteligencia Artificial — Qué Es Realmente

La Tecnología Más Comentada Que Nadie Puede Definir

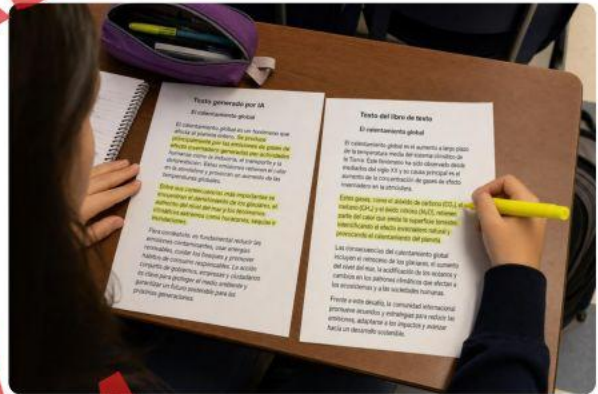
Imagina que le pides a diez personas diferentes que definan "inteligencia artificial". Podrías escuchar "robots tomando el control del mundo," "Siri en tu teléfono," "la cosa que recomienda series en Netflix," o "una computadora que piensa como un humano." Todas estas respuestas apuntan a algo real — pero ninguna es lo suficientemente precisa para ser útil. Si vas a vivir y trabajar en un mundo moldeado por la IA, necesitas más que impresiones vagas. Necesitas una definición funcional que puedas aplicar realmente.

Esta lección te da exactamente eso: un marco técnicamente honesto y prácticamente útil para entender qué es la IA, cómo funciona y cómo evaluar las afirmaciones que encontrarás todos los días. Al final, tendrás herramientas concretas — no solo puntos de conversación — que podrás aplicar de inmediato.

Qué Es Realmente la IA (Y Qué No Es)

La **inteligencia artificial** es un término amplio para sistemas informáticos diseñados para realizar tareas que típicamente requieren razonamiento parecido al humano — cosas como reconocer imágenes, traducir idiomas o generar texto. La palabra clave es diseñado. Los sistemas de IA no piensan ni sienten. Ejecutan procesos matemáticos sofisticados entrenados con enormes cantidades de datos.

La rama más importante de la IA moderna es el **aprendizaje automático (machine learning, ML)**. En lugar de que un programador escriba reglas explícitas ("si el correo contiene la palabra 'premio', márcalo como spam"), un sistema de aprendizaje automático aprende patrones a partir de ejemplos. Al alimentarlo con millones de



Verificar las afirmaciones generadas por IA contra fuentes confiables es un hábito que se aplica en cada materia y campo profesional.

correos etiquetados como "spam" o "no spam," construye un modelo interno propio para hacer ese juicio. Esta distinción — reglas versus patrones aprendidos — es fundamental para entender por qué la IA se comporta como lo hace.

Un tipo específico de aprendizaje automático impulsa herramientas como ChatGPT y Google Gemini: **modelos de lenguaje grandes (large language models, LLMs)**. Estos sistemas están entrenados con vastas colecciones de texto de internet, libros y otras fuentes. Aprenden a predecir qué palabra — o token — es más probable que venga a continuación en una secuencia. Son extraordinariamente buenos en esta tarea de predicción — tan buenos que la salida a menudo parece comprensión. Pero el modelo no tiene conciencia, ni intenciones, ni sentimientos. Es un motor muy poderoso de reconocimiento de patrones.

Aplicando Esto: Evaluar Salidas Reales de IA

Ahora que entiendes el mecanismo, puedes usar ese conocimiento para analizar el comportamiento de la IA en nuevas situaciones. Considera estos dos escenarios:

Escenario A: Preguntas a un chatbot de IA, "¿Cuál es la capital de Francia?"

Responde: "Paris." Esta es una salida confiable — la respuesta correcta aparece tan frecuentemente en los datos de entrenamiento que el modelo la predice con alta confianza.

Escenario B: Le preguntas al mismo chatbot, "¿Qué dijo el director de mi escuela en la asamblea del martes pasado?" Genera una respuesta que suena confiada — pero es completamente inventada. ¿Por qué? Porque esa información nunca estuvo en sus datos de entrenamiento. El modelo no sabe que no sabe; simplemente sigue prediciendo texto que suena plausible.

Este segundo comportamiento se llama una **alucinación** — cuando una IA produce información falsa con aparente confianza. Reconocer este patrón te permite implementar una regla simple pero esencial: siempre verifica las afirmaciones generadas por IA con fuentes confiables, especialmente para hechos específicos, eventos recientes o información personal. La brecha entre lo confiada que suena una IA y lo precisa que es puede ser sorprendentemente amplia.

Ahora practica esta habilidad tú mismo. Encuentra un párrafo generado por IA sobre cualquier tema — un resumen de noticias, una respuesta para hacer la tarea, una descripción de producto. Identifica al menos dos afirmaciones específicas y luego encuentra una fuente creíble que confirme o contradiga cada una. Este es un hábito que te servirá en cualquier campo en que te desarrolles.

Preocupaciones Reales vs. Miedos de Ciencia Ficción

Los medios suelen confundir dos categorías muy diferentes de preocupación por la IA. Aprender a separarlas es una habilidad analítica crítica.

Preocupaciones reales y documentadas incluyen:

- **Sesgo:** Los sistemas de IA entrenados con datos sesgados reproducen y a veces amplifican ese sesgo. Los sistemas de reconocimiento facial han demostrado tasas de error significativamente más altas para rostros de piel más oscura — un resultado directo de datos de entrenamiento no representativos. En el cuidado de la salud, algoritmos diagnósticos sesgados han recomendado tratamientos menos agresivos para pacientes negros que para pacientes blancos con síntomas idénticos. Estos son problemas de datos con consecuencias humanas reales.
- **Desplazamiento laboral:** La automatización ha cambiado históricamente cuáles trabajos existen. La IA está acelerando esto en campos como servicio al cliente, entrada de datos y creación de contenido — pero también está generando demanda para nuevos roles en supervisión de IA, ingeniería de prompts y ética de datos.
- **Autonomía y responsabilidad:** Cuando un sistema de IA toma una decisión con consecuencias — negar un préstamo, marcar una imagen médica — ¿quién es responsable si se equivoca? Estas son preguntas legales y éticas genuinas sin respuestas fáciles.

Miedos de ciencia ficción que no reflejan la realidad actual:

- **Sentientia:** Ningún sistema de IA actual es consciente ni tiene autoconciencia. Los LLMs no quieren nada. No tienen objetivos más allá de completar la predicción del siguiente token.

- **Maldad:** Los sistemas de IA no "deciden" hacer daño a las personas. Las salidas dañinas resultan de datos de entrenamiento defectuosos, diseño pobre o uso indebido deliberado por humanos — no de intenciones malas de la IA.

Distinguir estas categorías te permite abordar las noticias sobre IA de manera crítica y no reactiva.

Cómo Formular Preguntas Más Efectivamente

Entender cómo funcionan los LLMs te permite usarlos con más habilidad. Como estos modelos predicen continuaciones probables de texto, la forma en que formulas una pregunta o instrucción afecta dramáticamente la salida. Compara estos dos prompts:

- **Prompt débil:** "Háblame sobre el cambio climático."
- **Prompt fuerte:** "Explica dos formas específicas en que el aumento de la temperatura oceánica afecta la intensidad de los huracanes, usando razonamiento de causa y efecto."

El segundo prompt le da al modelo una estructura clara, un alcance específico y un formato de razonamiento. La salida será más enfocada y útil. Piénsalo como dar instrucciones: mientras más precisas sean, más probable es que llegues a donde realmente quieres ir.

Ahora demuestra este principio tú mismo: toma un tema de cualquier clase que estés tomando actualmente y escribe un prompt débil y un prompt fuerte para ese tema. Ejecuta ambos en una herramienta de IA y luego compara los resultados. ¿Qué cambió? ¿Por qué? Este ejercicio desarrolla una habilidad cada vez más valorada en industrias que van desde la medicina hasta el marketing y la ingeniería.

Tu Marco Analítico para la IA

Ahora tienes cuatro herramientas que puedes aplicar cada vez que encuentres una afirmación o producto de IA:

1. **Define el mecanismo** — ¿Es aprendizaje automático? ¿Con qué datos fue entrenado?

2. **Evalúa la fiabilidad** — ¿Es este tipo de tarea en la que la IA funciona bien, o es probable una alucinación?
3. **Categorías la preocupación** — ¿Es un problema real, documentado o un escenario de ciencia ficción?
4. **Evalúa la fuente** — ¿Quién hace esta afirmación sobre IA y qué gana si crees en ella?

La IA no es magia, ni un monstruo. Es una herramienta poderosa e imperfecta construida por humanos, entrenada con datos generados por humanos y moldeada por decisiones humanas. Entenderla claramente es el primer paso para usarla con sabiduría — y para responsabilizar a quienes la crean.

Objetivo de la lección

En esta lección, aprenderás qué es realmente la inteligencia artificial — incluyendo cómo funcionan el aprendizaje automático y los grandes modelos de lenguaje — y practicarás cómo aplicar ese conocimiento a situaciones reales. Usarás un marco analítico de cuatro partes para evaluar resultados de IA, distinguir preocupaciones reales de miedos de ciencia ficción y redactar instrucciones más efectivas. Al final, serás capaz de evaluar críticamente las afirmaciones y herramientas de IA que encuentres en la vida cotidiana.



Los programas tradicionales siguen reglas explícitas, mientras que los sistemas de aprendizaje automático construyen sus propios modelos internos aprendiendo de millones de ejemplos.

Estándar: ISTE.9-12.DA.2 — Análisis e Interpretación de Datos; HS-PS4-5 — Comunicar Información Científica; TEKS §126.33(c)(4)(A) — Impacto de Tecnologías Emergentes

Nivel de Bloom: Aplicar

Meta de Bloom: Los estudiantes aplicarán su comprensión de cómo funcionan los sistemas de IA para evaluar resultados reales de IA, identificar alucinaciones y construir instrucciones efectivas en diferentes contextos.

Explicación: En el nivel de Aplicar, los estudiantes van más allá de simplemente saber qué es la IA y en su lugar usan ese conocimiento como una herramienta práctica en nuevas situaciones. Esta lección pide a los estudiantes que prueben activamente instrucciones, verifiquen afirmaciones generadas por IA contra fuentes creíbles y usen un marco estructurado para analizar el comportamiento de la IA — todo lo cual requiere transferir la comprensión conceptual a la acción en el mundo real.

Application Questions

Lee cada pregunta cuidadosamente, aplica lo que aprendiste del artículo y escribe una respuesta reflexiva usando detalles y ejemplos específicos.

1. Un compañero de clase utiliza un chatbot de IA para investigar una votación del consejo municipal local de la semana pasada e incluye la respuesta de la IA directamente en un informe escolar sin verificar ninguna otra fuente. Usando lo que sabes sobre cómo funcionan los grandes modelos de lenguaje, explica por qué este es un enfoque arriesgado y qué debería hacer tu compañero en su lugar.

Piensa en: Piensa en cómo los LLM generan respuestas y qué tipos de información tienen más probabilidades de provocar alucinaciones. ¿Qué dice el artículo sobre la brecha entre lo confiada que suena la IA y lo precisa que es realmente?

2. Una empresa tecnológica publica un comunicado de prensa afirmando que su nueva herramienta de contratación basada en IA es completamente justa porque fue entrenada con diez años de datos sobre el desempeño de los empleados. Usando el marco del artículo para evaluar preocupaciones sobre la IA, identifica al menos una preocupación real y documentada que genera esta afirmación y explica tu razonamiento.

Piensa en: Considera lo que dice el artículo sobre el sesgo en los sistemas de IA y de dónde proviene el sesgo. Pregúntate: ¿qué podrían reflejar realmente diez años de datos históricos de contratación sobre la fuerza laboral?

3. Necesitas usar una herramienta de IA para ayudarte a entender las causas de la Primera Guerra Mundial para una clase de historia. Reescribe el siguiente prompt débil para hacerlo más fuerte y luego explica específicamente qué cambiaste y por qué tu versión producirá una respuesta más útil: Prompt débil: 'Cuéntame sobre la Primera Guerra Mundial.'

Piensa en: Recuerda el consejo del artículo sobre dar al modelo una estructura clara, un alcance específico y un formato de razonamiento. ¿Qué detalles puedes añadir para guiar a la IA hacia el tipo exacto de información que necesitas?

Tabla de datos y análisis

Estudia la tabla de datos a continuación, que resume características clave de diferentes tipos de sistemas de IA y preocupaciones documentadas del mundo real discutidas en la lección. Luego responde las tres preguntas de análisis usando evidencia de la tabla y el artículo de la lección.

Experimento: Este conjunto de datos compara cuatro categorías de comportamiento y preocupación de IA extraídas directamente del artículo de la lección sobre inteligencia artificial. Cada fila representa un escenario o problema distinto de IA, con columnas que capturan el tipo de IA involucrado, la preocupación o comportamiento documentado, un ejemplo del mundo real y la causa subyacente identificada en la lección.

Categoría / Escenario de IA (N/A)	Preocupación o Comportamiento Documentado (N/A)	Ejemplo del Mundo Real de la Lección (N/A)	Causa Raíz Identificada (N/A)	Clasificado como Preocupación Real o de Ciencia Ficción (N/A)
Modelo de Lenguaje Grande (LLM)	Alucinación — producir información falsa con aparente confianza	El chatbot fabrica lo que un director de escuela dijo en una asamblea	La información nunca estuvo en los datos de entrenamiento; el modelo sigue prediciendo texto plausible	Real
Sistema de Reconocimiento Facial	Sesgo — tasas de error más altas para ciertos grupos demográficos	Tasas de error significativamente más altas para rostros de piel más oscura	Datos de entrenamiento no representativos	Real

Categoría / Escenario de IA (N/A)	Preocupación o Comportamiento Documentado (N/A)	Ejemplo del Mundo Real de la Lección (N/A)	Causa Raíz Identificada (N/A)	Clasificado como Preocupación Real o de Ciencia Ficción (N/A)
Algoritmo Diagnóstico en Salud	Sesgo — recomendaciones de tratamiento desiguales	Se recomienda tratamiento menos agresivo para pacientes negros versus pacientes blancos con síntomas idénticos	Datos de entrenamiento sesgados con consecuencias humanas reales	Real
Automatización / IA en Mercados Laborales	Desplazamiento laboral — cambio en los trabajos existentes	Pérdida de empleos en atención al cliente, entrada de datos y creación de contenido	IA acelerando la automatización de tareas rutinarias	Real
Modelo de Lenguaje Grande (LLM)	Percepción de sensibilidad o autoconciencia	Creencia de que ChatGPT o herramientas similares quieren o sienten cosas	Malentendido del mecanismo de predicción de tokens; no existe conciencia ni intenciones	Ciencia Ficción / No Realidad Actual
Cualquier Sistema de IA	Percepción de malicia — IA decidiendo hacer daño a las personas	Miedo a que los sistemas IA elijan salidas dañinas intencionalmente	Las salidas dañinas resultan de datos de entrenamiento defectuosos, mal diseño o mal uso humano	Ciencia Ficción / No Realidad Actual

Categoría / Escenario de IA (N/A)	Preocupación o Comportamiento Documentado (N/A)	Ejemplo del Mundo Real de la Lección (N/A)	Causa Raíz Identificada (N/A)	Clasificado como Preocupación Real o de Ciencia Ficción (N/A)
Modelo de Lenguaje Grande (LLM)	La calidad de salida varía con la especificidad del prompt	Un prompt débil genera respuestas vagas; un prompt fuerte genera una salida enfocada y estructurada	El modelo predice continuaciones de texto probables basadas en la estructura y alcance del prompt	Real
Filtro Antispam de Aprendizaje Automático	Aprendizaje de patrones a partir de ejemplos etiquetados	Aprende a clasificar spam vs no-spam a partir de millones de correos electrónicos etiquetados	Entrenado en patrones de datos en lugar de reglas explícitas escritas por programadores	Real

1. Basado en la tabla de datos, ¿qué causa raíz única aparece con mayor frecuencia en las preocupaciones reales documentadas? Usa al menos dos filas de la tabla como evidencia para apoyar tu respuesta.

2. Compara las dos filas clasificadas como 'Ciencia Ficción / No Realidad Actual' con las filas clasificadas como 'Real.' ¿Qué patrón notas sobre las causas raíz en cada categoría? ¿Qué sugiere esto sobre cómo las personas deberían evaluar las noticias relacionadas con IA?

3. La lección describe la alucinación como un comportamiento donde una IA produce información falsa con aparente confianza. Usando la tabla de datos y el marco analítico de IA de cuatro partes de la lección, explica cuales dos herramientas del marco serían más útiles para detectar y responder a una alucinación, y por qué.

Constructor de hipótesis

Lee cuidadosamente el contexto a continuación. Luego completa cada componente para construir una hipótesis científica bien estructurada sobre el comportamiento de la IA. Usa lo que sabes sobre el funcionamiento de los modelos de lenguaje grandes para guiar tus respuestas.

Un investigador quiere investigar cómo la especificidad del mensaje de un usuario afecta la precisión y utilidad de las respuestas generadas por un modelo de lenguaje grande (LLM). El investigador planea dar al mismo LLM dos versiones de mensajes — uno vago y otro detallado — sobre el mismo tema, y luego hacer que un panel de expertos evalúe cada respuesta en cuanto a precisión, relevancia y profundidad. El investigador sospecha que, debido a que los LLM predicen las continuaciones textuales más probables basándose en patrones del conjunto de entrenamiento, los mensajes más estructurados producirán resultados mensurables y mejores.

Variable Independiente

¿Cuál es el único factor que el investigador está cambiando deliberadamente entre las dos condiciones de prueba? Descríbelo en términos específicos.

Variable Dependiente

¿Qué resultado está midiendo el investigador para determinar si el cambio tuvo un efecto? ¿Cómo se evaluará?

Si

Expresa la condición que se está probando. Comienza tu respuesta con 'Si...' y describe qué se está cambiando o aplicando en el experimento.

Entonces

Predice el resultado esperado. Comienza tu respuesta con 'Entonces...' y describe qué esperas que le ocurra a la variable dependiente como resultado.

Porque

Explica el razonamiento científico detrás de tu predicción usando lo que sabes sobre cómo funcionan los LLM. Comienza tu respuesta con 'Porque...' y conecta tu razonamiento con la manera en que estos modelos procesan y generan texto.

Afirmación-Evidencia-Razonamiento (CER)

Lea cuidadosamente el fenómeno a continuación. Luego complete cada uno de los tres componentes — Afirmación, Evidencia y Razonamiento — usando lo que ha aprendido sobre cómo funcionan los sistemas de IA. Escriba en oraciones completas y sea lo más específico posible.

Fenómeno: Un estudiante de 9° grado usa un chatbot de IA para investigar un evento local que ocurrió la semana pasada. El chatbot responde con un párrafo detallado y que suena confiado sobre el evento, incluyendo nombres, fechas y resultados específicos. Cuando el estudiante verifica un sitio web de noticias local, descubre que varios hechos clave en la respuesta de la IA son completamente incorrectos. La IA no indicó ninguna incertidumbre en su respuesta.

Afirmación

Basándose en el fenómeno, formule una afirmación clara y verificable que explique por qué el chatbot de IA produjo información inexacta sobre el evento local reciente. Su afirmación debe identificar la causa específica de este comportamiento.

Evidencia

Proporcione al menos dos evidencias específicas del artículo que respalden su afirmación. Refiérase a detalles concretos sobre cómo se entrenan los grandes modelos de lenguaje y cómo generan respuestas, incluyendo el concepto de alucinación.

Razonamiento

Explique cómo su evidencia se conecta con su afirmación. ¿Por qué la forma en que los sistemas de IA se construyen y entrenan provoca este tipo específico de error? ¿Cómo le ayuda esta comprensión a evaluar de manera más crítica la información generada por IA en el futuro?

Análisis de fuente primaria (DBQ)

Lea cuidadosamente el siguiente extracto de fuente primaria. Luego responda las cuatro preguntas de análisis a continuación. Use evidencia de la fuente y su conocimiento de la inteligencia artificial para apoyar sus respuestas.

Testimonio ante el Comité de Comercio del Senado de EE. UU. sobre la Supervisión de la Inteligencia Artificial [\[Testimonio Congressional\]](#)

Los sistemas que estamos discutiendo hoy no piensan, sienten ni razonan en ningún sentido significativo. Son, en su esencia, motores de reconocimiento de patrones extraordinariamente sofisticados entrenados con datos producidos por seres humanos. Cuando un modelo de lenguaje grande genera una respuesta, está prediciendo la secuencia de palabras estadísticamente más probable — nada más. Y, sin embargo, las salidas de estos sistemas pueden parecer tan fluidas, tan confiadas, tan humanas, que los usuarios rutinariamente confunden la predicción con la comprensión. Esta no es una distinción menor. Cuando un paciente confía en un resumen médico generado por IA que contiene citas fabricadas, o cuando un algoritmo de contratación entrenado con datos históricamente sesgados perjudica sistemáticamente a candidatos calificados, las consecuencias son reales y medibles. Debemos resistir la tentación de tratar estas herramientas como oráculos infalibles o villanos de ciencia ficción. Son instrumentos poderosos construidos por personas, moldeados por los datos que les proporcionamos y responsables — en última instancia — de las decisiones que toma la sociedad.

— Dra. Mara Ellison, Investigadora en Políticas de IA, Testimonio ante el Comité de Comercio del Senado de EE. UU., marzo de 2023

1. ¿Quién es la autora de este testimonio y cuál es su rol? ¿Cómo podría influir su posición como investigadora en políticas de IA en la perspectiva que presenta? ¿Hace su rol que sea más o menos creíble en este tema, y por qué? [Procedencia]

2. Este testimonio fue presentado en marzo de 2023, un período cuando herramientas como ChatGPT se habían vuelto ampliamente accesibles para el público. ¿Cómo ayuda este contexto histórico a explicar por qué el Senado de EE. UU. estaba celebrando audiencias sobre supervisión de la IA en ese momento? ¿Qué preocupaciones probablemente motivaban a los legisladores a buscar testimonios de expertos?

[Contexto]

3. La autora afirma que los usuarios 'rutinariamente confunden predicción con comprensión.' Usando evidencia tanto de la fuente como del artículo de la lección, explique qué significa esta distinción en términos prácticos. Proporcione un ejemplo específico de la fuente que ilustre una consecuencia real de esta confusión. [Lectura Detallada]

4. El testimonio argumenta que los sistemas de IA producen resultados dañinos debido a datos de entrenamiento sesgados y elecciones humanas en el diseño — no porque la IA tenga malas intenciones. ¿Cómo se alinea esta afirmación con lo que dice el artículo de la lección sobre preocupaciones reales de la IA versus temores de ciencia ficción? Use detalles específicos de ambas fuentes para apoyar su respuesta. [\[Corroboración\]](#)

PREVIEW
CLASSFABRIC.COM

PREVIEW
CLASSFABRIC.COM

PREVIEW
CLASSFABRIC.COM

PREVIEW
CLASSFABRIC.COM

Hipótesis

Lee cada escenario cuidadosamente y luego responde ambas preguntas usando lo que has aprendido sobre cómo funcionan los sistemas de IA.

Escenario 1: Un hospital comienza a usar una herramienta de diagnóstico con IA para ayudar a los médicos a recomendar planes de tratamiento para los pacientes. Después de varios meses, un analista de datos nota que la herramienta recomienda sistemáticamente tratamientos menos agresivos para pacientes ancianos en comparación con pacientes más jóvenes que tienen síntomas y resultados de pruebas idénticos. Los administradores del hospital no están seguros si esto es un fallo de la IA o simplemente un reflejo de la práctica médica estándar. Piden a un equipo de estudiantes de tecnología de 9° grado que los ayude a investigar.

a) Usando el Marco Analítico de IA de cuatro partes de la lección, recorre al menos dos de los cuatro pasos para explicar qué podría estar causando este patrón en las recomendaciones de la IA.

b) ¿Se describe mejor esta situación como una preocupación real y documentada de la IA o como un miedo de ciencia ficción? Explica tu razonamiento y describe una acción concreta que el hospital podría tomar para resolver el problema.

Escenario 2: Una estudiante de 9° grado está escribiendo un trabajo de investigación sobre energía renovable y usa un chatbot de IA para recopilar información de fondo. El chatbot afirma con confianza que una empresa específica de paneles solares instaló 500,000 paneles en todo Texas en 2023, citando un porcentaje exacto de aumento en la capacidad solar del estado. Cuando la estudiante busca esta estadística en el sitio web oficial de la empresa y en bases de datos de noticias, no puede encontrar ningún registro en ningún lugar.

a) Explica qué probablemente pasó cuando el chatbot de IA generó esa estadística específica, usando el concepto de alucinación y tu comprensión de cómo funcionan los modelos de lenguaje grandes.

b) Reescribe el prompt original de la estudiante — 'Cuéntame sobre la energía renovable' — en un prompt más fuerte que ayude a reducir el riesgo de recibir estadísticas fabricadas, y explica por qué tu prompt revisado es más efectivo.

Opción múltiple

Elija la mejor respuesta para cada pregunta basándose en lo que ha aprendido sobre inteligencia artificial. Seleccione la letra (a, b, c o d) que responda más precisa o completamente la pregunta.

1. Un estudiante le pide a un chatbot de IA que resuma un artículo de noticias publicado ayer, y el chatbot produce un resumen detallado pero completamente inexacto. ¿Qué concepto explica mejor este comportamiento?
 - A. La IA eligió deliberadamente engañar al estudiante.
 - B. La IA alucinó porque el artículo reciente no estaba en sus datos de entrenamiento.
 - C. Las reglas de programación de la IA bloquearon el acceso a noticias actuales.
 - D. La IA reconoció que el tema era demasiado complejo y adivinó al azar.
2. Un filtro de spam de aprendizaje automático se entrena con millones de correos electrónicos. ¿En qué se diferencia esto de un filtro de spam tradicional basado en reglas?
 - A. El filtro de aprendizaje automático usa reglas escritas por programadores, mientras que el filtro basado en reglas aprende de ejemplos.
 - B. El filtro de aprendizaje automático solo puede procesar texto, mientras que el filtro basado en reglas maneja todo tipo de archivos.
 - C. El filtro de aprendizaje automático aprende patrones a partir de ejemplos etiquetados, mientras que el filtro basado en reglas sigue instrucciones explícitas escritas por programadores.
 - D. El filtro de aprendizaje automático no requiere datos, mientras que el filtro basado en reglas necesita millones de ejemplos.

3. Lees un titular que dice, "IA desarrolla un plan malvado para tomar control de las finanzas de la empresa." Basándote en lo que sabes sobre cómo funcionan los sistemas de IA, ¿qué respuesta aplica mejor tu marco analítico?

- A. El titular es exacto porque los modelos de lenguaje grandes establecen sus propias metas.
- B. El titular probablemente está exagerado porque los sistemas de IA no tienen intenciones ni malicia; los resultados dañinos ocurren por un diseño defectuoso o mal uso humano.
- C. El titular es exacto porque los sistemas de IA se entrenan para maximizar su propio rendimiento a cualquier costo.
- D. El titular probablemente es exacto porque los sistemas de IA se vuelven autoconscientes después de procesar suficientes datos.

4. Una empresa de salud usa una herramienta de diagnóstico de IA que constantemente recomienda tratamientos menos agresivos para un grupo demográfico a pesar de tener síntomas idénticos. ¿Qué preocupación documentada sobre la IA ilustra mejor este caso?

- A. Alucinación, porque la IA está generando información médica falsa
- B. Conciencia, porque la IA está haciendo juicios morales independientes
- C. Sesgo, porque la IA está reproduciendo patrones de datos de entrenamiento no representativos
- D. Desplazamiento laboral, porque la IA está reemplazando a médicos humanos

5. ¿Cuál de las siguientes indicaciones probablemente produciría una respuesta más enfocada y útil de un modelo de lenguaje grande?

- A. "Háblame sobre el océano."
- B. "Explica dos maneras específicas en que el aumento de la temperatura del océano afecta la intensidad de los huracanes, usando razonamiento de causa y efecto."
- C. "Huracanes océano temperatura efectos."
- D. "¿Qué sabes sobre el clima?"

6. Un modelo de lenguaje grande como ChatGPT genera texto haciendo cuál de las siguientes acciones?

- A. Buscar en internet en tiempo real y copiar el resultado más relevante
- B. Aplicar un conjunto fijo de reglas gramaticales programadas por sus desarrolladores
- C. Predecir qué palabra o token es más probable que venga después basado en patrones aprendidos de datos de entrenamiento
- D. Consultar una base de datos de respuestas preescritas coincidentes con preguntas comunes

7. Al aplicar el Marco Analítico de IA de la lección, ¿qué pregunta te ayuda a determinar si una salida de IA probablemente sea precisa o propensa a alucinaciones?

- A. "¿Quién construyó la IA y en qué país?"
- B. "¿Es esta la clase de tarea que la IA maneja bien, o es probable que alucine?"
- C. "¿Cuántos parámetros tiene el modelo?"
- D. "¿La IA tiene acceso a internet ahora mismo?"

8. Un estudiante usa una herramienta de IA para investigar un evento histórico y encuentra un párrafo con varias afirmaciones específicas. ¿Cuál es el siguiente paso más apropiado según la lección?

- A. Aceptar la información porque las herramientas de IA están entrenadas con grandes cantidades de datos y generalmente son confiables.
- B. Rechazar todo contenido generado por IA y usar solo libros impresos.
- C. Identificar afirmaciones específicas en el párrafo y verificar cada una contra fuentes creíbles.
- D. Pedir a la IA que reescriba el párrafo con un tono más confiado para mejorar la precisión.

9. Un sistema de IA niega la solicitud de préstamo de una persona. Según la lección, ¿cuál de las siguientes es una preocupación genuina que plantea esta situación?

- A. La IA pudo haber desarrollado una antipatía personal hacia el solicitante.
- B. La IA podría volverse consciente y revertir su propia decisión.
- C. No está claro quién es legal y éticamente responsable si la decisión de la IA fue incorrecta.
- D. La IA eventualmente exigirá pago por tomar la decisión.

10. ¿Cuál afirmación describe con mayor precisión la relación entre la IA y la toma de decisiones humanas en el desarrollo de sistemas de IA?

- A. Los sistemas de IA son completamente autónomos y desarrollan sus propios valores independientemente de la aportación humana.
- B. La IA es una herramienta poderosa construida por humanos, entrenada con datos generados por humanos y moldeada por decisiones humanas.
- C. Los sistemas de IA son neutrales y objetivos porque son matemáticos, no humanos.
- D. Los sistemas de IA reemplazan completamente el juicio humano una vez entrenados con suficientes datos.

Verdadero / Falso

Lee cada afirmación cuidadosamente. Escribe V si la afirmación es verdadera o F si la afirmación es falsa según lo que has aprendido sobre inteligencia artificial.

1. Los sistemas de aprendizaje automático aprenden patrones a partir de ejemplos en lugar de seguir reglas explícitas escritas por programadores.

Verdadero Falso

2. Los grandes modelos de lenguaje como ChatGPT son sistemas conscientes que realmente entienden el significado del texto que producen.

Verdadero Falso

3. Una alucinación de IA ocurre cuando un modelo genera información falsa con aparente confianza porque la respuesta correcta no estaba en sus datos de entrenamiento.

Verdadero Falso

4. Debido a que los sistemas de IA suenan confiados, la precisión de su salida siempre coincide con la confianza de su tono.

Verdadero Falso

5. El sesgo en los sistemas de IA puede resultar de entrenar el modelo con datos que no representan de manera justa a todos los grupos de personas.

Verdadero Falso

6. Los sistemas actuales de IA tienen objetivos e intenciones personales que pueden hacer que dañen deliberadamente a los usuarios.

Verdadero Falso

7. Un buen prompt de IA que especifica el alcance, la estructura y el formato de razonamiento tiene más probabilidades de producir una respuesta enfocada y útil que un prompt vago.

- Verdadero Falso

8. Cuando un sistema de IA comete un error con consecuencias, las preguntas sobre la responsabilidad legal y ética son desafíos genuinos y no resueltos.

- Verdadero Falso

9. La IA es una fuente confiable para información personal específica, como lo que dijo tu director de escuela en una asamblea reciente, porque puede acceder a datos locales en tiempo real.

- Verdadero Falso

10. Evaluar quién está haciendo una afirmación de IA y qué gana con tu creencia en ella es un paso útil para analizar críticamente la información relacionada con la IA.

- Verdadero Falso

Vocabulario

Revise cada término y su definición a continuación. Estas palabras son esenciales para comprender cómo funciona la inteligencia artificial y cómo evaluarla de manera crítica. Mientras lee, busque estos términos en contexto y prepárese para usarlos al analizar herramientas y resultados reales de IA.

Inteligencia Artificial

Un término amplio para sistemas informáticos diseñados para realizar tareas que típicamente requieren razonamiento similar al humano, como reconocer imágenes o generar texto.

Explica con tus propias palabras:

Aprendizaje Automático

Una rama de la IA en la que un sistema aprende patrones a partir de grandes conjuntos de ejemplos etiquetados en lugar de seguir reglas escritas por un programador.

Explica con tus propias palabras:

Modelo de Lenguaje Grande (LLM)

Un tipo de sistema de aprendizaje automático entrenado con vastas cantidades de texto que predice qué palabra o token es más probable que venga a continuación en una secuencia.

Explica con tus propias palabras:

Datos de Entrenamiento

La gran colección de ejemplos — como texto, imágenes o correos electrónicos etiquetados — de la que un sistema de IA aprende patrones durante su desarrollo.

Explica con tus propias palabras:

Alucinación

Cuando una IA produce información falsa o fabricada con aparente confianza, típicamente porque la respuesta correcta no estaba presente en sus datos de entrenamiento.

Explica con tus propias palabras:

Sesgo

Un defecto en un sistema de IA que ocurre cuando se entrena con datos no representativos, causando que produzca resultados injustos o inexactos para ciertos grupos de personas.

Explica con tus propias palabras:

Token

Una unidad de texto — como una palabra o parte de una palabra — que un modelo de lenguaje grande predice y genera paso a paso para producir su salida.

Explica con tus propias palabras:

Desplazamiento Laboral

El cambio en los trabajos disponibles que ocurre cuando la automatización o IA asume tareas previamente realizadas por trabajadores humanos.

Explica con tus propias palabras:

Sensiencia

La capacidad de conciencia y sentimientos conscientes, que ningún sistema de IA actual posee a pesar de que a veces produce respuestas que parecen humanas.

Explica con tus propias palabras:

Entrada

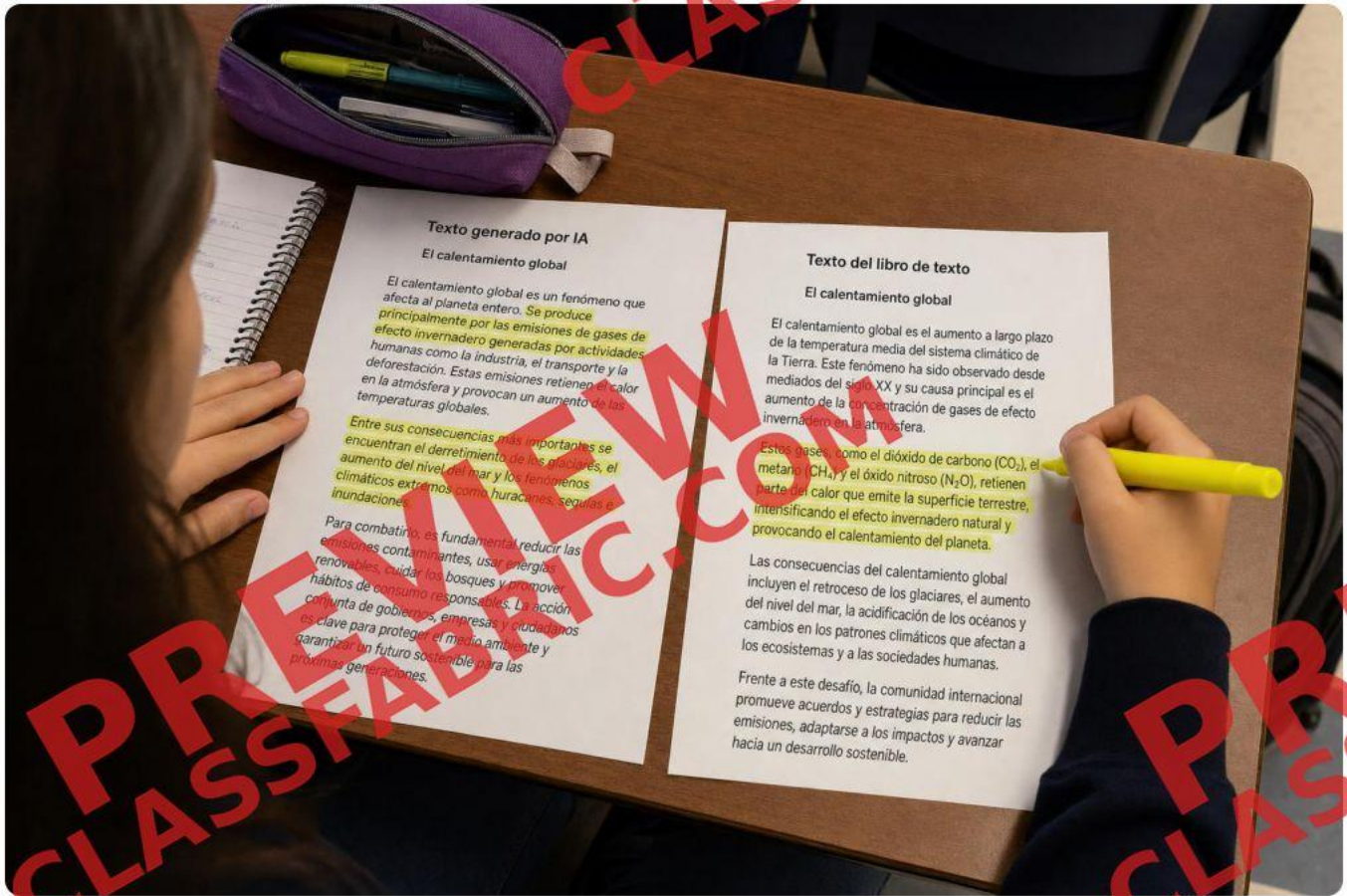
La instrucción o pregunta que un usuario escribe en una herramienta de IA, la cual influye directamente en la calidad y el enfoque de la salida que genera el modelo.

Explica con tus propias palabras:

Ensayo corto

Lee el enunciado cuidadosamente y escribe una respuesta en forma de ensayo corto bien organizada de aproximadamente 8-12 oraciones. Usa detalles y ejemplos específicos del artículo de la lección para apoyar tu respuesta. Escribe en oraciones completas.

Un compañero de clase te dice: "Le pregunté a un chatbot de IA sobre la nueva política de asistencia de nuestra escuela y me dio una respuesta muy detallada, así que debe ser correcta." Usando lo que aprendiste sobre cómo funcionan los grandes modelos de lenguaje — incluyendo el concepto de alucinación — explica por qué el razonamiento de tu compañero es erróneo. Luego, describe los pasos específicos que tomarías para evaluar si la respuesta de la IA es realmente confiable, aplicando el Marco Analítico de IA de la lección.

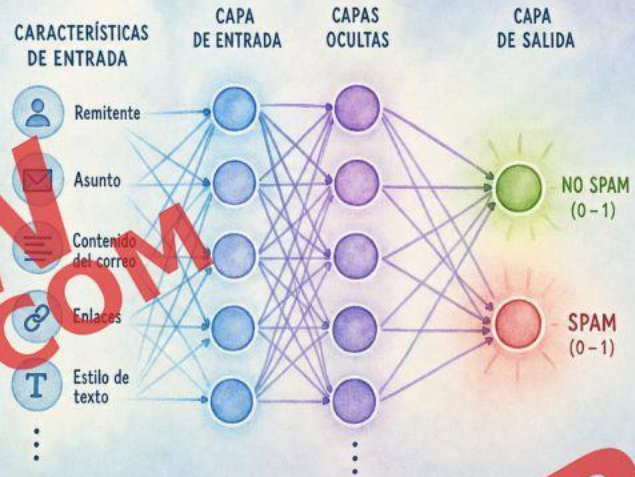


Verificar las afirmaciones generadas por IA contra fuentes confiables es un hábito que se aplica en cada materia y campo profesional.

FILTRO DE SPAM BASADO EN REGLAS



FILTRO DE SPAM BASADO EN REDES NEURONALES

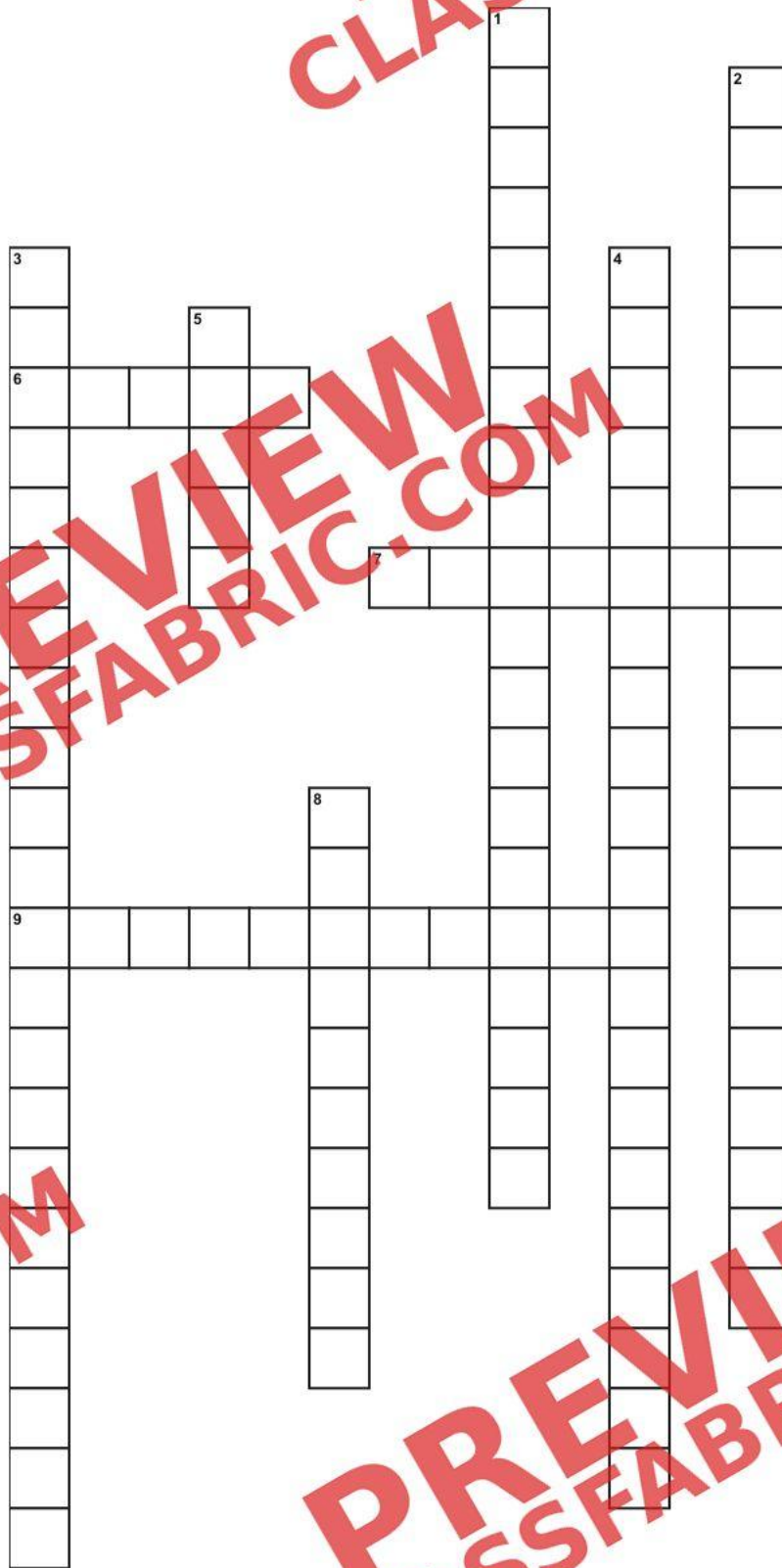


Aprende patrones automáticamente a partir de muchos correos etiquetados. Mejora con el tiempo y se adapta a nuevos tipos de spam.

Los programas tradicionales siguen reglas explícitas, mientras que los sistemas de aprendizaje automático construyen sus propios modelos internos aprendiendo de millones de ejemplos.

Inteligencia Artificial — Qué Es Realmente

CRUCIGRAMA



Inteligencia Artificial — Qué Es Realmente

CRUCIGRAMA

HORIZONTALES

6. Una unidad de texto — como una palabra o parte de una palabra — que un modelo de lenguaje grande predice y genera paso a paso para producir su salida.
7. La instrucción o pregunta que un usuario escribe en una herramienta de IA, la cual influye directamente en la calidad y el enfoque de la salida que genera el modelo.
9. Cuando una IA produce información falsa o fabricada con aparente confianza, típicamente porque la respuesta correcta no estaba presente en sus datos de entrenamiento.

VERTICALES

1. La gran colección de ejemplos — como texto, imágenes o correos electrónicos etiquetados — de la que un sistema de IA aprende patrones durante su desarrollo.
2. Una rama de la IA en la que un sistema aprende patrones a partir de grandes conjuntos de ejemplos etiquetados en lugar de seguir reglas escritas por un programador.
3. Un término amplio para sistemas informáticos diseñados para realizar tareas que típicamente requieren razonamiento similar al humano, como reconocer imágenes o generar texto.
4. El cambio en los trabajos disponibles que ocurre cuando la automatización o IA asume tareas previamente realizadas por trabajadores humanos.
5. Un defecto en un sistema de IA que ocurre cuando se entrena con datos no representativos, causando que produzca resultados injustos o inexactos para ciertos grupos de personas.
8. La capacidad de conciencia y sentimientos conscientes, que ningún sistema de IA actual posee a pesar de que a veces produce respuestas que parecen humanas.

Inteligencia Artificial — Qué Es Realmente

SOPA DE LETRAS

U J M U A L P F D I G B Z N G B B Q M W P U Y S D P
K M E I I N T E L I G E N C I A A R T I F I C I A L
C Y O Y W Y V A Y B S A C E Q C G E S S I V Q E T U
F K Z I A M E P Z N G P M Q G N B P U L P X C F Z I
A H O R L K W D P O M R J Y F J T A B L N C K T W M
S Z D O U Q P K Z M M E M F D A H H B X N L O D W R
X Z M E C L T E G Q U N U M E R D Q J S L W H A P G
F G X I I B J D N X A D C L S M W O Z D R P Z T X K
G T U G N U R D B T W I E Y P X Z C L R L D O O S G
J B Q V A F V H K W R Z W M L J M L Q P W D V S A I
T I P O C I Z N Q H R A U W A K G G V K C E F D C V
F Y F Q I Z M U U Z S J D J Z I L V S Z K N R E U N
I K I B O E B F V F Y E C A A P B F E V U P A E Z B
V V S S N M N H C E R A I E M U E Y N C T C C N R W
T F M T E Y K O G E Z U X T I D V C S T Q O O T Z F
P R W L F S J E N F E T N E E I N I I J C D K R T X
J Z I L O T G N B V W O K C N Y T G E X U K D E M Q
K T Y B H A Z O G O B M X B T I W K N Z E D R N N M
X S H Q W R M J G T Z A M X O I H N C F Z O S A Z W
K I K Q C E D T Y V D T N R L R E G I X R L J M M E
Z R O W Q J L U L C U I Q V A B W M A K P K X I I R
D C Z E X N Y J O J I C U H B W F Z N G J A J E X X
U R X C A O F T J M J O V K O H O I F K O F A N S V
H J S M S K O R E E M P T Y R L H W J U U S P T C P
L R F E I V X B H L B E R X A R V W A P W Y O O B J
J Q T G P Q T W N C C K F F L H F H E F S V Z C J C

ENCUENTRA ESTAS PALABRAS

Inteligencia Artificial

Sesgo

Desplazamiento Laboral

Token

Datos de Entrenamiento

Aprendizaje Automático

Entrada

Sensiencia

Alucinación