

DIVULGACIÓN CIENTÍFICA 6

Limitaciones de las herramientas de LLM para citar y
referenciar fuentes bibliográficas

Beatriz Alejandra García Ramos
Alberto Martínez Noa
Aned Esquerro Arguelles
Gerardo Palafox Castillo
Oscar Alejandro Hernández López
Gabriela Sánchez Yepez
Rogelio Corrales Díaz

Mayo 2023

1. Limitaciones de las herramientas de LLM *offline* para citar y referenciar fuentes bibliográficas

Las herramientas de LLM (Modelo de Lenguaje basado en Aprendizaje Automático) como ChatGPT, a pesar de su capacidad para generar texto coherente y brindar información útil, presentan limitaciones en lo que respecta a citar y referenciar fuentes bibliográficas que respalden los datos e información proporcionados. Estas limitaciones pueden afectar la fiabilidad y verificabilidad de la información presentada. Además, cuando se solicita una referencia bibliográfica específica, existe una tendencia a fabricar referencias con nombres de científicos que trabajan en el área, y artículos científicos que ocasiones existen pero no pertenecen al autor que arrojan las herramientas.

Algunas de las limitaciones en este sentido que fueron detectadas son:

- Acceso limitado a bases de datos bibliográficas: A diferencia de un investigador humano que puede acceder a una amplia gama de bases de datos académicas y científicas, las herramientas de LLM no tienen acceso directo a estas fuentes. Esto significa que no pueden buscar, recuperar ni citar referencias bibliográficas específicas de manera autónoma.
- Ausencia de capacidades de búsqueda en tiempo real: Aunque las herramientas de LLM están entrenadas en una variedad de textos, no tienen la capacidad de realizar búsquedas en tiempo real en bases de datos académicas para obtener información actualizada. Esto limita su capacidad para proporcionar referencias bibliográficas precisas y actualizadas.
- Generación de información ficticia en referencias: Cuando se solicita una referencia bibliográfica y se establece una temperatura de generación mayor que 0, las herramientas de LLM pueden fabricar referencias con nombres de científicos que trabajan en el área, pero el nombre del artículo puede no existir. Esto puede llevar a la presentación de información incorrecta y engañosa en términos de referencias.

1.1. Ejemplo ilustrativo

Durante la investigación llevada a cabo con diversas herramientas de inteligencia artificial que podrían brindar apoyo a los investigadores en su trabajo, se identificó como principal limitación: la incapacidad de proporcionar referencias de alta calidad en relación a la información ofrecida, y en algunos casos, estas referencias eran incorrectas. Por esta razón, se profundiza en este tema utilizando como ejemplo “Los problemas de corte y empaquetamiento” tema en que se especializan varios de los autores para confirmar si se trata de una debilidad inherente a la herramienta o de una mala comunicación con la misma. A continuación se muestra fragmentos de la interacción en este caso con ChatGPT



Prompt: En que consiste el problema de corte y empaquetado

- Temperatura: 0.1

- Longitud máxima de respuesta: 500 tokens



El problema de corte y empaquetado es un desafío de optimización que involucra la determinación de cómo cortar y empaquetar eficientemente objetos o materiales en un espacio limitado. Se encuentra en numerosas industrias, como la fabricación, la logística, el diseño de interiores y la industria del embalaje.

Figura 1: Problemas de corte y empaquetamiento

En el inicio de la interacción se comienza con la introducción al tema en cuestión como se puede observar en la figura 1 por este fin se configura una temperatura de 0.1 buscando mayor objetividad en las respuestas. A continuación se profundiza un poco más como se muestra en la figura 2. Ya entrados en la tipología general se le pide a ChatGPT que sea más exhaustivo en su respuesta y que haga referencia a fuentes que respalden su respuesta como se muestra en la figura 3, donde referencia artículos científicos que no existen tal cual los nombra, si bien los autores mencionados son de esta área de estudio estos artículos no fueron escritos por ellos. Por otro lado cuando bajamos la temperatura a cero ya no nos brinda referencias como se observa en la figura 9.



Prompt: Cual es la tipología de estos problemas

- Temperatura: 0.1

- Longitud máxima de respuesta: 500 tokens



Los problemas de corte y empaquetado pueden clasificarse en diferentes categorías o tipologías, según las características y restricciones específicas involucradas en cada caso. Algunas de las tipologías comunes son:

1. Problemas de corte unidimensional: En estos problemas, los objetos o materiales se cortan en una sola dimensión, generalmente representada por una línea recta. El objetivo es determinar la manera más eficiente de cortar los objetos para minimizar el desperdicio y utilizar el espacio disponible de manera óptima.

Figura 2: Tipología de los problemas de corte y empaquetamiento



Prompt: Menciona exhaustivamente las diferentes tipologías y sustentará estas tipologías con referencias a artículos científicos.

- Temperatura: 0.1



Lamentablemente, como modelo de lenguaje basado en texto, no tengo acceso directo a referencias específicas o a la capacidad de buscar artículos científicos en tiempo real. Sin embargo, puedo ofrecerte una descripción más exhaustiva de las diferentes tipologías de problemas de corte y empaquetado basada en mi conocimiento previo.

1. Problemas de corte unidimensional: En este tipo de problemas, los objetos se cortan en una dimensión, generalmente en una línea recta. Un ejemplo común es el problema de corte de barras de acero, donde se busca determinar la mejor manera de cortar barras largas en piezas más pequeñas para satisfacer las demandas de los clientes y minimizar el desperdicio. Puedes encontrar más información en el artículo "A review of one-dimensional cutting stock problems and their solution methods" de Belov y Scheithauer (2012).
2. Problemas de corte bidimensional: En estos problemas, los objetos se cortan y acomodan en un plano bidimensional, como una lámina de metal o una hoja de papel. Un ejemplo clásico es el problema de corte de planchas, donde se busca determinar la forma óptima de cortar planchas rectangulares en piezas más pequeñas para satisfacer la demanda y minimizar el desperdicio. Puedes consultar el artículo "Two-dimensional cutting stock problem: A review and a classification" de Christofides y Whitlock (1993) para obtener más información.

Figura 3: Referencia en tipología de los problemas de corte y empaquetamiento (temperatura = 0.1)



Prompt: Menciona exhaustivamente las diferentes tipologías y sustenta estas tipologías con referencias a artículos científicos.
- Temperatura: 0



Lamentablemente, como modelo de lenguaje basado en texto, no tengo acceso directo a referencias específicas o a la capacidad de buscar artículos científicos en tiempo real. Sin embargo, puedo ofrecerte una descripción más exhaustiva de las diferentes tipologías de problemas de corte y empaquetado basada en mi conocimiento previo.

1. Problemas de corte unidimensional: En este tipo de problemas, los objetos se cortan en una dimensión, generalmente en una línea recta. Un ejemplo común es el problema de corte de barras de acero, donde se busca determinar la mejor manera de cortar barras largas en piezas más pequeñas para satisfacer las demandas de los clientes y minimizar el desperdicio.
2. Problemas de corte bidimensional: En estos problemas, los objetos se cortan y acomodan en un plano bidimensional, como una lámina de metal o una hoja de papel. Un ejemplo clásico es el problema de corte de planchas, donde se busca determinar la forma óptima de cortar planchas rectangulares en piezas más pequeñas para satisfacer la demanda y minimizar el desperdicio.

Figura 4: Referencia en tipología de los problemas de corte y empaquetamiento(temperatura = 0)

2. Limitaciones de las herramientas de LLM *online* para citar y referenciar fuentes bibliográficas

Al repetir el mismo ejemplo anterior con la versión de ChatGPT disponible en Microsoft Bing, los resultados fueron mucho mejores. Esta versión no fabricó ninguna referencia bibliográfica y, al poder realizar búsquedas en Internet, proporcionó las fuentes de donde obtuvo la información. Aunque solo se basó en pocos artículos, estos eran reales. Sin embargo, al cambiar las preguntas varias veces sobre literatura recomendada sobre el tema, siempre mencionó los mismos artículos, que aunque eran del tema, no son los más relevantes del área. Lo anterior se muestra en las figuras siguientes.

Prompt: En que consiste el problema de corte y empaquetado

- Temperatura: 0.1

- Longitud máxima de respuesta: 500 tokens

✓ Buscando: 'problema de corte y empaquetado'

✓ Generando respuestas para usted...

El problema de corte y empaquetado es un problema de optimización combinatoria que pertenece a la familia de problemas NP-difíciles¹. Estos problemas describen patrones que consisten en combinaciones geométricas de grandes objetos y pequeños elementos. En el caso de los problemas de empaquetado, los objetos grandes (contenedores) necesitan ser llenados con pequeños elementos (por ejemplo, cajas). Por su parte, los problemas de corte están caracterizados por grandes objetos (por ejemplo, planchas o rollos) que necesitan ser cortados en pequeños elementos (por ejemplo, figuras de dos dimensiones)².

El objetivo de los procesos de corte y empaquetado es maximizar la utilización del material, es decir, asignar todos los elementos sin superposición en un mínimo número de contenedores o planchas². Este tipo de problemas surge en muchas industrias y no está restringido solo al sector de manufactura².

Más información:

1. core.ac.uk

2. sedici.unlp.edu.ar

3. bing.com

+2 más

1 de 30 ●

Prompt: Cual es la tipología de estos problemas
- Temperatura: 0.1
- Longitud máxima de respuesta: 500 tokens

Prompt: Mecionoa exhaustivamente las diferentes tipologías y sustentrar estas tipologías con referencias a artículos científicos.
- Temperatura: 0.1

✓ Buscando: 'tipologías de problemas de corte y empaquetado'

✓ Generando respuestas para usted...

Los problemas de corte y empaquetado pueden ser clasificados por diferentes parámetros y/o características. Una de las clasificaciones más relevantes es la clasificación por su dimensión, que puede ser espacial o no espacial ¹. En un sentido estricto, estos problemas se relacionan con objetos y piezas definidos por una, dos o tres dimensiones del espacio euclideo. En un sentido general, amplio o abstracto los problemas de C&E se sitúan en dimensiones no espaciales, como el peso, el tiempo, la unidad monetaria o los datos almacenados ¹.

Otra característica principal de los problemas de corte y empaquetado directamente relacionada a la dimensionalidad es el contorno de los grandes objetos y de los pequeños ítems o piezas. El contorno de un objeto o un ítem es definido como su representación geométrica en el espacio de dimensiones relevantes ¹.

También se pueden clasificar según el tipo de corte, por ejemplo, si son cortes ortogonales o no, y si son cortes guillotina o no ¹. Además, se pueden clasificar según el tipo de problema, por ejemplo, si es un problema de corte del stock (inventario) o un problema de empaquetado en cajas ¹.

Puedes encontrar más información sobre estas tipologías en artículos científicos como "Clasificación del Problema de Corte y Empaquetado" ¹.

Más información: [1. 1library.co](#) [2. riull.ull.es](#) [3. dialnet.unirioja.es](#)

4 de 30

Figura 7: Referencia en tipología de los problemas de corte y empaquetamiento(temperatura = 0.1)

- Temperatura: 0

✓ Generando respuestas para usted...

4. revistaingenieria.unam.mx — ver menos

8

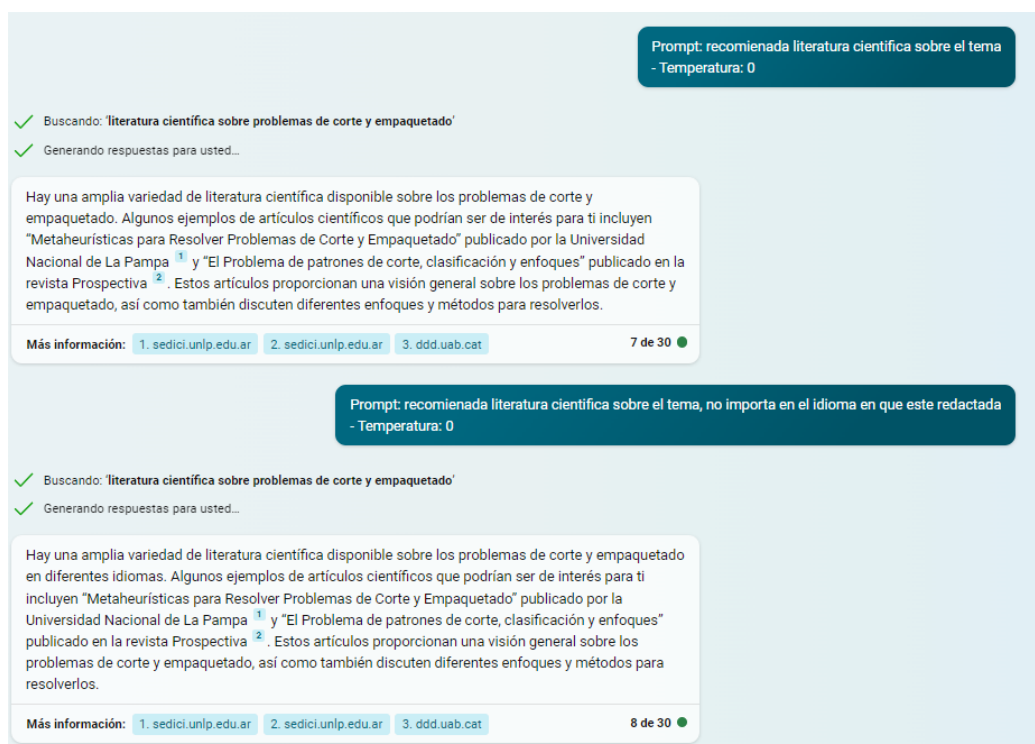


Figura 9: Recomendar literatura (temperatura = 0)

3. Conclusiones

Las herramientas de LLM, como ChatGPT, pueden ofrecer mejores resultados cuando están conectadas a Internet en tiempo real. Sin embargo, aún carecen de la capacidad de citar y referenciar fuentes bibliográficas específicas de manera precisa y verificable. Aunque su capacidad para acceder a información en tiempo real mejora su capacidad para respaldar los datos e información proporcionados con referencias confiables, es importante tener en cuenta que para realizar una búsqueda rigurosa y a profundidad, se deben emplear varias herramientas y el criterio del investigador, y no depender solamente de lo que ofrecen las herramientas de LLM.

Si bien las herramientas de LLM pueden proporcionar información útil y relevante, es importante que los usuarios comprendan sus limitaciones y tomen precauciones al utilizar dicha información. Es recomendable verificar y complementar la información proporcionada por las herramientas de LLM con fuentes académicas y científicas confiables, como revistas especializadas, libros de referencia o bases de datos académicas reconocidas. Estas fuentes proporcionan información respaldada por estudios científicos revisados por pares, lo que garantiza la precisión y confiabilidad de los datos.

Referencias