



**Institución:** Universidad de Colima

**Escuela:** Facultad de Ingeniería Mecánica y Eléctrica

**Carrera:** Ingeniería en Computación Inteligente

**Materia:** Procesamiento del Lenguaje Natural

**Profesor:** Carrillo Zepeda Oswaldo

**Alumno:** Herrera Escareño Kevin Alejandro

**Grado y grupo:** 6°B

**Nombre de la actividad:** GPT 3

**Fecha:** 16 de Mayo de 2022

**Lugar:** Colima, Col.



## GPT-3

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. La versión completa de GPT-3 cuenta con 96 capas y una capacidad de 175.000 millones de parámetros de aprendizaje automatizado, lo cual supera la magnitud de su predecesor, GPT-2.



**Figura 1.** Logotipo de GPT-3.

GPT-3 fue introducido en mayo de 2020 y, hasta julio de 2020, se encontraba en fase beta. Es parte de una tendencia en sistemas de procesamiento de lenguaje natural (NLP) basados en "representaciones de lenguaje pre-entrenadas". 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.

## Antecedentes

Entre las arquitecturas empleadas en procesamiento de lenguaje natural se encuentran las de redes neuronales que utilizan un modelo de aprendizaje profundo denominado Transformer, introducido en 2017. Los modelos GPT-n se encuentran basados en dicha arquitectura de aprendizaje profundo.

El 11 de junio de 2018, los investigadores e ingenieros de OpenAI publicaron una investigación original sobre modelos generativos aplicados a sistemas de inteligencia artificial para procesamiento de lenguaje, los cuales podrían ser pre-entrenados con enormes cantidades de texto a través de bases de datos, en un proceso al que denominaron "pre-entrenamiento generativo" (generative pre-training o GP). Como consecuencia, este modelo podía prescindir de la supervisión humana y eliminaba el tiempo invertido en el etiquetado de los datos.

## Modelo GPT

El modelo GPT (Generative Pre-Training Transformer) fue publicado en el año 2018 por OpenAI. Está basado en la arquitectura del modelo Transformer, sobre el que se realiza una secuencia de aprendizaje semi-supervisado: un pre-entrenamiento no supervisado y un afinado supervisado.



Se utilizó un corpus con 7000 libros sin publicar para el entrenamiento previo, el cual ayudó al modelo a aprender sobre la estructura del lenguaje. Este entrenamiento utiliza un total de 117 millones de parámetros para la red neuronal, lo que supone un indicador que permite comparar la eficiencia entre modelos.

GPT resultó ser un gran avance, ya que, en el momento de su publicación, sus resultados superaban a la mayoría de modelos similares existentes. No obstante, fue superado con creces por su sucesor, el modelo GPT-2.

## **Modelo GPT-2**

Sucesor directo de GPT, el modelo GPT-2 fue publicado en el año 2019 por OpenAI. Contando con un número de parámetros para la red neuronal más de 10 veces mayor que su predecesor (1.500 millones de parámetros), fue pre-entrenado con 40 GigaBytes (GB) de datos, extraídos de las páginas web mejor valoradas de Reddit.

El modelo resultó ser tan potente, que sus desarrolladores decidieron no publicar la versión completa de 1.500 millones de parámetros, por miedo a que fuera utilizado con fines ilícitos, tales como la generación de noticias falsas que potencien la desinformación.

Así, la primera versión publicada de este modelo contaba con 117 millones de parámetros. No obstante, a medida que otros modelos similares con mayor potencia eran publicados, OpenAI decidió revelar versiones mejores, hasta finalmente dar a conocer el modelo al completo. Actualmente, existen cuatro versiones publicadas de GPT-2, en función de la cantidad de parámetros que requiere la red neuronal:

- 117M
- 345M
- 774M
- 1.500M

## **Funcionamiento**

GPT-3 es una inteligencia artificial, más en concreto se trata de un modelo de machine learning compuesto de algoritmos que están entrenados para reconocer patrones en los datos y aprender a través de los ejemplos. Dicho con palabras un poco más técnicas, es una serie de algoritmos que componen una Red Neuronal Recurrente con Memoria a largo plazo (LSTN Long Short-term memory).

Es decir, este programa analiza el texto o datos dados para, después, ofrecer predicciones de palabras y frases en base a las suministradas; a partir de ahí, la inteligencia artificial es capaz de continuar las frases o responder a preguntas.

## **Ventajas**

Las aplicaciones de GPT-3 se pueden ver en las tecnologías del lenguaje humano (procesamiento de lenguaje, habla, traducción, búsqueda o asistentes virtuales), en la generación de imágenes o audio



e incluso en las ayudas a la programación de 'software'. Es el caso de sistemas como GPT-2 (OpenAI), BERT y T-5 (Google), ROBERTA (Facebook) o CTRL (Salesforce).

Dentro de las ventajas que ofrece GPT-3, se encuentran:

- **Rastreo de ideas rápido:** Al ser una red neuronal, GPT-3 es capaz de, a partir de una serie de conceptos clave, rastrear y mostrar determinadas ideas que la IA considera primordiales en internet.
- **Optimiza el tiempo del research:** Al momento de realizar un research, GPT-3 es capaz de recolectar información acerca de lo que se está escribiendo, ahorrando así una gran cantidad de tiempo y esfuerzo.
- **Capta y desarrolla premisas:** Con solo desarrollar encabezados, será capaz de redactar las premisas que se expondrán en cada uno de los títulos.

## Capacidades

Con GPT-3 el equipo había duplicado la capacidad de su predecesor, GPT-2, haciendo del nuevo modelo el más robusto a la fecha. La gran cantidad de parámetros que posee GPT-3 lo otorga un mejor nivel de precisión respecto a anteriores modelos de menor capacidad, incluso supera en diez veces la capacidad del modelo Turing NLG de Microsoft.

Sesenta por ciento de los datos de pre-entrenamiento para GPT-3 provienen de una versión filtrada de Common Crawl consistente de 410 mil millones de elementos con codificación de pares de bytes. Otras fuentes son WebText2 que aporta 19.000 millones de elementos que representan el 22% del total; 12 mil millones de elementos provienen de Books1, representando el 8%; 55.000 millones elementos de Libros, y 3.000 millones de elementos de Wikipedia, representando el 3%. GPT-3 fue entrenado con cientos de miles de millones de palabras y es capaz de codificar en CSS, JSX, Python, entre otros lenguajes de programación.

## Bibliografía

Coco Solution. (2021, 23 marzo). Ventajas de usar GPT-3 para hacer contenido de redacción. Recuperado 14 de mayo de 2022, de <https://cocosolution.com/usar-gpt-3-copywriting/>.

Gómez, D. V. (2021, julio). Generación Automática de textos con GPT-2. Universidad de La Laguna. <https://riull.ull.es/xmlui/bitstream/handle/915/24743/Generacion%20Automatica%20de%20Textos%20con%20GPT-2.pdf?sequence=1>.

González, C. (2021, 12 marzo). ¿Qué es GPT-3?: la inteligencia artificial que se encargará de escribir por ti. BBVA NOTICIAS. Recuperado 14 de mayo de 2022, de <https://www.bbva.com/es/que-es-gpt-3-la-inteligencia-artificial-que-se-encargara-de-escribir-por-ti/>.

