

GPT3

Que es

Generative Pre-trained Transformer 3, conocida por sus siglas (GPT-3), es un modelo de lenguaje autorregresivo que emplea aprendizaje profundo para producir textos que simulan la redacción humana. Es la tercera generación de los modelos de predicción de lenguaje perteneciente a la serie GPT, creados por OpenAI, un laboratorio de investigación de inteligencia artificial con sede en San Francisco

Historia sobre gpt3 (antecedentes)

Según The Economist, la mejora de los algoritmos, los potentes ordenadores y el aumento de los datos digitalizados han impulsado una revolución en el aprendizaje automático, con nuevas técnicas en la década de 2010 que han dado lugar a "rápidas mejoras en las tareas", incluida la manipulación del lenguaje. Una arquitectura utilizada en el procesamiento del lenguaje natural (PLN) es una red neuronal basada en un modelo de aprendizaje profundo que se introdujo por primera vez en 2017: el Transformer. Hay una serie de sistemas de PNL capaces de procesar, minar, organizar, conectar, contrastar, comprender y generar respuestas a preguntas.

El 11 de junio de 2018, los investigadores e ingenieros de OpenAI publicaron su artículo original sobre modelos generativos -modelos de lenguaje-sistemas de inteligencia artificial- que podían ser preentrenados con un enorme y diverso corpus de texto a través de conjuntos de datos, en un proceso que denominaron preentrenamiento generativo (GP). Los autores describieron cómo los rendimientos de la comprensión del lenguaje en el procesamiento del lenguaje natural (PLN) se mejoraron en GPT-3 a través de un proceso de "preentrenamiento generativo de un modelo de lenguaje en un corpus diverso de texto sin etiquetar, seguido de un ajuste fino discriminatorio en cada tarea específica." Esto eliminó la necesidad de supervisión humana y de etiquetado manual, que requiere mucho tiempo.

En febrero de 2020, Microsoft presentó su Turing Natural Language Generation (T-NLG), del que se decía que era el "mayor modelo de lenguaje jamás publicado, con 17.000 millones de parámetros"

Como funciona

GPT-3 permite crear contenido escrito gracias a su algoritmo lingüístico preentrenado. Es decir, funciona como una herramienta de predicción del lenguaje, por lo que es capaz de crear textos a partir de una orden o una tarea que se le encargue.

A este tipo de modelo de IA también se le conoce como machine learning, se entrena con varios parámetros (órdenes) que se convierten en datos que luego se traducen en respuestas. Para que se entienda mejor: para que GPT-3 funcione hay que alimentarlo con una porción de texto, el usuario debe darle un texto inicial para que, a partir de este, el sistema cree uno completamente nuevo.

Ventajas o beneficios

Automatización de procesos

Permite que las máquinas hagan de forma automática tareas que para los humanos resultan repetitivas y tediosas.

Reduce el error humano

Al reducir la intervención de los humanos en ciertos procesos, acaba con las posibilidades de que estos puedan cometer errores. Por ejemplo, una errata al introducir un dato en la contabilidad de un negocio.

Potencia la creatividad

Al liberar a los trabajadores de tareas repetitivas y poco motivadoras, la mente de estos es mucho más libre para dedicarse al proceso creativo.

Aporta precisión

Al ser capaz de tomar decisiones por sí misma, la inteligencia artificial da lugar a procesos productivos mucho más eficientes y con una menor tasa de error.

Agiliza la toma de decisiones

La inteligencia artificial es capaz de analizar miles de datos en apenas minutos y además tener en cuenta posibles actualizaciones de los mismos. La información bien sintetizada y actualizada ayuda a los profesionales a tomar decisiones estratégicas.

Capacidades que tiene

tiene 96 capas y 175 mil millones de parámetros. Es decir, “1.000 veces más que modelos previos como BERT (un modelo especializado en tareas de procesamiento de lenguaje natural que cuenta con 110 millones de parámetros)”, tal y como recalca de Pablo.

El experto cuenta que para entrenar el sistema se han utilizado en total 500 mil millones de palabras, extraídas de textos de la web, Wikipedia y varias colecciones de libros en abierto. Y también menciona que la forma de entrenamiento empleada se conoce como auto-supervisada, “puesto que son algoritmos que tratan de predecir la siguiente palabra, cualquier texto escrito por una persona nos sirve para entrenarlos”.

Detrás de este modelo de inteligencia artificial está OpenAI, una organización puesta en marcha por empresarios como Elon Musk, cofundador de Tesla y SpaceX, y Reid Hoffman, cofundador de LinkedIn, que nació con el objetivo de investigar y democratizar el acceso a la inteligencia artificial general. En sus orígenes, se trataba de una organización sin ánimo de lucro. Sin embargo, en 2020, se ha convertido en una empresa y se ha asociado con Microsoft con el fin de conseguir nuevos avances tanto en el ámbito del lenguaje, con modelos GPT-3, como en la robótica y la visión artificial.