

LLMs – MODELOS DE LENGUAJE GRANDES

Una discusión sobre modelos generativos en la academia

Francisco Cruz
Doctor en Ciencias de la Computación



Los Modelos de Lenguaje Grandes (LLMs, por sus siglas en inglés) [1] son modelos de inteligencia artificial que están diseñados para comprender y generar lenguaje humano. Estos modelos son entrenados en vastas cantidades de texto para aprender patrones y estructuras lingüísticas, y luego se utilizan para realizar tareas relacionadas con el lenguaje, como la traducción automática, la generación de texto y la respuesta a preguntas.

Uno de los LLMs más conocidos en la actualidad es GPT (Generative Pre-trained Transformer) [2] desarrollado por OpenAI, que es un enfoque basado en redes neuronales y utiliza la arquitectura Transformer. Estos modelos se denominan “grandes” porque tienen un gran número de parámetros y se entrenan en grandes conjuntos de datos, lo que les permite capturar mejor la complejidad y sutileza del lenguaje humano. Particularmente, ChatGPT es una versión específica de GPT que se ha ajustado y entrenado específicamente para tareas de conversación y diálogo. ChatGPT está optimizado para responder a preguntas y participar en interacciones conversacionales.

Los LLMs han demostrado un gran avance en tareas de procesamiento del lenguaje natural, mejorando la calidad de las traducciones automáticas, la generación de texto coherente y la capacidad de respuesta en conversaciones con chatbots y asistentes virtuales. Sin embargo, también plantean desafíos en cuanto a su consumo de recursos computacionales y la necesidad de grandes conjuntos de datos de entrenamiento, además de su uso ético. Hoy en día existe una creciente preocupación de los riesgos del mal uso de esta herramienta, ya sea para estafas, fakenews o plagio [3].

