

INTELIGENCIA ARTIFICIAL GENERATIVA

ENRIQUE PUERTAS

Escuela de Arquitectura, Ingeniería y Diseño
UNIVERSIDAD EUROPEA DE MADRID

La inteligencia artificial generativa es una tecnología fascinante y prometedora que permite generar contenidos nuevos y originales y que está cambiando la forma en que interactuamos con el mundo digital. Desde los primeros sistemas basados en redes neuronales, a las más avanzadas técnicas basadas en transformers y modelos de difusión, como los pulares chatGPT o Dall-e, estos sistemas están en boca de todo el mundo y han provocado que nos replanteemos la forma en la que se estaban realizando muchas tareas. El objetivo de este artículo es ayudar a entender lo que hay por debajo y el cómo funcionan estos modelos, y presentar el increíble abanico de posibilidades que se abren en todos los sectores gracias a estas nuevas tecnologías dentro del campo de Inteligencia Artificial.

PALABRAS CLAVE •

Inteligencia Artificial Generativa, Transformers, Modelos de Difusión, Redes Generativas Adversarias

CÓMO CITAR ESTE ARTÍCULO •

Puertas, Enrique. 2023. "Inteligencia Artificial Generativa" en: UEM STEAM Essentials. Enlace web UEM :: http://projectbasedschool.universidadeuropea.es/escuela/escuela/steam_essentials

INTRODUCCIÓN

En ocasiones ocurren casualidades maravillosas que producen un gran impacto en nuestra sociedad. Cuando Alexander Fleming descubrió la penicilina en 1928 fue porque se le contaminó de moho una placa de Petri, y cuando en 1895 Wilhelm Conrad Röntgen descubrió los Rayos X, en realidad estaba experimentando con tubos de rayos catódicos. A lo largo de la historia de la ciencia, numerosos descubrimientos han sido el resultado de la serendipia, un término que se refiere a hallazgos afortunados e inesperados realizados mientras se buscaba algo distinto.

A finales del año 2022, fruto también de una inesperada casualidad, las aparece el sistema de Inteligencia Artificial chatGPT, que permite generar textos y dar respuesta a preguntas realizadas en Lenguaje Natural, y que ha supuesto toda una revolución por su impacto y las implica-

ciones socioculturales. Aunque hoy en día es un sistema que se está usando de forma masiva (es ya el software con el mayor crecimiento de usuarios de la historia), en realidad chatGPT se creó con un propósito totalmente distinto al que luego se ha acabado convirtiendo. ChatGPT se creó inicialmente como una herramienta auxiliar para mejorar el sistema de Generación de textos GPT-3 (Generative Pretrained Transformer versión 3) creado por la empresa OpenAI, y que era capaz de escribir de manera coherente textos que son sintáctica y gramaticalmente correctos, pero que tenía el problema de que muchas veces los textos generados eran inseguros desde el punto de vista del lenguaje y el contenido. Es decir, a veces generaba textos políticamente incorrectos (machistas, racistas, lenguaje ofensivo, etc.). ChatGPT fue creado inicialmente como una herramienta auxiliar para mejorar la interacción conversacional entre

la IA y los usuarios humanos, con el fin de que el sistema principal (GPT-3) aprendiera generar respuestas más precisas, contextualizadas y seguras. Y mientras se probaba el sistema chatGPT para mejorar las respuestas de GPT-3, la gente empezó a usar la herramienta como un asistente para escribir trabajos, resolver dudas, realizar tareas de forma más eficaz... teniendo tanto éxito que pilló con el pie cambiado incluso a sus propios creadores. Aunque posiblemente no haya sido una casualidad al nivel de la que permitió el descubrimiento de la penicilina o los rayos X, este uso inesperado de la herramienta ha supuesto toda una revolución y ha producido numerosos cambios en nuestra sociedad. El futuro dirá el alcance de estos cambios.

LOS CONCEPTOS DE INTELIGENCIA ARTIFICIAL E INTELIGENCIA ARTIFICIAL GENERATIVA

Pero en este artículo no se va a hablar chatGPT, al menos de forma exclusiva, sino de la tecnología subyacente que ha permitido la creación de este sistema que parece que va a cambiar la forma de trabajar de las personas, y muchos de los procesos de las empresas: Vamos a hablar de la Inteligencia Artificial Generativa (IAG), que es una disciplina dentro de la IA y el Deep Learning, que se enfoca en el uso de algoritmos para crear contenido de nuevo como por ejemplo textos, imágenes, videos, o música. Para entender mejor qué es y cómo funciona la IAG, vamos primero a explicar la relación que existe entre los términos "Inteligencia Artificial", "Machine Learning" (Aprendizaje Automático), "Deep Learning" e "Inteligencia Artificial Generativa", términos interconectados y muy relacionados, pero que no son los mismo.

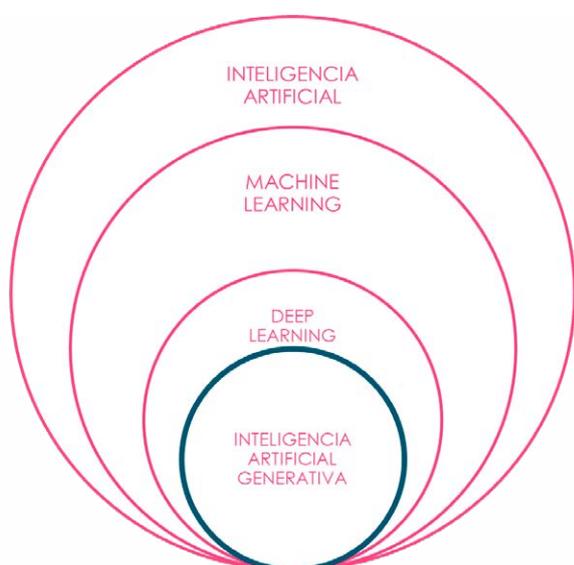
1) » Inteligencia Artificial (IA): La IA es la disciplina que busca desarrollar sistemas que puedan imitar la inteligencia humana en tareas específicas. Estos sistemas pueden incluir desde algoritmos simples hasta sistemas más complejos basados en modelos cognitivos.

2) » Machine Learning (Aprendizaje Automático): El Machine Learning es una rama de la IA que se centra en desarrollar algoritmos y modelos capaces de aprender y mejorar automáticamente a través de la experiencia. Estos modelos aprenden a partir de datos y utilizan algoritmos y técnicas como la regresión, clasificación y agrupamiento para tomar decisiones o hacer predicciones.

3) » Deep Learning (Aprendizaje Profundo): El Deep Learning es una subrama del Machine Learning que se basa en usar una familia de algoritmos llamados redes neuronales artificiales. Estas redes neuronales son capaces de aprender y extraer características complejas a través de múltiples capas, lo que les permite realizar tareas como reconocimiento de imágenes, procesamiento de lenguaje natural y mucho más.

4) » Inteligencia Artificial Generativa: La Inteligencia Artificial Generativa es un campo emergente dentro de la IA que se centra en desarrollar sistemas capaces de generar contenido nuevo y original. Estos sistemas utilizan técnicas como el aprendizaje no supervisado y las redes generativas adversarias para crear imágenes, música, texto y otros tipos de contenido.

Por tanto, la Inteligencia Artificial Generativa es un campo dentro de la Inteligencia Artificial que se centra en la creación de modelos capaces de generar contenidos nuevos y realistas a partir de patrones aprendidos en conjuntos de datos existentes usando Machine Learning, más concretamente un tipo de Machine Learning que se llama Deep Learning y que se caracteriza por usar Redes Neuronales para el aprendizaje y descubrimiento de patrones. Estas relaciones atienden a la "jerarquía" de los términos, aunque existen otras clasificaciones en otros ejes conceptuales como pueden ser la similitud con la inteligencia humana o la forma de entrenar los algoritmos, por ejemplo.



ÁMBITOS Y EJEMPLOS

En los últimos años, la IA generativa ha experimentado avances significativos y ha expandido sus aplicaciones a una amplia gama de dominios, como el arte, la música, la moda, la arquitectura y muchos otros. En algunos de ellos, está transformando la forma en que creamos, diseñamos y comprendemos el mundo que nos rodea. En otros, está mejorando y haciendo más eficientes los procesos y operaciones existentes.

El hecho de que la IA generativa se utilice en muchos dominios también implica que sus modelos pueden manejar di-

Figura 1 » Muestra la relación entre estas cuatro áreas:

ferentes tipos de datos, desde lenguaje natural hasta audio o imágenes. A continuación se muestran algunos ejemplos de estos dominios en los que se está usando:

Generación de texto

Una de las mayores aplicaciones de la IA generativa, es su capacidad para producir nuevo contenido en lenguaje natural. De hecho, los algoritmos de IA generativa se pueden utilizar para generar textos originales, como artículos, novelas, poesías o códigos de programación, por ejemplo. Su capacidad para crear contenido coherente en varios idiomas ha hecho que la IAG se empiece a utilizar de forma habitual para generar textos de diferentes géneros y con diferentes estilos. Un ejemplo destacado dentro del dominio de la generación de textos es el ya mencionado GPT-3, desarrollado por OpenAI (*Introducing ChatGPT, n.d.*). Este modelo ha sido utilizado para crear artículos periodísticos, narrativas creativas o dar respuestas a preguntas complejas, por nombrar solo algunos de los ejemplos de uso de esta herramienta.



Figura 2 » Imágenes generadas por IAG.

Uno de los sectores que más se está beneficiando de la aparición de la IAG ha sido el de la publicidad y el marketing. Las empresas pueden utilizar la IAG para generar descripciones de productos, mensajes en redes sociales, anuncios con impacto y mensajes personalizados para sus clientes. Esto permite ahorrar tiempo y recursos, al tiempo que se mantiene un alto nivel de calidad en el contenido generado.

Otro caso de uso muy relevante en la generación de texto, por la cantidad de puertas que abre, es el campo de la traducción automática. La IAG ha demostrado ser una herramienta válida y valiosa para esta tarea. Los modelos basados en IAG han mejorado significativamente la calidad de las traducciones automáticas al capturar mejor el contexto y las sutilezas del lenguaje y esto ha facilitado la comunicación y el intercambio de información entre personas que hablan diferentes idiomas.

Generación de Imágenes

Uno de los primeros y más conocidos ejemplos de IA generativa en creación de imágenes son las arquitecturas de Redes Generativas Adversarias (GAN), (Goodfellow et al., 2014). El objetivo de las GAN es generar imágenes realistas que sean indistinguibles de imágenes reales. Para ello se hace competir dos redes neuronales una contra otra; La primera trata de generar imágenes que parezcan “reales”, mientras que la segunda trata de detectar “fraudes” (imágenes falsas) Esta capacidad ha tenido varias aplicaciones interesantes en el ámbito empresarial, como generar conjuntos de datos sintéticos para entrenar modelos de visión por computadora, generar imágenes de productos realistas y generar imágenes realistas para aplicaciones de realidad virtual y realidad aumentada.

En 2021 se introdujo en este campo un nuevo modelo de IA generativa creado por OpenAI (los mismos que estaban detrás del GPT-3 y chatGPT), llamado DALL-E (**This Avocado Armchair Could Be the Future of AI | MIT Technology**

Review, n.d.). A diferencia de las GAN, el modelo DALL-E está diseñado para generar imágenes a partir de descripciones en lenguaje natural (las GAN toman un vector con ruido aleatorio como entrada) y puede generar una amplia gama de imágenes, que podrían parecer poco realistas pero que aun así representan los conceptos deseados. Este tipo de modelos de IAG para generar imágenes tienen un gran potencial en industrias creativas como la publicidad, el diseño de productos y la moda, entre otras, para crear imágenes originales únicas y creativas.

Generación de audio y música

LA IAG también ha demostrado su potencial para crear música y sonidos originales, y se aplica para tareas de composición musical, producción de efectos de sonido o creación de paisajes sonoros.

En el ámbito de la composición musical, la IAG ha sido utilizada para generar piezas musicales en diferentes estilos y géneros. Los modelos de IAG pueden analizar grandes

conjuntos de datos musicales existentes y aprender las estructuras y patrones subyacentes para generar nuevas composiciones musicales. Un ejemplo destacado es el proyecto “Daddy’s Car” de Sony CSL (Díaz-Limon, 2016), donde se utilizó la IAG para componer una canción completa en el estilo de los Beatles.

Además de la composición musical, la IAG ha sido aplicada en la creación de efectos de sonido para películas, videojuegos y producciones audiovisuales. Los modelos de IAG pueden generar sonidos personalizados que se ajusten a las necesidades específicas de cada proyecto.

La IAG también ha encontrado aplicaciones en la creación de paisajes sonoros y ambientes atmosféricos. Los modelos de IAG pueden generar sonidos ambientales que imiten entornos naturales o crear paisajes sonoros inmersivos para experiencias virtuales o de realidad aumentada. Esto permite a los creadores audiovisuales enriquecer sus producciones con sonidos envolventes y cautivadores.

Un caso de uso de la IAG para generar audio ha sido el de la clonación y generación de voces. Algunas plataformas permiten crear modelos de voces que imitan a la perfección el tono y timbre de voz de una persona real.

TÉCNICAS QUE POSIBILITAN LA IAG

En los comienzos de la IAG, los modelos se basaban en técnicas que, aunque innovadoras, tenían muchísimas limitaciones en cuanto a la calidad y coherencia de los resultados generados. Esos primeros tipos incluyen modelos basados en reglas heurísticas y estadísticas simples, que sentaron las bases para desarrollos más sofisticados. A medida que avanzaba la investigación en aprendizaje automático, surgieron aproximaciones más avanzadas, como las Redes Neuronales Recurrentes (RNN por sus siglas en inglés) y las Redes Generativas Adversarias (también llamadas GAN por sus siglas en inglés) que ofrecieron saltos cualitativos en la

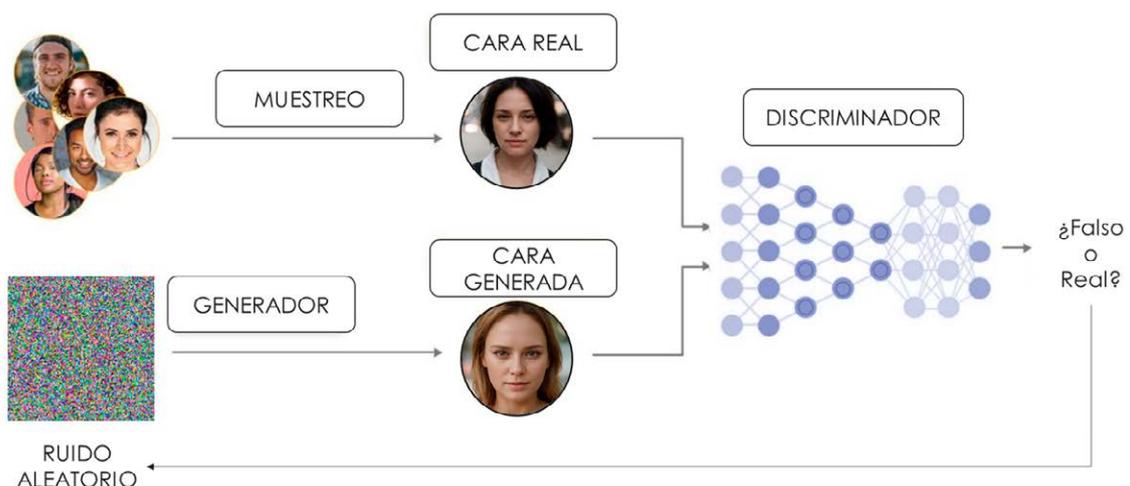
capacidad generativa de los modelos, y que ya eran capaces de generar de forma automatizada textos e imágenes con bastante buena calidad (aunque no perfectos).

En el caso de las GANs, se entrenan dos modelos: un generador y un discriminador, que compiten entre sí para mejorar la calidad del contenido generado. El generador aprende a generar contenido nuevo y el discriminador aprende a distinguir entre el contenido generado y el real. Esto permite a la IAG crear contenido nuevo y único, contenido que nunca antes ha sido visto o experimentado, lo que la hace especialmente útil en campos como el arte, la música y el diseño.

Las Redes Neuronales Recurrentes (Rumelhart et al., 1986) son un tipo de arquitectura de redes neuronales artificiales que están diseñadas para procesar y analizar secuencias de datos, donde la salida de cada paso depende de la entrada actual y de la información anterior almacenada en su memoria interna. A diferencia de las redes neuronales tradicionales, que procesan datos de forma independiente, las RNN tienen conexiones recurrentes (de ahí su nombre) que les permiten mantener y actualizar información a lo largo de una secuencia. Esto las hace especialmente adecuadas para el procesamiento de datos secuenciales, como texto, audio o series temporales.

Las RNN se han usado mucho para la generación de texto, como por ejemplo en los sistemas de texto predictivo y autocompletado de los teclados de teléfonos móviles, donde la red neuronal recibe una secuencia de palabras y se le pide que genere la siguiente palabra o palabras más probables en función del contexto anterior. Otra aplicación importante ha sido la generación de texto a partir de una descripción. Por ejemplo, proporcionando una frase inicial y pedirle a la red neuronal que genere un párrafo completo que siga el estilo y la estructura del texto de entrenamiento. Esto se logra alimentando la salida generada en cada paso de tiempo nuevamente a la red como entrada en el siguiente paso, permitiendo que la red “recuerde” y genere texto coherente y contextualmente relevante.

Figura 3 » Entrenamiento de una GAN



Además, las RNN se han combinado también con otras técnicas, como las redes generativas adversariales (GAN), para generar texto más realista. Estas combinaciones han llevado a avances significativos en la generación de texto, permitiendo la creación de historias, poesía, música y otros tipos de contenido textual.

Sin embargo, tanto las Redes Neuronales Recurrentes, como las Redes Generativas Adversarias, presentan algunos problemas que las limitan.

En el caso de las GAN, puede aparecer un problema conocido como “modo colapso” o “colapso del modo de generación”. Esto ocurre cuando el generador de la GAN produce contenido que sigue siempre el mismo patrón o muestra poca variabilidad en las salidas generadas. En lugar de generar una amplia diversidad de resultados, la red generadora se estanca en la producción de un número limitado de ejemplos o patrones, lo que puede llevar a una falta de originalidad y diversidad de los contenidos generados. Este problema puede surgir debido a varios factores, pero el más habitual es el desequilibrio entre el generador y el discriminador en el proceso de entrenamiento. Si el discriminador es demasiado efectivo en identificar las falsificaciones generadas por el generador, puede dificultar que el generador explore y genere nuevas variaciones de contenido.

Por otro lado, Las Redes Neuronales Recurrentes pueden tener dificultades para reconocer y capturar relaciones complejas entre palabras en textos largos, debido por un lado a un problema conocido como “desvanecimiento o explosión del gradiente”, y por otro por la limitación de su memoria a corto plazo.

El desvanecimiento o explosión del gradiente ocurre durante el entrenamiento de las RNN cuando los gradientes (la información de capas anteriores) que se propagan a través de las capas de la red se vuelven extremadamente pequeños o extremadamente grandes. Esto puede llevar a que la red tenga dificultades para aprender y capturar correlaciones a largo plazo entre palabras en la secuencia de texto.

Además, las RNN tienen una memoria a corto plazo limitada, lo que significa que a medida que se procesa una secuencia de texto larga, la red puede olvidar información relevante al depender únicamente de la información más reciente. Esto puede dificultar la capacidad de la red para capturar y recordar relaciones de largo alcance en el texto. Estos problemas pueden llevar a que las RNN tengan dificultades para reconocer relaciones sutiles y dependencias de largo alcance en el texto, lo que puede afectar su capacidad para comprender y generar secuencias de texto coherentes y contextualmente relevantes.

Para abordar estas limitaciones de las GANs y de las RNNs,

surgieron dos arquitecturas de IA Generativa que son las que realmente han supuesto toda una revolución en el campo de la IA Generativa: Las redes “transformers” y los modelos de difusión. La primera particularmente en el procesamiento del lenguaje natural (PLN), y la segunda en la generación de imágenes. Vamos a ver en qué consisten estas dos tecnologías y cómo se han entrenado para que sean capaces de generar textos e imágenes muy difíciles de diferenciar de contenidos creados por humanos.

La aparición de las arquitecturas transformer ha revolucionado el campo del procesamiento del lenguaje natural (PLN) y ha tenido un impacto significativo en la inteligencia artificial generativa (IAG). Los transformers destacan por su capacidad para gestionar secuencias de datos con dependencias a largo plazo, superando a las arquitecturas recurrentes y convolucionales que dominaban anteriormente el campo.

Los transformers fueron introducidos por Vaswani et al. (2017) en su trabajo titulado “Attention is All You Need”. El modelo transformer es una arquitectura de red neuronal que se basa exclusivamente en mecanismos de atención para ponderar la influencia de diferentes partes de una secuencia de entrada. La idea central es que no todas las partes de la secuencia son igualmente importantes, y los mecanismos de atención permiten al modelo centrarse en las partes más relevantes para realizar una tarea. Es decir, mientras que en las redes recurrentes una palabra está relacionada con las anteriores con un peso que depende de la “distancia” entre las palabras, con los transformers se hace un cálculo de los “pesos” de las relaciones entre las palabras dentro del contexto del texto, y ese peso depende de lo relacionadas que estén las palabras más que de lo lejos que estén una de otra.

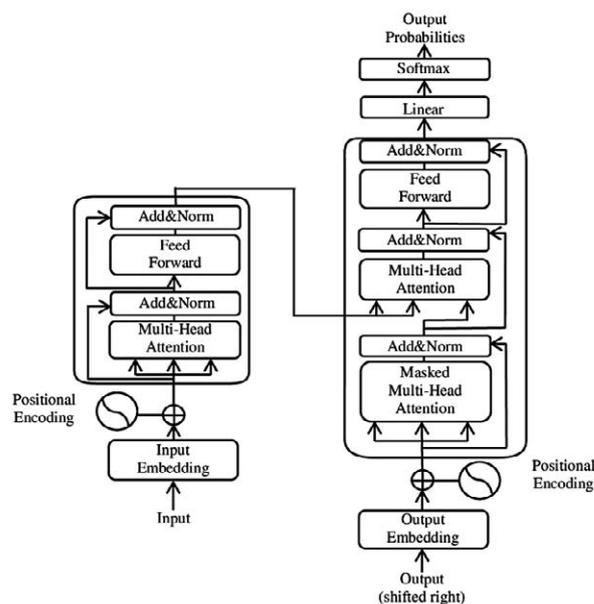


Figura 4 » Arquitectura Transformer. Yuening Jia - DOI:10.1088/1742-6596/1314/1/012186

Los transformers funcionan de la siguiente manera: En primer lugar, el texto de entrada se divide en palabras o tokens individuales. Cada palabra se representa utilizando lo que se conoce como word embeddings (Mikolov et al., 2013), que son vectores numéricos que capturan el significado y las características de cada palabra. Es decir, cada palabra se representa como una lista de números que es una especie de "ADN" de esa palabra para un significado y un contexto concreto. Estos embeddings ayudan a la red a entender el contexto y las relaciones entre las palabras. Además, se agrega información posicional a los embeddings para que la red pueda entender el orden y la posición de las palabras en la secuencia. Esta información posicional es importante porque ayuda a la red a capturar las estructuras y las dependencias en el texto.

A continuación, el modelo original de transformers usaba un sistema de **codificador-decodificador**: El codificador recibe la secuencia de embeddings y procesa cada palabra de manera individual. Utiliza los mecanismos de atención para calcular la importancia o relevancia de cada palabra en relación con las demás palabras de la secuencia. Esto permite que la red capture las relaciones y las dependencias entre las palabras de manera más eficiente.

La arquitectura codificador-decodificador es la clave de las arquitecturas transformer. El codificador se encarga de en-

Son los mecanismos de atención los que permiten solventar el problema de memoria a corto plazo que tenían las Redes Neuronales Recurrentes.

MODELOS DE DIFUSIÓN

En el año 2016 se publicó el artículo titulado "Improved Techniques for Training GANs" (Técnicas mejoradas para entrenar GANs) (Salimans et al., 2016), en el que se introdujo un enfoque novedoso para el entrenamiento de las Redes Generativas Adversariales (GANs) mediante la difusión de ruido en lugar de alimentar una entrada de ruido aleatorio directamente al generador. Esta técnica de difusión permitió generar imágenes más realistas y de mayor calidad, y solventar el problema del colapso que limitaba la variedad en el contenido generado.

Los modelos de difusión son una forma de modelo generativo que se basa en la idea de revertir un proceso de difusión, que es un proceso estocástico que degrada sucesivamente la información de los datos originales hasta convertirlos en ruido puro. El proceso de generación se conceptualiza como la inversión de este proceso, es decir, partiendo de un patrón de ruido aleatorio y refinándolo progresivamente hasta alcanzar una muestra coherente y de alta calidad.



Figura 5 » Proceso de degradación y reconstrucción de los modelos de difusión.

tender el texto de entrada y extraer información relevante, mientras que el decodificador utiliza esa información para generar texto nuevo y coherente. El decodificador también utiliza mecanismos de atención para enfocarse en las partes relevantes del texto de entrada durante la generación.

Los mecanismos de atención son fundamentales en las arquitecturas transformer. Permiten que la red se enfoque en las partes más importantes del texto de entrada mientras genera texto nuevo. La atención se calcula mediante la comparación de las palabras en una secuencia y la asignación de pesos a cada palabra en función de su relevancia.

En cada paso, se agrega una pequeña cantidad de ruido a la imagen existente, lo que hace que la imagen se vuelva más y más degradada. Este ruido puede ser aleatorio o seguir alguna distribución específica. A medida que se repiten estos pasos de degradación, la imagen se va desvaneciendo y perdiendo detalles, hasta que finalmente se convierte en ruido puro.

Una vez que se ha generado el ruido, se invierte el proceso para generar imágenes nuevas. Esto se hace mediante un proceso llamado muestreo inverso o reconstrucción.

Se toma el ruido generado y se aplica una serie de pasos de reconstrucción para revertir el proceso de degradación.

En cada paso de reconstrucción, se toma el ruido existente y se le aplica una transformación inversa para recuperar gradualmente los detalles y la estructura de la imagen original. Estos pasos de reconstrucción se repiten hasta que se obtiene una imagen final que es una versión generada de la imagen original de alta calidad. El proceso de muestreo inverso funcionaría de la siguiente manera:

- 1) » Comienzas con un vector de ruido que actúa como una especie de “código” para la imagen reconstruida.
- 2) » Este vector de ruido se pasa a través de una red generativa, que es una red neuronal diseñada para generar imágenes. La red generativa toma el vector de ruido como entrada y produce una imagen reconstruida como salida.
- 3) » Durante la generación de la imagen reconstruida, la red generativa utiliza una serie de transformaciones y operaciones para ir agregando gradualmente detalles y estructuras a partir del vector de ruido.
- 4) » A medida que se repiten estos pasos, la imagen reconstruida se va refinando y se acerca cada vez más a la imagen original. La red generativa ajusta los parámetros internos en cada paso para mejorar la calidad y la similitud con la imagen original.
- 5) » Finalmente, después de un número determinado de pasos o iteraciones, se obtiene la imagen reconstruida que se asemeja a la imagen original en la medida de lo posible.

Además, el uso de textos en los modelos de difusión puede desempeñar un papel importante en la generación de imágenes nuevas y específicas. Los textos se utilizan como una descripción o guía para generar imágenes que se ajusten a ciertos conceptos o características deseadas. Al combinar el texto con el proceso de difusión y reconstrucción, se puede guiar la generación de imágenes para que se ajusten a la descripción o concepto proporcionado en el texto. Esto permite una generación más controlada y específica de imágenes.

La generación de imágenes, por ejemplo, estos modelos han demostrado ser capaces de producir imágenes realistas y detalladas que son a menudo indistinguibles de las fotografías reales. Esto los hace particularmente útiles para aplicaciones como la creación de arte digital, el diseño de modas y la visualización arquitectónica, aunque también plantea retos y controversia por el mal uso que se pueda hacer de esta nueva tecnología.

Estos modelos no sirven solo para generar imágenes, sino que también se pueden aplicar en otros ámbitos como la

generación de audio, donde los modelos de difusión han sido aplicados con éxito para la generación de música y síntesis de voces, ofreciendo mejoras en la naturalidad y la expresividad del audio generado. Además, en la química y la biología, los modelos de difusión están facilitando el descubrimiento de nuevas moléculas y el diseño de proteínas, al permitir la exploración de espacios de alta dimensionalidad de una manera más controlada y eficiente.

CONTROVERSIA Y MAL USO

Aunque la IAG es una tecnología con un enorme potencial para mejorar la vida de las personas y hacernos más productivos en el trabajo, no está exenta de polémica por el riesgo que existe de que se use de forma inapropiada y con fines poco éticos e incluso ilícitos. La generación de textos, imágenes y audios artificiales muy difíciles de distinguir de contenidos reales puede ser usado para perpetrar fraudes o esparcir noticias engañosas, las famosas “fake news”. Con la IAG se pueden crear muy fácilmente noticias con textos e imágenes falsos que intenten hacer pasar un bulo por algo real (Aïmeur et al., 2023). Por otro lado, la clonación de voces, se puede usar para imitar la voz de personas conocidas con el fin de engañar a las personas y realizar estafas, por ejemplo.

Además del uso fraudulento de la IAG, también está el impacto negativo que ésta puede tener sobre ciertas profesiones y negocios. El sector de creación de contenidos audiovisuales está en general bastante preocupado por la facilidad con la que ahora se pueden crear canciones o ilustraciones, por ejemplo. Algunas plataformas como Spotify han empezado a prohibir en la plataforma obras generadas con Inteligencia Artificial y algunos sectores profesionales como los guionistas de series y películas están intentando “blindarse” protestando para que se introduzca en sus convenios la prohibición del uso de inteligencia artificial para la generación automatizada de guiones.

Por último también está sobre la mesa el tema de la propiedad intelectual. ¿A quién pertenece la autoría de un contenido generado mediante IAG? ¿Qué pasa con los autores de las obras que se han usado para entrenar los modelos generativos? Estas y otras cuestiones están aún abiertas y tendremos que esperar un tiempo para ver cómo se resuelven.

CONCLUSIÓN

La evolución de la inteligencia artificial en la generación de contenido ha experimentado avances muy significativos en los últimos años, desde modelos iniciales basados en reglas heurísticas y estadísticas simples hasta el desarrollo de arquitecturas más sofisticadas como las redes neuronales recurrentes y las redes generativas adversarias. La introducción de los transformers y los modelos de difusión ha permitido superar limitaciones previas y ha ampliado el alcance de la IA generativa en diversos dominios, incluyendo el arte, la música, la moda y la arquitectura. Asimismo, la creación de sistemas como chatGPT ha revolucionado

la interacción conversacional en lenguaje natural, demostrando el impacto significativo de la IA generativa en la sociedad actual. A pesar del enorme potencial de esta tecnología y de la gran cantidad de puertas que abre, no podemos perder de vista los peligros que puede acarrear si no se gestiona de forma adecuada. Por eso habrá que seguir explorando y desarrollando tecnologías de IA generativa para seguir transformando y mejorando diversos aspectos de nuestra vida cotidiana, pero sin perder de vista los aspectos ético y el posible mal uso que se pueda hacer de ella.

REFERENCIAS BIBLIOGRÁFICAS

Aimeur, E., Amri, S., & Brassard, G. (2023). *Fake news, disinformation and misinformation in social media: a review*. Social Network Analysis and Mining 2023 13(1), 1-36. <https://doi.org/10.1007/S13278-023-01028-5>

Díaz-Limon, J. A. (2016). *Daddy's Car: La Inteligencia Artificial como Herramienta Facilitadora de Derechos de Autor*. Revista La Propiedad Inmaterial, 22. <https://heinonline.org/HOL/Page?hand=hein.journals/revpropin22&sid=83&div=&collection=>

Goodfellow, I. J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., & Bengio, Y. (2014). *Generative Adversarial Networks*. Science Robotics, 3(January), 2672-2680. <https://arxiv.org/abs/1406.2661v1>

Introducing ChatGPT. (n.d.). Retrieved January 26, 2024, from <https://openai.com/blog/chatgpt>

Mikolov, T., Sutskever, I., Chen, K., Corrado, G., & Dean, J. (2013). *Distributed Representations of Words and Phrases and their Compositionality*. Advances in Neural Information Processing Systems. <https://arxiv.org/abs/1310.4546v1>

Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). *Learning representations by back-propagating errors*. Nature 1986 323:6088, 323(6088), 533-536. <https://doi.org/10.1038/323533a0>

Salimans, T., Goodfellow, I., Zaremba, W., Cheung, V., Radford, A., & Chen, X. (2016). *Improved Techniques for Training GANs*. Advances in Neural Information Processing Systems, 2234-2242. <https://arxiv.org/abs/1606.03498v1>

This avocado armchair could be the future of AI | MIT Technology Review. (n.d.). Retrieved January 26, 2024, from <https://www.technologyreview.com/2021/01/05/1015754/avocado-armchair-future-ai-openai-deep-learning-nlp-gpt3-computer-vision-common-sense/>

BIOGRAFÍA

Enrique Puertas es Ingeniero Informático y Doctor en Tecnologías de la Información Aplicadas. Es profesor del departamento de Ciencias, Computación y Tecnología de la Universidad Europea de Madrid, entidad en la que también ha dirigido el Máster Universitario en Big Data Analytics. Es director del grupo de investigación "Inteligencia Artificial e Interacción humano-Máquina", y miembro del grupo "IASalud", y sus líneas de investigación se centran en el uso de técnicas de Inteligencia Artificial y Big Data aplicadas al ámbito clínico y sistemas inteligentes de transporte. Es también co-autor del libro "Manual práctico de Inteligencia Artificial en entornos sanitarios". En su faceta docente cuenta con más de 20 años de experiencia como profesor en el área de informática, impartiendo asignaturas, en grado y postgrado, relacionadas con programación, Interfaces de Usuario, Inteligencia Artificial, y Big Data. Como investigador ha participado en más de una veintena proyectos y contratos de investigación, europeos, nacionales, y regionales, y es autor de más de cincuenta publicaciones en revistas y congresos internacionales. Dentro de estas publicaciones cabe destacar 24 de ellas por estar indexadas en índices JRCR/SJR y cuenta con un SEXENIO de investigación reconocido por ANECA-CNEAI.

