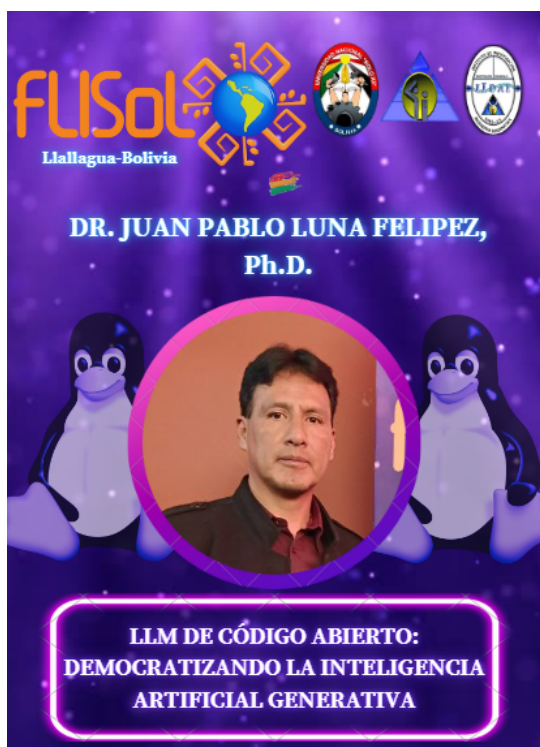

LLM DE CÓDIGO ABIERTO: DEMOCRATIZANDO LA INTELIGENCIA ARTIFICIAL GENERATIVA



JUAN PABLO LUNA FELIPEZ, Ph.D.

jplunaf@gmail.com

Ingeniería Informática

Universidad Nacional “Siglo XX”

Lla Laguna, Bolivia

RESUMEN

La inteligencia artificial generativa se enfoca en crear contenido original como texto, imágenes o código, usando modelos de lenguaje a gran escala (LLM) entrenados con enormes volúmenes de datos. Estos modelos, como los GPT, permiten aplicaciones como chatbots, traducción automática y generación de contenido. Mientras que muchos LLM son propietarios, los de código abierto ofrecen ventajas como transparencia, personalización, seguridad y menor dependencia de grandes empresas. Proyectos como Bloom, LLaMA, Falcon y otros promueven el acceso libre a esta tecnología, impulsando una IA más abierta, ética e innovadora.

1. INTRODUCCIÓN

La inteligencia artificial (IA) es un campo de estudio que se enfoca en la creación de máquinas y sistemas que pueden realizar tareas que normalmente requieren inteligencia humana, como el aprendizaje, la resolución de problemas, la toma de decisiones y la percepción.

2. DESARROLLO

Inteligencia Artificial Generativa

Es un tipo de inteligencia artificial (IA) que se centra en la creación de contenido nuevo y original, como texto, imágenes, audio, video y código. Aprende patrones y relaciones a partir de grandes conjuntos de datos para generar contenido similar al que se ha utilizado para entrenarla. La IA generativa utiliza diversas técnicas de aprendizaje automático, como las redes neuronales convolucionales (CNNs) y las redes neuronales recurrentes (RNNs), para procesar y comprender los datos de entrenamiento.

LLMs

LLMs (Large Language Models) son modelos de lenguaje de gran escala que utilizan técnicas de aprendizaje automático para procesar y generar texto. Estos modelos son entrenados con grandes cantidades de datos de texto y pueden aprender a reconocer patrones y relaciones en el lenguaje natural.

Aplicaciones

Chatbots y asistentes virtuales. Generación de contenido. Análisis de texto. Traducción automática.

Generar correos electrónicos, artículos de noticias, libros y guiones de películas, programas de computadora, software completo, artículos científicos.

Bases

- Estadísticas (frecuencias).
- Redes Bayesianas.
- Modelos de Markov.
- MHHM.
- Redes Neuronales Artificiales.
- Transformadores (pocos millones de parámetros).
- Redes Neuronales Convolutivas (CNN).
- Redes Neuronales Recurrentes (RNN).

Evolución y modelos

Comprensión de texto. Generación de texto. Modelos híbridos.

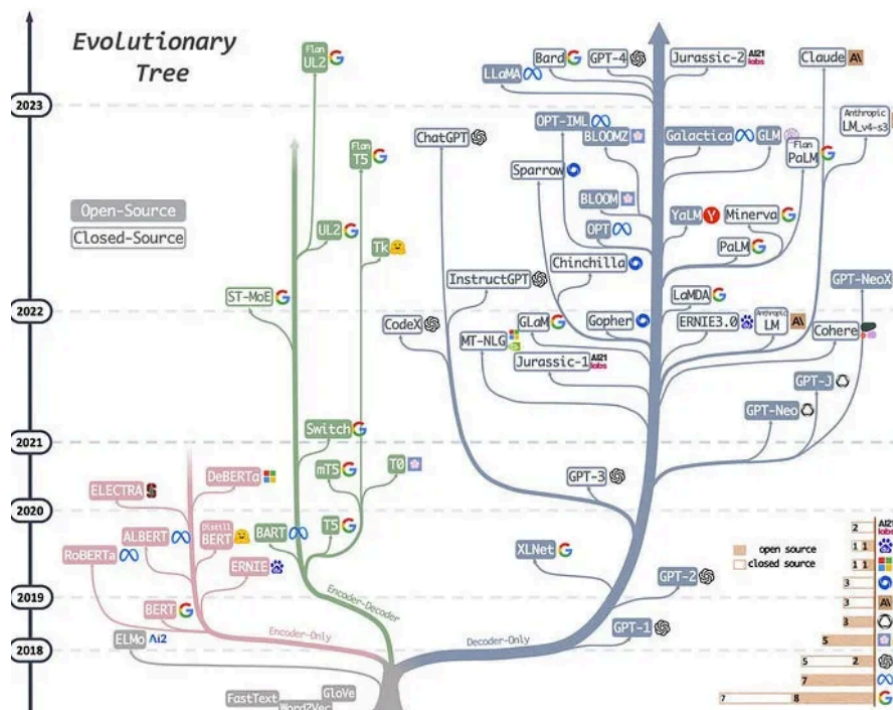


Figura 1: Evolución y modelos.

Parámetros

- GPT-1: 100 millones de parámetros.
- GPT-2: 1.500 millones de parámetros.
- GPT-3: 175.000 millones de parámetros.
- GPT-4: 1 billón de parámetros.
- El cerebro humano tiene 100 billones de conexiones.

Entrenamiento

Estos modelos ven más de 100 mil millones de palabras durante su entrenamiento, que es más de 100 veces lo que un humano escuchará o leerá en su vida! Aprenden mucho más lentamente que nosotros, pero tienen acceso a muchos (¡muchos!) más datos.

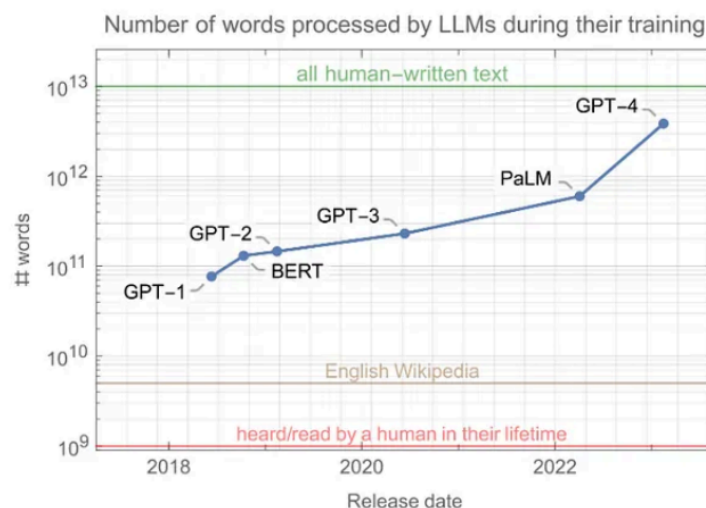


Figura 2: Gráfico del entrenamiento

Refinando Chatbots

- OpenAI: InstructGPT, ChatGPT (diciembre 2022).
- Anthropic lanzó Claude.
- Google lanzó Bard, luego Gemini.
- Meta lanzó LLaMA.

Dos meses después de su lanzamiento, ChatGPT ya tenía 100 millones de usuarios, el crecimiento de producto más rápido de la historia.

LLM propietarios

GPT, Bard, Gemini, Copilot.

Sus LLM son propietarios y sólo pueden ser utilizados por los clientes tras comprar una licencia. Esa licencia viene acompañada de derechos, pero también de posibles restricciones sobre cómo utilizar la LLM, así como de información limitada sobre los mecanismos que hay detrás de la tecnología. Controlados principalmente por OpenAI, Microsoft y Google.

LLM de código Abierto

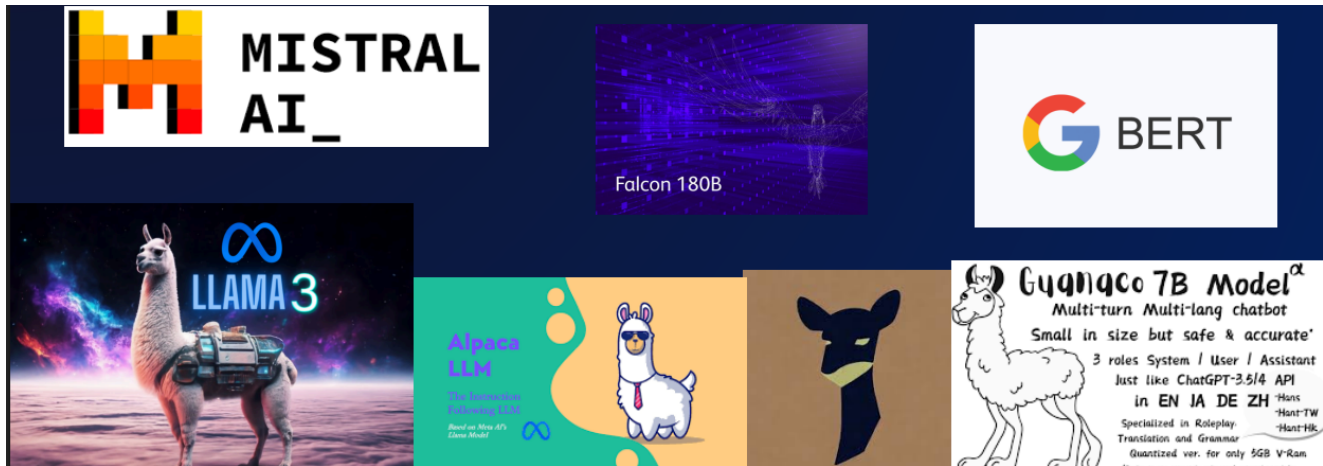


Figura 3: LLM de código abierto

Ventajas

- Mayor seguridad y privacidad de los datos.
- Ahorro de costes y reducción de la dependencia de proveedores.
- Transparencia del código y personalización del modelo de lenguaje.
- Apoyo activo de la comunidad y fomento de la innovación.
- Abordar la huella medioambiental de la IA.

Licencias

- MMistral Apache2 <https://github.com/simonw/llm-mistral/blob/main/LICENSE>
- Bloom MIT https://github.com/dptsa-300/start_with_bloom
- Bert Apache2 <https://github.com/google-research/bert/blob/master/LICENSE>
- Alpaca Apache2 https://github.com/tatsu-lab/stanford_alpaca
- OpenLlama Apache2 https://github.com/openlm-research/open_llama
- Falcon TII Falcon LLM License Version 1.0
<https://github.com/Decentralised-AI/falcon40b/blob/main/LICENSE.txt>
- Llama ACUERDO DE LICENCIA COMUNITARIA META LLAMA 3 <https://llama.meta.com/llama3/license/>

- Tabla Licencias y otros:
<https://info.deci.ai/hubfs/Downloadable%20Assets/%5BCheatSheet%5D-Top-Ten-List-of-LLMs-in-Open-Source-Arena.pdf>

Herramientas:

<https://deepinfra.com/chat>

<https://poe.com/>

Comparativa GTP vs Vicuña:

<https://chat.lmsys.org/>

3. CONCLUSIÓN

Los LLM de código abierto representan un gran paso hacia la democratización de la inteligencia artificial generativa. Frente a los modelos propietarios, ofrecen transparencia, adaptabilidad y libertad para investigadores, desarrolladores y empresas. Gracias al trabajo colaborativo de la comunidad y a la disponibilidad de licencias abiertas, estos modelos impulsan la innovación, reducen la dependencia tecnológica y abren el acceso a una IA más ética, accesible y sostenible para todos.



Figura 4: Fotografía de presentación de la ponencia