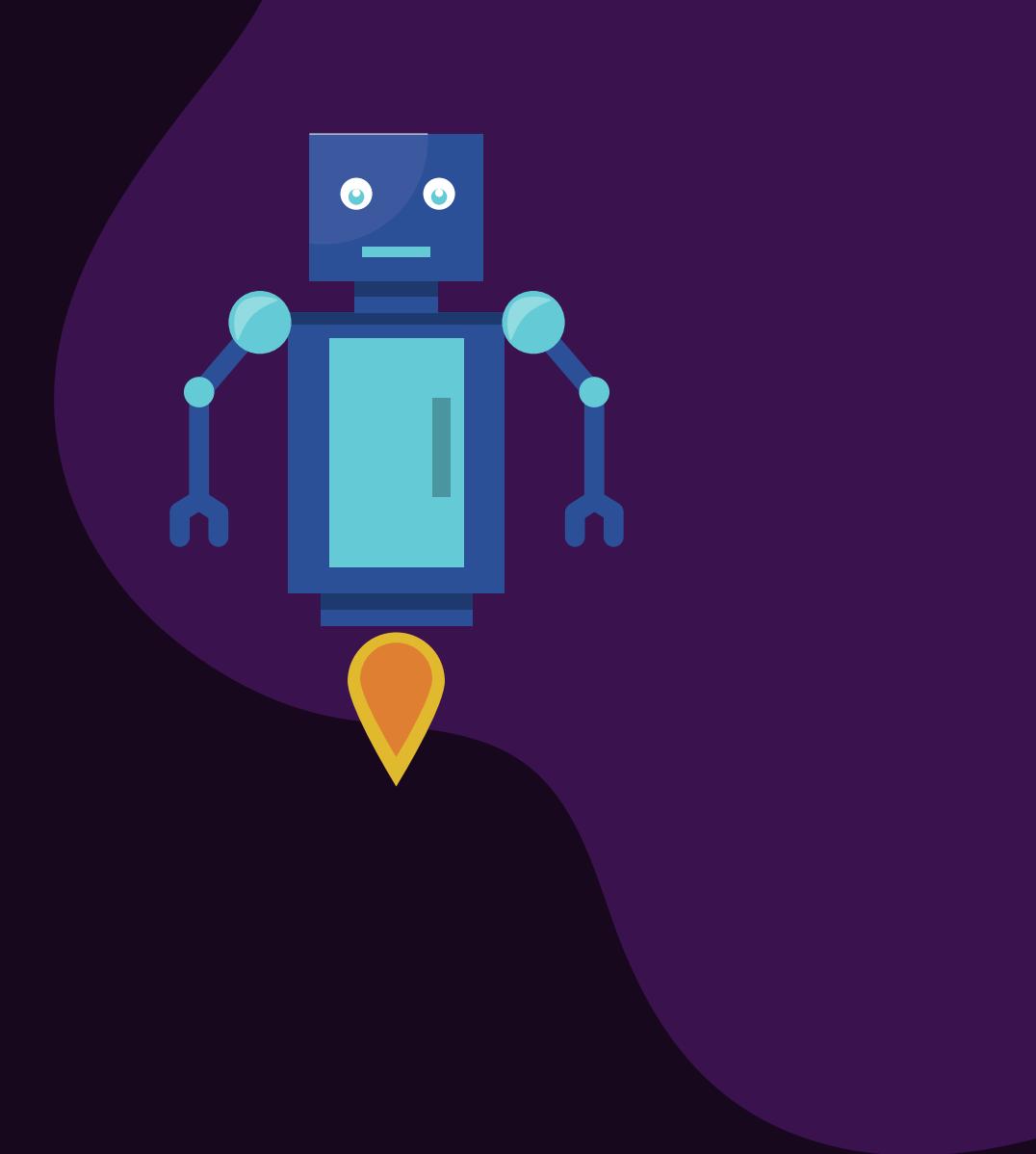
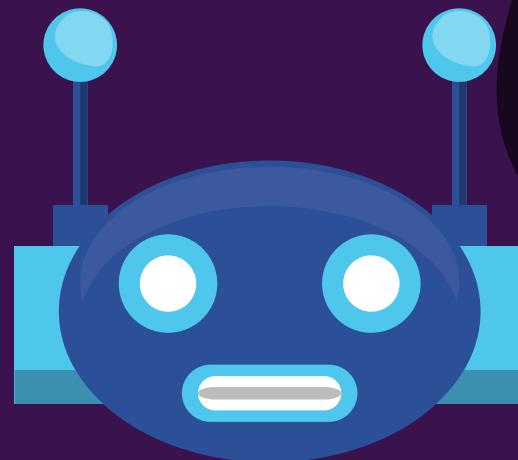


INTELIGENCIA ARTIFICIAL

Clase 12

Tutor: José Andrés Montenegro Santos
Miércoles, 18 de diciembre de 2024

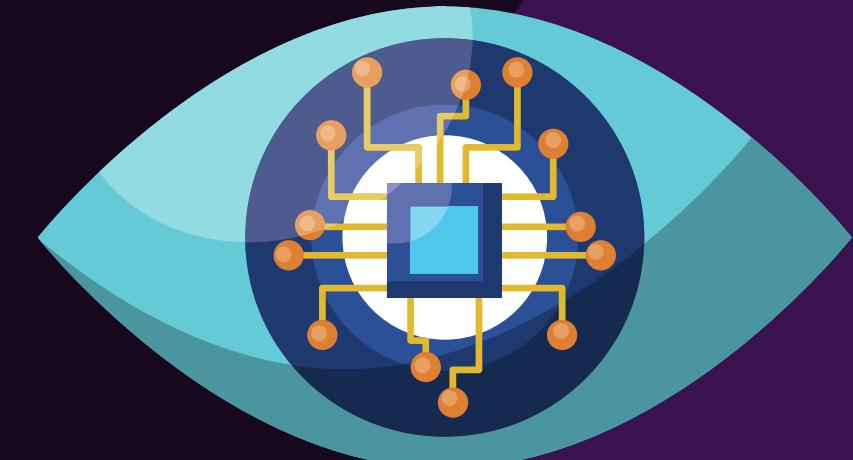
INTELIGENCIA ARTIFICIAL GENERATIVA



¿Qué es?

La IA generativa se basa en sofisticados modelos de aprendizaje automático llamados modelos de aprendizaje profundo , algoritmos que simulan los procesos de aprendizaje y la toma de decisiones del cerebro humano. Estos modelos funcionan mediante la identificación y programación de los patrones y relaciones en grandes cantidades de datos, y luego usando esa información para comprender las solicitudes o preguntas del lenguaje natural de los usuarios y responder con nuevo contenido relevante.

La inteligencia artificial generativa, también conocida como IA generativa o IA gen, es un tipo de IA capaz de crear nuevos contenidos e ideas, como conversaciones, historias, imágenes, videos y música. Es capaz de aprender lenguaje humano, lenguajes de programación, arte, química, biología o cualquier otro tema complejo. Reutiliza aquello que sabe para resolver nuevos problemas.



¿COMO FUNCIONA?

Capacitación

La IA generativa comienza con un modelo fundacional, un modelo de aprendizaje profundo que sirve de base para múltiples tipos de aplicaciones de IA generativa. Los modelos fundacionales más comunes hoy en día son los modelos de lenguaje de gran tamaño (LLM), creados para aplicaciones de generación de texto, pero también hay modelos básicos para la generación de imágenes, videos, sonido y música, así como modelos fundacionales multimodales que pueden admitir varios tipos de generación de contenido.

El resultado de este entrenamiento es una red neuronal de parámetros (representaciones codificadas de las entidades, patrones y relaciones en los datos) que pueden generar contenido de forma autónoma en respuesta a entradas o instrucciones.

Ajuste

Metafóricamente hablando, un modelo fundacional es generalista: sabe mucho sobre muchos tipos de contenido, pero a menudo no puede generar tipos específicos de resultados con la precisión o fidelidad deseadas. Para ello, el modelo debe ajustarse a una tarea específica de generación de contenido. Esto se puede hacer de varias maneras.

- **Ajuste**

El ajuste implica alimentar el modelo con datos etiquetados específicos de la aplicación de generación de contenido: preguntas o instrucciones que es probable que reciba la aplicación y las respuestas correctas correspondientes en el formato deseado. Por ejemplo, si un equipo de desarrollo intenta crear un chatbot de atención al cliente, crearía cientos o miles de documentos que contienen preguntas de atención al cliente etiquetadas y respuestas correctas; posteriormente, alimentaría el modelo con esos documentos.

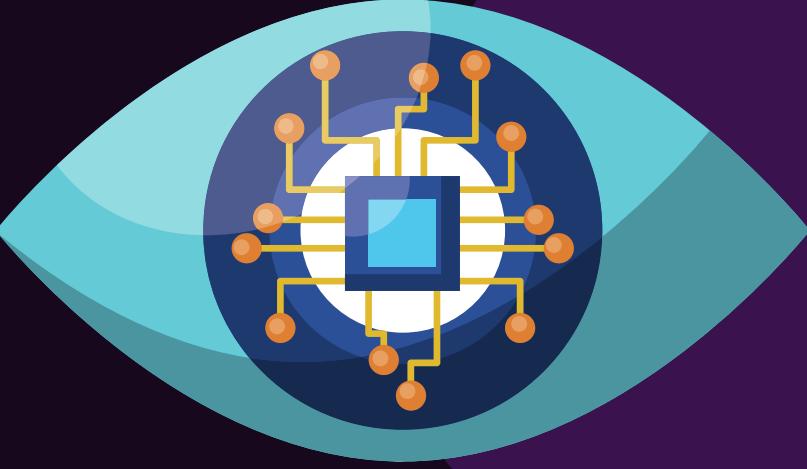
- **Aprendizaje por refuerzo con retroalimentación humana (RLHF, por sus siglas en inglés)**

En las RLHF, los usuarios humanos responden al contenido generado con evaluaciones que el modelo puede utilizar para actualizar el modelo para una mayor precisión o relevancia. A menudo, en las RLHF hay personas que "asignan una calificación" a diferentes resultados en respuesta a la misma pregunta.

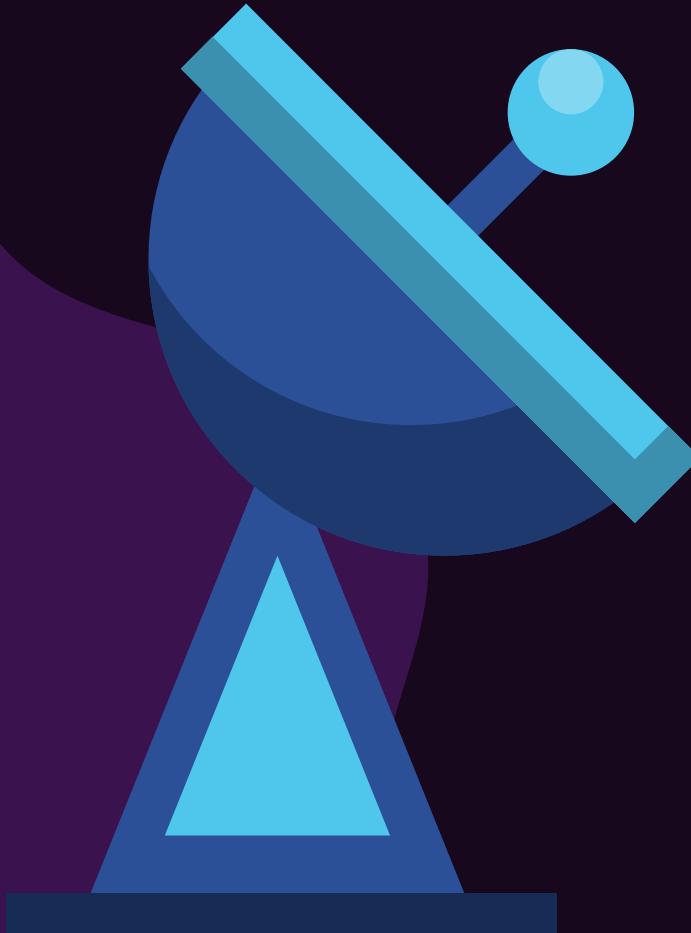
Generación, evaluación, más ajuste

Los desarrolladores y usuarios evalúan continuamente los resultados de sus aplicaciones de IA generativa y ajustan adicionalmente el modelo, incluso una vez a la semana, para aumentar su precisión o relevancia. (En cambio, el propio modelo fundacional se actualiza con mucha menos frecuencia, quizá cada año o cada 18 meses.)

Otra opción para mejorar el rendimiento de una aplicación de IA generativa es la generación aumentada de recuperación (RAG). La RAG consiste en un marco para ampliar el modelo fundacional y utilizar fuentes relevantes fuera de los datos de entrenamiento, para complementar y ajustar los parámetros o representaciones en el modelo original. La RAG puede garantizar que una aplicación de IA generativa siempre tenga acceso a la información más actualizada.



CONTENIDO QUE PUEDE CREAR LA IA



Contenido

Texto

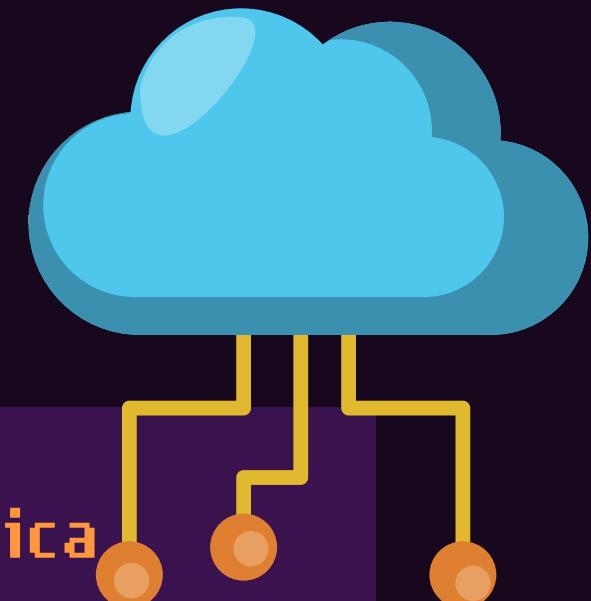
Los modelos generativos, especialmente aquellos basados en transformadores, pueden generar texto coherente y relevante para el contexto, desde instrucciones y documentación hasta folletos, correos electrónicos, textos para sitios web, blogs, artículos, informes, documentos e incluso escritura creativa. También pueden realizar tareas de escritura repetitivas o tediosas (redactar resúmenes de documentos o metadescpciones de páginas web), dejando tiempo a los redactores para trabajos más creativos y de mayor valor.

Imágenes y video

La generación de imágenes, como DALL-E, Midjourney y Stable Diffusion, puede crear imágenes realistas o arte original; también puede realizar transferencia de estilo, traducción de imagen a imagen y otras tareas de edición de imágenes o mejora de imágenes. Las herramientas de video de IA generativa emergente pueden crear animaciones a partir de indicaciones de texto, y pueden aplicar efectos especiales al video existente de manera más rápida y rentable que otros métodos.

Sonido, voz y música

Los modelos generativos pueden sintetizar voz y contenido de audio de sonido natural para usarlos en chatbots de IA habilitados por voz y asistentes digitales, narración de audiolibros y otras aplicaciones. La misma tecnología puede generar música original que imita la estructura y el sonido de las composiciones hechas por profesionales.



Contenido

Código de software

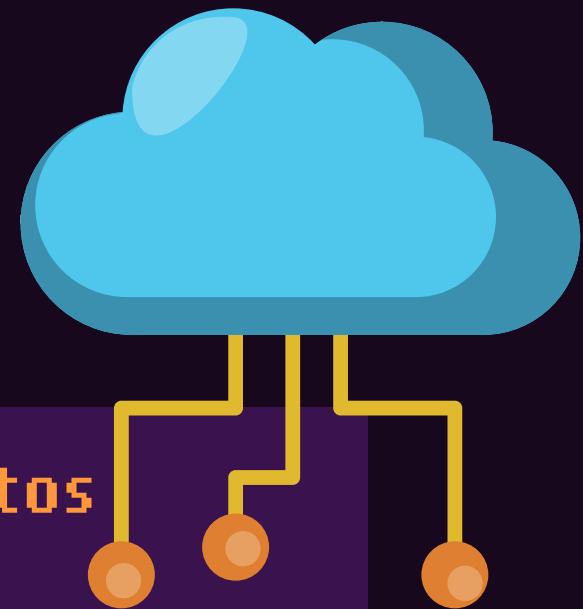
La IA generativa puede generar código original, autocompletar fragmentos de código, traducir entre lenguajes de programación y resumir la funcionalidad del código. Permite a los desarrolladores crear prototipos, refactorizar y eliminar errores en aplicaciones rápidamente, al tiempo que ofrece una interfaz de lenguaje natural para tareas de programación.

Diseño y arte

Los modelos de IA generativa pueden generar obras de arte y diseño únicos, así como ayudar en el diseño gráfico. Las aplicaciones incluyen la generación dinámica de entornos, personajes o avatares, además de efectos especiales para simulaciones virtuales y videojuegos.

Simulaciones y datos sintéticos.

Los modelos de IA generativa pueden entrenarse para generar datos sintéticos o estructuras sintéticas basadas en datos reales o sintéticos. Por ejemplo, la IA generativa se aplica en el descubrimiento de fármacos para generar estructuras moleculares con las propiedades deseadas, lo que ayuda al diseño de nuevos compuestos farmacéuticos.





BENEFICIOS



Creatividad mejorada

Las herramientas de generación de inteligencia artificial pueden inspirar creatividad a través de una lluvia de ideas automatizada, dando lugar a múltiples versiones novedosas de contenido. Estas variaciones también pueden servir como puntos de partida o referencias que ayuden a escritores, artistas, diseñadores y otros creadores a superar los bloqueos creativos.

Toma de decisiones mejorada (y más rápida)

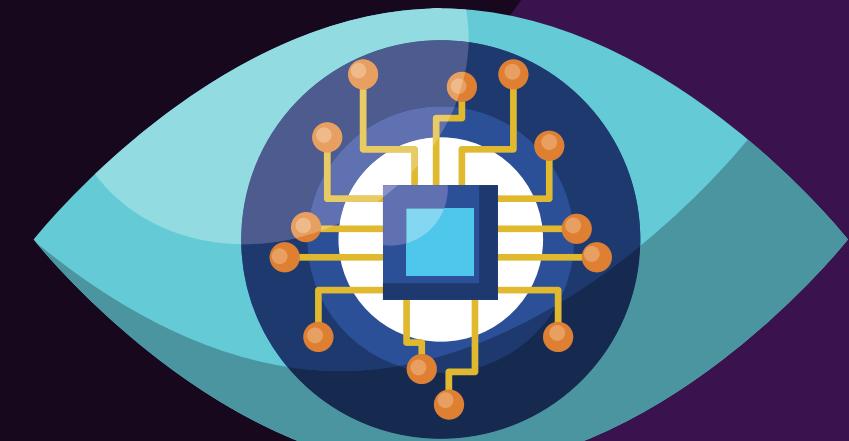
La IA generativa destaca en el análisis de grandes conjuntos de datos, la identificación de patrones y la extracción de insights significativos, y luego en la generación de hipótesis y recomendaciones basadas en esos conocimientos para apoyar a ejecutivos, analistas, investigadores y otros profesionales en la toma de decisiones más inteligentes y basadas en datos.

Personalización dinámica

En aplicaciones como los sistemas de recomendación y la creación de contenido, la IA generativa puede analizar las preferencias y el historial del usuario y generar contenido personalizado en tiempo real, lo que lleva a una experiencia de usuario más personalizada y atractiva.

Disponibilidad constante

La IA generativa opera continuamente sin fatiga, proporcionando disponibilidad las 24 horas para tareas como chatbots de atención al cliente y respuestas automatizadas.



**DESAFIOS,
LIMITACIONES Y
RIESGOS**



Desafíos, limitaciones y riesgos

"Alucinaciones" y otros resultados inexactos

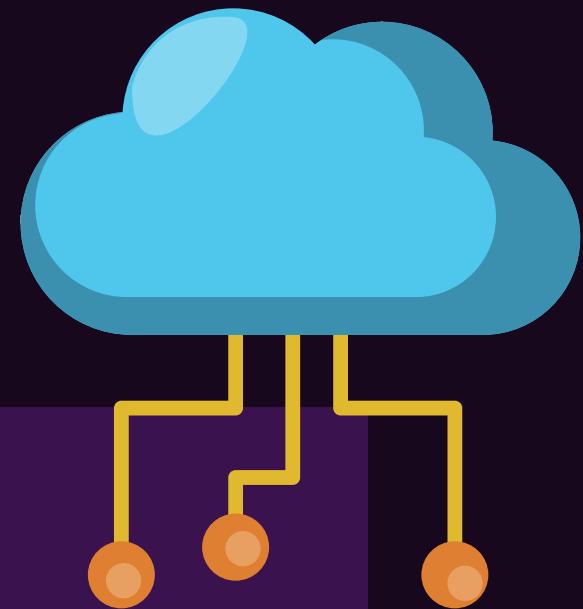
Una alucinación de IA es un resultado de IA generativa que no tiene sentido o es totalmente inexacta, pero, con demasiada frecuencia, parece completamente plausible.

Resultados inconsistentes

Debido a la naturaleza variacional o probabilística de los modelos de IA generativa, las mismas entradas pueden dar lugar a resultados un poco o muy diferentes. Esto puede ser indeseable en ciertas aplicaciones, como los chatbots de atención al cliente, donde se esperan o desean resultados coherentes. A través de la ingeniería rápida, al ajustar o redactar de forma iterativa las instrucciones, los usuarios pueden llegar a instrucciones que ofrezcan de forma coherente los resultados que desean de sus aplicaciones de IA generativa.

Sesgo

Los modelos generativos pueden aprender los sesgos sociales presentes en los datos de entrenamiento, o en los datos etiquetados, las fuentes de datos externas o los evaluadores humanos utilizados para ajustar el modelo, y generar contenido sesgado, injusto u ofensivo como resultado. Para evitar resultados sesgados de sus modelos, los desarrolladores deben garantizar que los datos de entrenamiento sean diversos, establecer pautas para evitar sesgos durante el entrenamiento y el ajuste, así como evaluar continuamente los resultados del modelo para determinar el sesgo y la precisión.



Desafíos, limitaciones y riesgos

Falta de explicabilidad y métricas

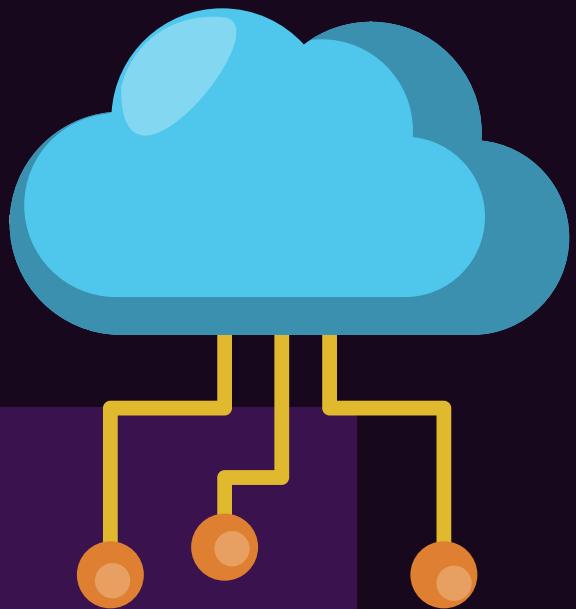
Muchos modelos de IA generativa son modelos de "caja negra", lo que significa que puede ser desafiante o imposible entender sus procesos de toma de decisiones; incluso los ingenieros o científicos de datos que crean el algoritmo subyacente pueden entender o explicar qué es exactamente lo que está sucediendo dentro de este y cómo llega a un resultado específico.

Amenazas a la seguridad, la privacidad y la propiedad intelectual

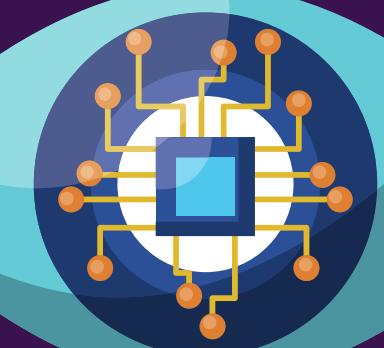
Los modelos de IA generativa pueden aprovecharse para generar correos electrónicos de phishing convincentes, identidades falsas u otro contenido malicioso que pueda engañar a los usuarios para que tomen medidas que comprometan la seguridad y la privacidad de los datos.

Deepfakes

Los deepfakes son imágenes, videos o audio generados o manipulados por IA y creados para convencer a las personas de que están viendo, mirando o escuchando a alguien hacer o decir algo que nunca hizo o dijo.



CASOS DE USO



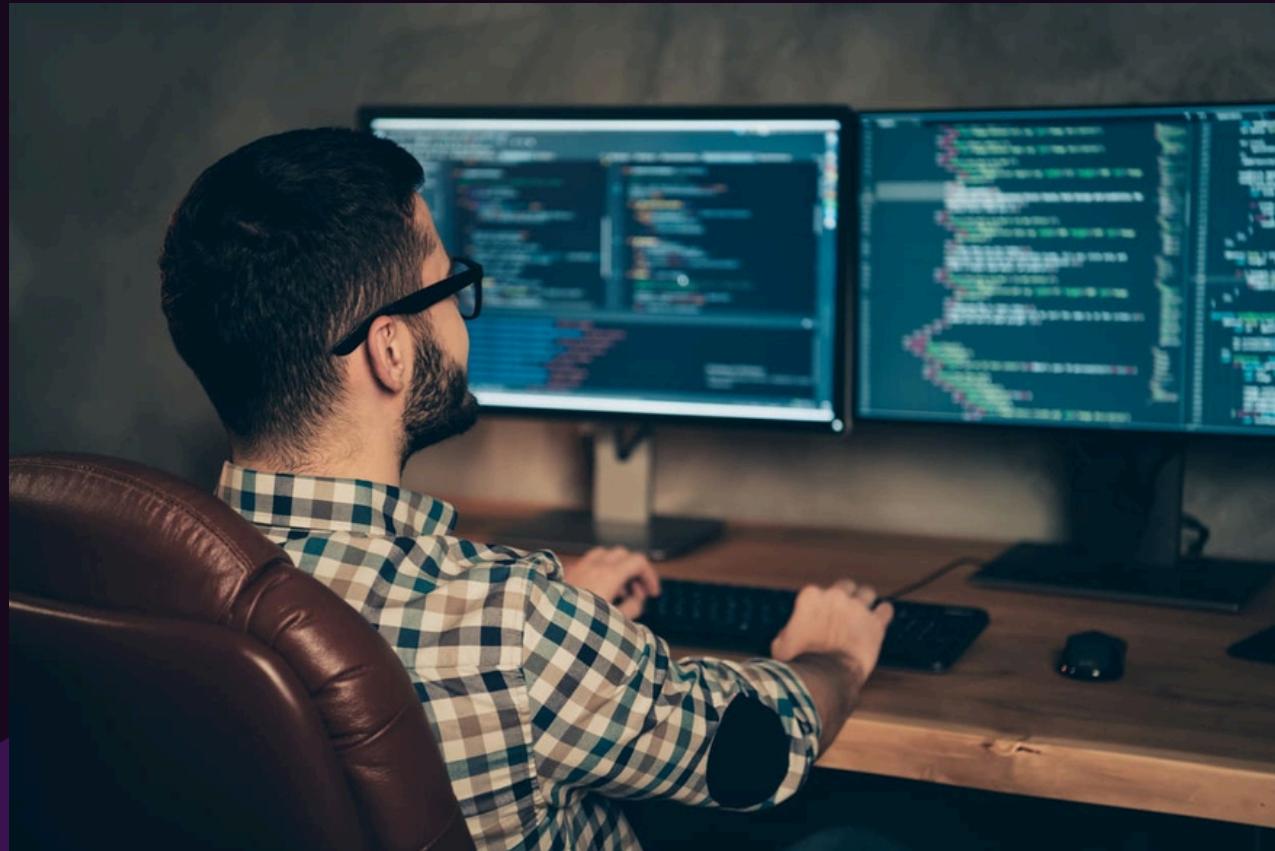
Experiencia del cliente

Las organizaciones de marketing pueden ahorrar tiempo y ampliar su producción de contenido mediante el uso de herramientas de IA generativa para redactar textos para blogs, páginas web, material publicitario, correos electrónicos y más. Pero las soluciones de IA generativa también pueden crear textos e imágenes de marketing muy personalizados en tiempo real en función de cuándo, dónde y a quién se entrega el anuncio.



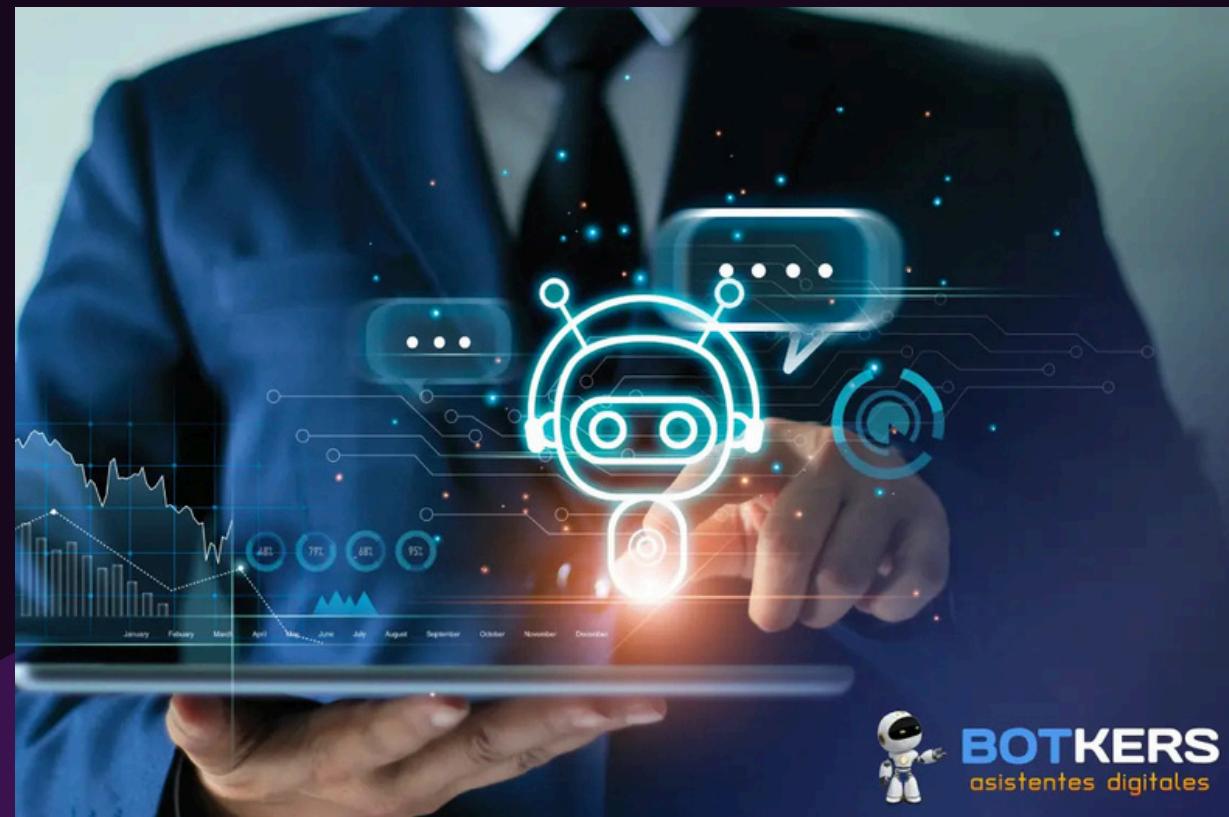
Desarrollo de software y modernización de aplicaciones

Las herramientas de generación de código pueden automatizar y acelerar el proceso de escritura de código nuevo. La generación de código también tiene el potencial de acelerar drásticamente la modernización de aplicaciones automatizando gran parte de la programación repetitiva necesaria para modernizar las aplicaciones heredadas para entornos de nube híbrida.



Trabajo digital

La IA generativa puede redactar o revisar rápidamente contratos, facturas, recibos y otros "trámites" digitales o físicos para que los empleados que la usan o gestionan puedan concentrarse en tareas de nivel superior. Esto puede acelerar los flujos de trabajo en prácticamente todas las áreas empresariales, incluidos recursos humanos, asuntos legales, adquisiciones y finanzas.



Ciencia, ingeniería e investigación

Los modelos de IA generativa pueden ayudar a los científicos e ingenieros a proponer soluciones novedosas a problemas complejos. Por ejemplo, en la atención médica, los modelos generativos se pueden aplicar para sintetizar imágenes médicas para entrenar y probar sistemas de imágenes médicas.

