

Análisis de las limitaciones de ChatGPT

José Luis Rodríguez y Alexandre Muñoz

October 2023

1 Imperfecciones del ChatGPT

ChatGPT es una implementación de la tecnología GPT, desarrollada por OpenAI, diseñada para mantener conversaciones de manera coherente y contextual en lenguaje natural. Este modelo de lenguaje se entrena previamente en una amplia variedad de datos textuales de la web y puede ser afinado para tareas específicas, lo que lo hace útil en aplicaciones como asistentes virtuales, atención al cliente automatizada y generación de contenido, mejorando la interacción entre humanos y máquinas en entornos de chat en línea.

Profundizando en las limitaciones actuales de ChatGPT destacamos las siguientes:

1. Falta de sentido común

Aunque puede generar respuestas similares a las humanas y tiene acceso a una gran cantidad de información, no posee sentido común de nivel humano, y el modelo carece también de los conocimientos previos que tenemos nosotros. Por lo que puede dar a veces respuestas sin sentido o inexactas.

2. Falta de inteligencia emocional

Aunque ChatGPT es competente en el procesamiento del lenguaje, puede tener dificultades para captar los sutiles matices de la comunicación humana. Por ejemplo, si un usuario utiliza el sarcasmo o el humor en su mensaje, ChatGPT puede no captar el significado deseado y proporcionar una respuesta inapropiada o irrelevante.

3. Limitaciones a la hora de gestionar varias tareas al mismo tiempo

El modelo funciona mejor cuando se le asigna una única tarea u objetivo. Si le pides a ChatGPT que realice varias tareas a la vez, le costará priorizarlas, lo que reducirá su eficacia y precisión. Además es más adecuado para generar contenidos breves.

4. Conocimiento limitado

Aunque pueden generar información útil, no tiene un conocimiento profundo de hechos o eventos específicos más allá de lo que han aprendido de

los datos de entrenamiento, lo que puede llevar a respuestas incorrectas o desactualizadas.

2 Incidente mencionado del abogado

La famosa noticia trata de que el abogado de un hombre que demandó a una aerolínea por daños personales utilizó ChatGPT para preparar una presentación, pero el bot de inteligencia artificial entregó casos falsos que el abogado presentó después ante el tribunal, lo que llevó a un juez a considerar sanciones mientras la comunidad jurídica lidia con uno de los primeros casos de "alucinaciones" de IA que hacen acto de presencia en los tribunales.

El problema es que aquella retahíla detallada de decisiones judiciales, citas y nombres tenía más de ciencia ficción que de informe legal contrastado. Cuando los abogados de Avianca se pusieron a buscar las referencias que enumeraba el escrito de Schwartz fueron incapaces de localizarlas. Ni siquiera el magistrado pudo, normal ya que el contenido era impreciso o incluso inventado.

Más que su falta de veracidad lo llamativo del informe es quién había sido el autor. El responsable de semejante despliegue de ficción judicial no fue Schwartz —no al menos de forma directa— sino la herramienta de la que echó mano para su informe: ChatGPT. Como ha acabado reconociendo el abogado, recurrió al motor de OpenAI para agilizar su investigación legal, una decisión que, visto lo ocurrido, no ha resultado demasiado atinada. "Se ha revelado como poco fiable", admite.

2.1 Implicaciones

Algunas de las implicaciones negativas incluyen:

1. **Errores y fallos de precisión:** Los modelos de lenguaje como ChatGPT pueden cometer errores y proporcionar respuestas inexactas o inadecuadas en situaciones legales complejas. Esto puede llevar a asesoramiento legal incorrecto o a la toma de decisiones incorrectas en casos legales.
2. **Responsabilidad ética y legal:** Si un abogado o un bufete de abogados utiliza un sistema de IA como ChatGPT para proporcionar asesoramiento legal y se cometen errores, esto podría dar lugar a cuestiones éticas y legales de responsabilidad y negligencia.
3. **Privacidad y confidencialidad:** La conversación entre un cliente y un abogado a menudo involucra información confidencial y sensible. El uso de sistemas de IA plantea preocupaciones sobre la privacidad y la seguridad de los datos, especialmente si la información se almacena o se utiliza de manera inapropiada.
4. **Pérdida de empleo:** La automatización de tareas legales básicas mediante IA podría reducir la demanda de empleados en roles paralegales y

administrativos, lo que podría tener un impacto negativo en el empleo en la industria legal.

5. **Dependencia tecnológica:** La dependencia excesiva de la tecnología puede hacer que los abogados confíen en sistemas de IA sin el debido escrutinio o supervisión, lo que podría socavar la toma de decisiones informadas y la interpretación legal.
6. **Despersonalización:** La automatización a través de sistemas de IA podría llevar a una experiencia legal despersonalizada para los clientes, lo que podría afectar negativamente la relación cliente-abogado y la calidad del servicio.

El uso de ChatGPT en abogacía "puede" mejorar la eficiencia y la accesibilidad de servicios legales, pero también plantea desafíos importantes relacionados con la ética, la precisión y la privacidad que deben ser abordados cuidadosamente. La supervisión y orientación de profesionales del derecho siguen siendo esenciales para garantizar que se cumplan los estándares éticos y legales en todas las interacciones con clientes y en la toma de decisiones legales.

3 Grafos de Conocimiento

Los grafos de conocimiento son estructuras de datos que representan información de manera interconectada, donde los nodos representan conceptos o entidades, y las aristas (conexiones) representan las relaciones entre estos conceptos. Utilizar grafos de conocimiento en conjunto con ChatGPT puede mejorar significativamente la capacidad de comprensión y generación de respuestas del modelo.

Una ventaja que tienen los grafos sobre los sistemas de Inteligencia Artificial es que ésta necesita un entrenamiento continuo de los modelos, incluyendo nueva información. Sin embargo, los grafos de conocimiento son capaces de encontrar relaciones complejas sin este entrenamiento constante.

¿Cómo los grafos proporcionan contexto estructurado a ChatGPT?

1. **Integración con ChatGPT:** El grafo de conocimiento se integra con ChatGPT como una fuente adicional de información. El modelo se entrena para consultar y utilizar este grafo en sus respuestas y tareas de comprensión. Puede hacerlo mediante consultas a la base de conocimiento durante la generación de respuestas o al responder preguntas basadas en el contenido del grafo.
2. **Mejora de la Comprensión Contextual:** Cuando un usuario hace una pregunta o proporciona un contexto, ChatGPT puede utilizar el grafo de conocimiento para obtener información relevante y mejorar su comprensión contextual. Esto permite respuestas más precisas y enriquecidas, ya que el modelo puede acceder a datos específicos en el grafo para respaldar sus respuestas.

3. **Resolución de Ambigüedad:** Los grafos de conocimiento ayudan a resolver ambigüedades en las consultas o preguntas. Por ejemplo, si un usuario pregunta "¿Quién es Juan?" y hay múltiples personas con ese nombre, el grafo de conocimiento puede proporcionar detalles contextuales para identificar al Juan relevante.
4. **Actualización Continua:** El grafo de conocimiento no es estático y debe actualizarse periódicamente para reflejar nuevos datos y cambios en el conocimiento. Esto asegura que el modelo tenga acceso a información actualizada y precisa.

En resumen, la integración de grafos de conocimiento en ChatGPT fortalece su capacidad para comprender, contextualizar y generar respuestas precisas y coherentes. Esta combinación de procesamiento de lenguaje natural y estructuras de conocimiento interconectadas tiene el potencial de mejorar significativamente las capacidades de chatbots y asistentes virtuales en una variedad de aplicaciones, desde la atención al cliente hasta la asesoría legal y la educación.