



EXPLORACIÓN DEL POTENCIAL DIDÁCTICO DE LAS ALUCINACIONES DE CHATGPT

Andrés Merino

Junio 2025

Facultad de Ciencias Exactas, Naturales y Ambientales





CONTENDIO

1. Introducción
2. ¿Cómo funciona ChatGPT?

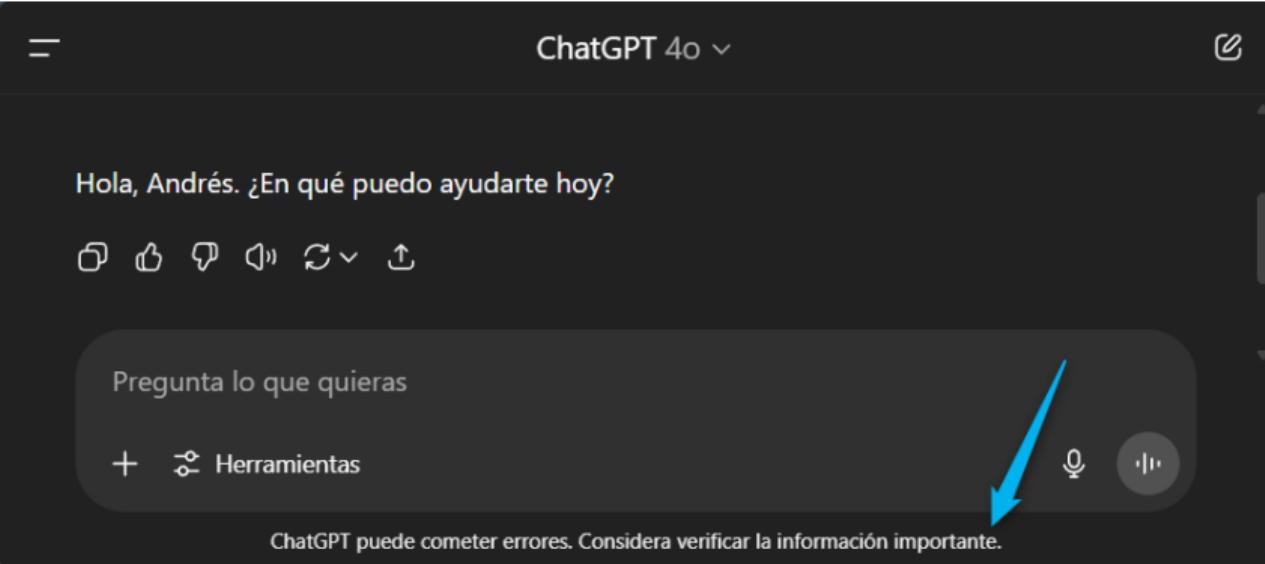


**Universidad
de Santander**
UDES

VISUALIZADA MINEDUCACIÓN | SNIES 3832



INTRODUCCIÓN



The screenshot shows a dark-themed ChatGPT interface. At the top, it says "ChatGPT 4o". Below that, a message from the AI reads: "Hola, Andrés. ¿En qué puedo ayudarte hoy?". Underneath the message are several small icons. A large input field below says "Pregunta lo que quieras". To the right of the input field is a button with a question mark icon and the number "0". A blue arrow points from this button down to a note at the bottom of the screen. The note reads: "ChatGPT puede cometer errores. Considera verificar la información importante.".

ChatGPT puede cometer errores: ¿prohibimos su uso en aula o lo aprovechamos para enseñar?



Objetivo de la charla

- Reflexionar sobre el funcionamiento de ChatGPT y su tendencia a generar respuestas incorrectas.
- Presentar una experiencia concreta en el aula donde las alucinaciones se usaron como recurso didáctico.
- Explorar el diseño de GPTs personalizados que inducen errores con fines pedagógicos.



Objetivo de la charla

- Reflexionar sobre el funcionamiento de ChatGPT y su tendencia a generar respuestas incorrectas.
- Presentar una experiencia concreta en el aula donde las alucinaciones se usaron como recurso didáctico.
- Explorar el diseño de GPTs personalizados que inducen errores con fines pedagógicos.

Caso de uso: Cálculo Diferencial



¿CÓMO FUNCIONA CHATGPT?



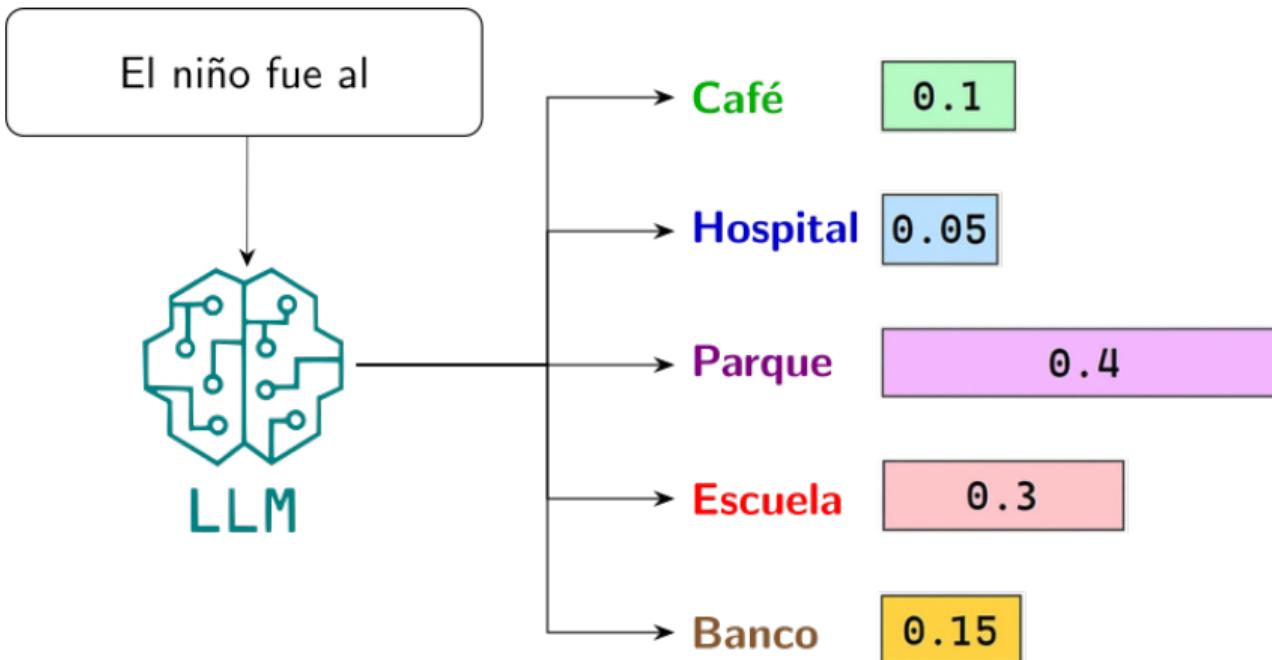
¿CUÁL ES LA PRÓXIMA PALABRA?

El niño fue al

- Café
- Hospital
- Parque
- Escuela
- Banco



¿CUÁL ES LA PRÓXIMA PALABRA?





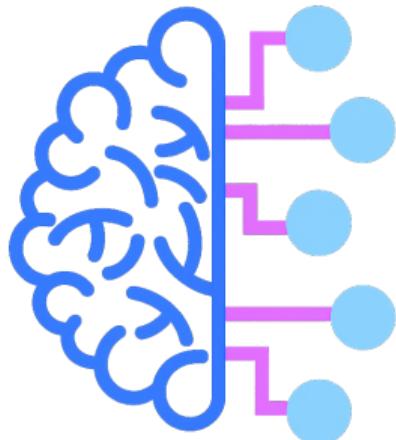
¿CÓMO FUNCIONA CHATGPT?

- ChatGPT es una herramienta que usa los Modelos GPT de OpenAI.
- Los modelos GPT de OpenAI son **modelos grandes de lenguaje** (*Large Language Model, LLM*).
- Su tarea principal es predecir la **palabra más probable** dada una secuencia anterior.
- No comprende el significado, solo calcula probabilidades a partir de patrones lingüísticos aprendidos.





PREDICCIÓN DE LA SIGUIENTE PALABRA



- El modelo asigna una **distribución de probabilidades** a cada posible palabra siguiente.
- No elige al azar: selecciona las opciones con mayor probabilidad.
- Por eso puede generar respuestas coherentes... o también **erróneas** de manera convincente.



¿CÓMO SE ENTRENÓ CHATGPT?

- GPT-3 fue entrenado con más de **300 mil millones de tokens** (~570 GB de texto limpio).
- Para GPT-4 se utilizó un volumen mucho mayor, aunque OpenAI no ha revelado cifras exactas.

Fuentes principales

- **Common Crawl** (filtrado y depurado).
- **WebText2, Books1, Books2, y Wikipedia en inglés.**



Gracias



Presentación

Contacto: aemerinot@puce.edu.ec