

¿QUÉ ES LA INTELIGENCIA ARTIFICIAL GENERATIVA (IAG)?

escrito por Víctor D. Parra | 4 mayo 2023

Es muy probable que últimamente no hayas dejado de oír hablar de los aparentes “superpoderes” de **ChatGPT**, **DALL-E** y similares. Aquí en **MyTips** ya habíamos hablado con anterioridad acerca de la Inteligencia Artificial, el Aprendizaje Automático (*Machine Learning*) y el Aprendizaje Profundo (*Deep Learning*) con la intención de acercarlos al público general, así que ya estábamos tardando en hablar sobre la Inteligencia Artificial Generativa.

La **inteligencia artificial generativa (IAG)** es una rama de la **inteligencia artificial** que se dedica a crear datos o contenidos completamente nuevos a partir de los existentes, utilizando algoritmos de aprendizaje automático.

Esta tecnología ha existido desde hace mucho tiempo, y se remonta a la creación del [chatbot Eliza](#) en el MIT en 1966, pero los avances recientes en la IA generativa están revolucionando la forma en que se crea el contenido. Por ejemplo, **una IAG puede generar imágenes, textos, vídeos, música o código que no existían antes, pero que se parecen a los que utilizó para aprender.**

¿En qué se diferencia la IAG de la Inteligencia Artificial “normal”?

Aquí en **MyTips** ya habíamos hablado sobre **Inteligencia Artificial (IA)** -más abajo os dejo el enlace a post relacionados, por si no los habías leído ya- pero ¿qué diferencia hay entre “esa” IA y esta otra IAG?

La principal diferencia radica en los objetivos y la forma en que se utilizan para resolver problemas de aprendizaje automático.

La inteligencia artificial de la que ya habíamos hablado en este [post](#) recibe el nombre de **IA discriminativa** y se enfoca principalmente en tareas de **clasificación/categorización** es decir, en identificar y separar patrones específicos en los datos o en predecir una salida o etiqueta en función de una entrada dada, sin preocuparse por generar datos nuevos. Por ejemplo, **una IA discriminativa podría ser entrenada para distinguir entre perros y gatos en imágenes, clasificando cada imagen como “perro” o “gato”.**

Por contra, la **IA generativa se enfoca en la generación de datos nuevos y realistas, es decir, en la creación de nuevos conjuntos de datos que sean similares a los datos originales con los que fue entrenada.** Durante su entrenamiento, estos modelos aprenden la distribución subyacente de los datos y, a partir de esta, generan datos completamente nuevos pero similares a los datos originales. Por ejemplo, **una IA generativa podría ser entrenada para generar imágenes completamente nuevas de perros y gatos que sean indistinguibles de las imágenes reales.**

Resumiendo y simplificando mucho, una IA discriminativa, a partir de “estudiar” miles de fotografías etiquetadas de perros y gatos (miles de ejemplos de lo que es un perro y lo que es un gato) “aprende” como distinguirlos/clasificarlos (animales de 4 patas, 2 orejas, las de los perros generalmente más grandes, un hocico que en los perros suele ser más pronunciado que en los gatos, etc) mientras que la IA generativa, a partir de “estudiar” miles de ejemplos “aprende” cuáles son las características de un gato, de forma que cuando le pidamos que se invente un gato será capaz de crear uno similar a uno real, con todas sus características.

Algunos ejemplos de IAG son:

- **GPT-3** (Generative Pre-trained Transformer 3) y **GPT-4**: Es un modelo de lenguaje natural de gran tamaño (LLM, de *Large Language Model*) basado en redes neuronales artificiales, capaz de generar textos coherentes y fluidos en varios idiomas. Fue desarrollado por la empresa **OpenAI** y es considerado uno de los más avanzados del mundo en la actualidad.
- **DALL-E**: Es un modelo de inteligencia artificial desarrollado también por OpenAI que puede crear imágenes realistas a partir de descripciones textuales. Por ejemplo, si se le pide que genere una imagen de un “elefante rosa tocando la guitarra”, DALL-E producirá una imagen de un elefante rosa tocando la guitarra de manera realista.
- **Stable Diffusion**: es un modelo de aprendizaje automático basado en la técnica de difusión, que se utiliza para generar imágenes realistas a partir de “ruido aleatorio”. Permite generar imágenes de alta calidad con una mayor eficiencia y calidad que otros modelos de aprendizaje automático; las imágenes generadas son más consistentes y menos propensas a tener errores. A diferencia de modelos de la competencia como DALL-E, Stable Diffusion es de código abierto y no limita artificialmente las imágenes que produce.
- **Midjourney**: Midjourney es una inteligencia artificial capaz de generar imágenes a partir de descripciones textuales, similar a Dall-e de OpenAI y al Stable Diffusion de código abierto. Es creación del laboratorio independiente de investigación Midjourney.
- **AIVA** (*Artificial Intelligence Virtual Artist*): Es un programa de inteligencia artificial que puede componer música original en varios géneros y estilos. Se le da una melodía o un género musical y el programa crea una pieza musical completa, incluyendo instrumentos y armonías.
- **DeepDream**: Es un programa desarrollado por Google que utiliza una red neuronal para analizar y modificar

imágenes existentes. El programa busca patrones en las imágenes y los amplifica, creando imágenes surrealistas y psicodélicas.

- **Bard**: es un bot conversacional desarrollado por Google, como respuesta a ChatGPT, al que muchos consideran una verdadera amenaza al prácticamente monopolio de Google en las búsquedas en internet. Está basado en [LaMDA](#) de *Language Model for Dialogue Applications* (modelo de lenguaje para aplicaciones de diálogo) una familia de modelos de lenguaje desarrollados por Google.
- **Runway ML**: Es una plataforma que permite a los usuarios crear modelos de inteligencia artificial generativa sin necesidad de programación. Ofrece herramientas para crear modelos de síntesis de voz, imágenes y video. Como ellos mismos dicen, "Runway es un nuevo tipo de suite creativa. Una donde la IA es un colaborador y donde cualquier cosa que puedas imaginar puede ser creada"... su eslogan comercial es *"Everything you need to make anything you want"*.

Estas quizás sean las de un uso "más serio" o que pueden dar una idea de las tremendas capacidades de la IAG; pero hay muchas más. Algunas son meros "ejercicios" académicos y otras no son más que simples entretenimientos como esta IA que no hace tanto [se hizo muy popular porque "daba vida" a fotografías antiguas](#). Otras muchas IAG están "embebidas" -o lo estarán- en software que utilizamos a diario, en forma de asistentes, chatbots, etc. Un ejemplo, [Copilot de GitHub](#) una IA que ayuda a los usuarios de esta plataforma para desarrolladores software, las intenciones de [Microsoft de dotar a su suite ofimática Office de un asistente basado en Inteligencia Artificial](#) o las de [Google para Gmail y Google Docs](#).

¿Cómo funciona la IAG?

No soy un experto en el tema, así que me limitaré a dar unas pinceladas sobre el principio de funcionamiento de la IAG.

Algunas de las técnicas empleadas en las IAG son las siguientes:

- Redes Generativas Antagónicas (GAN, siglas en inglés de *Generative Antagonic Networks*)
- Modelos Autorregresivos (AR)
- Modelos Variacionales (VAE)

Redes generativas antagónicas

Las [redes generativas antagónicas](#) (GAN, por sus siglas en inglés) es una de las técnicas más populares y potentes para crear IAG.

Fueron presentadas por [Ian Goodfellow](#) y sus colegas en 2014. Goodfellow había trabajado como científico investigador en **Google Brain** (un equipo de investigación de **Deep Learning** bajo el paraguas de Google AI) y fue director de aprendizaje automático en **Apple**. Goodfellow ha realizado varias contribuciones importantes al campo del aprendizaje profundo (aquí [sus artículos en Google Académico](#))

Las GAN son dos redes neuronales artificiales -una red generadora y otra discriminadora- que compiten entre sí. La red generadora se esfuerza en crear datos falsos con el objetivo de engañar a la red discriminadora, quien a su vez intenta distinguir entre los datos reales y los falsos. De esta manera, ambas redes se entrenan mutuamente, mejorando sus respectivas capacidades para generar y detectar datos.

Modelos Autoregresivos (AR)

Se trata de un tipo de modelo en el que se predice el valor de una variable basándose en sus valores previos. Un modelo AR utiliza la información pasada para hacer una predicción de la variable en el futuro. Estos modelos son útiles en la generación de texto, donde se pueden utilizar las palabras anteriores para predecir la siguiente palabra, por ejemplo, en la función de “Teclado Predictivo” de un *smartphone*.

Modelos Variacionales (VAE)

Los modelos variacionales (VAE) son una clase de modelos generativos que se utilizan para la creación de imágenes y otros tipos de datos complejos. Estos modelos se basan en la teoría de la probabilidad y en la capacidad de aprendizaje de las redes neuronales.

Por ejemplo, pueden crear imágenes de caras que no existen, pero que se parecen a las caras reales. Los VAE operan mediante tres etapas:

1. Primero, toman una entrada -siguiendo con el ejemplo anterior, la imagen de una cara- y la transforman en un código compuesto por dos partes: una media y una varianza. Estas dos partes definen una distribución de probabilidad, generalmente una gaussiana, que representa la entrada de forma compacta.
2. A continuación, los VAE eligen al azar un valor de esa distribución y lo utilizan para reconstruir la entrada original, como la imagen. Este valor es una variable aleatoria latente que captura las características esenciales de la entrada.
3. Finalmente, comparan la entrada original con la reconstrucción y ajustan los parámetros del modelo para minimizar la diferencia entre ambas. De esta manera, el modelo aprende a generar datos similares a los que ha “visto”.

Los VAE tienen varias aplicaciones, como eliminar ruido en imágenes, crear nuevas imágenes, reducir dimensiones o clasificar objetos.

Así fue como una IAG fue capaz de pintar un cuadro imitando el estilo de Rembrandt. 347 años después de su fallecimiento, tenemos un cuadro nuevo.

¿Para qué sirve la IAG?

La IAG tiene muchas aplicaciones prácticas en diferentes sectores e industrias. Algunos ejemplos son:

- **Arte y entretenimiento:** la IAG puede crear obras de arte originales, como pinturas, esculturas o música, imitando el estilo de artistas famosos o creando uno propio. También puede generar guiones, personajes o escenarios para películas, videojuegos o cómics.
- **Medicina y salud:** la IAG puede generar imágenes médicas sintéticas, como radiografías, resonancias o ecografías, que sirvan para entrenar a los médicos o a los sistemas de diagnóstico automático. También puede generar modelos tridimensionales de órganos o tejidos humanos para la investigación biomédica o la impresión 3D.
- **Educación y formación:** la IAG puede generar textos educativos o informativos sobre cualquier tema, adaptados al nivel y al idioma del lector. También puede generar ejercicios, exámenes o simulaciones para evaluar el aprendizaje o el desempeño de los estudiantes o los trabajadores.
- **Marketing y publicidad:** la IAG puede generar imágenes o vídeos de productos o servicios que no existen todavía, pero que se ajustan a las preferencias o necesidades de los clientes potenciales. También puede generar slogans, logos o campañas publicitarias personalizadas para cada segmento de mercado.
- **Seguridad y privacidad:** la IAG puede generar datos

sintéticos que preserven la privacidad de las personas o las empresas, pero que mantengan las características estadísticas de los datos originales. Así se pueden compartir o analizar los datos sin riesgo de revelar información sensible o confidencial.

Conclusiones

La IAG es una [tecnología innovadora y disruptiva que abre un mundo de posibilidades para la creatividad y la solución de problemas](#). Sin embargo, también plantea algunos desafíos éticos y legales que hay que tener en cuenta, como el riesgo de generar contenidos falsos (automatizar *phishing*) o engañosos (*fake news*) que puedan afectar a la verdad; un ejemplo, esta [app que permite “clonar tu voz”](#), con multitud de usos “malvados”.

Otro frente es cómo afectan a los derechos de autor (como denuncian los [guionistas de Hollywood](#) o los [ilustradores japoneses](#)). Las IAG se entrenan con contenidos desarrollados por otros, estilos que luego pueden imitar.

Hay experiencia del uso de la tecnología para manipular a la opinión pública, por ejemplo, el [escándalo de Facebook y Cambridge Analítica](#) ¿cómo podría polarizarse la sociedad en un fuego cruzado de mensajes optimizados para una determinada audiencia, repletos de fotografías, vídeos y/o audios falsos indistinguibles de los reales?



Fotografía falsa de la detención de Donald Trump, generada por una IAG

[Hay también preocupación de que destruya empleos “creativos”](#) (redacción de textos, diseño gráfico, ilustradores, actores de doblaje, programadores, etc) que hasta ahora parecían encontrarse a salvo de la automatización pura y dura o de la Inteligencia Artificial discriminativa.

Los más optimistas afirman que como en anteriores ocasiones en la Historia, la aparición de una nueva tecnología implica la desaparición de algunos trabajos, pero también la creación de otros completamente nuevos que nadie podía haber imaginado hasta entonces, como por ejemplo la de [prompt engineer](#), la persona que sabe cómo hacerle peticiones a la IA para que genere los mejores resultados de la manera más efectiva y eficiente. Otro empleo, menos glamuroso (y menos lucrativo) es el de [entrenador de IA](#), un trabajo a menudo [“relocalizado” en países con mano de obra muy barata](#). Pero un estudio de Goldman Sachs publicado el pasado 26 de marzo señalaba que, según estimaciones realizadas a partir de datos de EEUU y Europa, [casi dos tercios de los trabajos actuales estaban expuestos a cierto grado de automatización de IA](#). La pregunta es entonces ¿cuál será el balance de nuevos puestos de trabajo creados

frente a los destruidos?

Por la experiencia del pasado, toda revolución tecnológica ha supuesto la desaparición de empleos, la transformación de otros y la creación de algunos completamente nuevos. La cuestión con la IA en general y la IAG en particular es ¿cuál será el balance final?

En un post anterior dedicado a la IA ya veíamos como Stephen Hawking y Elon Musk advertían sobre los posibles peligros de la IA. Lo cierto es que el desarrollo parece imparable, y es rapidísimo, por eso, de nuevo Elon Musk (curiosamente uno de los socios fundadores de OpenAI, la empresa de ChatGPT), Steve Wozniak, Harari y otros [firmaron un documento pidiendo una moratoria en el desarrollo de la IAG](#). Recientemente han sido también noticia las [declaraciones de Geoffrey Hinton](#), ex ingeniero de Google y bautizado como “el padrino de la IA”. Aunque cree que ahora mismo la IA no es más inteligente que los humanos, muy pronto lo será y advierte de algunos riesgos: la creación de contenido digital cuya veracidad será imposible de comprobar para el “usuario promedio”, el reemplazo de trabajadores en una amplia gama de oficios y que los sistemas de IA se conviertan un día en armas autónomas, especies de “robots asesinos”.

Este post ha sido creado parcialmente empleando IAG: su estructura, algunos aspectos de su redacción, etc han sido revisados por una IAG (en concreto, ChatGPT 3.5 y ChatGPT 4 a través de su integración en el buscador Bing de Microsoft). La imagen que ilustra la cabecera de este post (ese robot sentado frente a una máquina de escribir en lo que parece el estudio de un escritor) fue generada por DALL-E. También he empleado la IAG para tratar de optimizar su posicionamiento SEO en buscadores...

Si bien mediante múltiples iteraciones y retoques podría haber ampliado el texto hasta convertirlo en [todo un libro](#), ChatGPT

todavía no es perfecto, y puede “alucinar” que es como sus creadores llaman a cuando la IAG directamente se inventa información que al estar tan bien redactada, puede parecer completamente cierta. Esto puede suponer un peligro si alguien pretende usar a una IAG como médico, para apoyo psicológico, etc.

Espero que este post escrito con la ayuda de una IAG os haya parecido interesante. Como siempre, tenéis disponible el apartado de Comentarios y muchas gracias por dedicar vuestro tiempo a leerme.

Para saber más

- [Estos son los 9 mejores libros para ponerse al día con el nuevo y apasionante mundo de la inteligencia artificial generativa, según los expertos](#) artículo del 28/04/2023 en la revista on line **Business Insider**.
- [Sobre algoritmos y puestos de trabajo](#), artículo de [Enrique Dans](#) en su blog sobre la preocupación por la posible pérdida de puestos de trabajo provocada por la Inteligencia Artificial.
- [Estamos todos despedidos... o no](#), artículo de José Luis Casal en su blog **Better Business for a Better World**, tratando el tema de la pérdida de empleos provocada por la automatización.
- [The Potentially Large Effects of Artificial Intelligence on Economic Growth \(Briggs/Kodnani\)](#) el informe de la consultora **Goldman Sachs** que mencionaba más arriba.
- [Deep Learning](#) acceso gratuito a la versión on line del libro escrito por Ian Goodfellow, Yoshua Bengio y Aaron Courville. Los autores lo definen como un recurso para ayudar a los estudiantes y a los profesionales a que se adentren al campo del aprendizaje automático en general, y en particular en el aprendizaje profundo.
- [Openai.com](#) la web de la empresa creadora de ChatGPT

- [*Dall-E: Creating images from text*](#) , entrada en el blog de **OpenAI** explicando mediante ejemplos cómo funciona.
- [*The Next Rembrandt*](#) un corto documental en YouTube resumiendo el proceso mediante el cual una IA fue capaz de “pintar” un retrato completamente original con el estilo de [Rembrandt](#). Aquí la [web del proyecto](#) para una explicación mucho más detallada.

Artículos relacionados, aquí en MyTips

[*NOCIONES SOBRE INTELIGENCIA ARTIFICIAL*](#)

[*¿VAN A ENVIARNOS LOS ROBOTS AL PARO?*](#)