TAREA 2

1-En la entrevista, Yann comenta sobre el fen ómeno manifestado por modelos de lenguaje natural conocido como alucinaci ón. ¿A qu é se refiere con este concepto? Explica en tus propias palabras.

En el contexto de modelos de lenguaje natural autoregresivos, como los LLMs, la alucinación se refiere a la producción de palabras o tokens basada en una probabilidad, la cual puede alejar la generación de texto de un conjunto de respuestas razonables. Si consideramos que estos errores son independientes a lo largo de la secuencia de tokens, lo que sucede es que con cada token generado, la probabilidad de mantenerse dentro de un conjunto de respuestas correctas disminuye, y esta disminución es exponencial. Este fenómeno se conoce como alucinación. Por tanto, la probabilidad de que una respuesta sea sin sentido aumenta exponencialmente con la cantidad de tokens generados

2- Minutos mas tarde, los locutores comienzan a discutir respecto a la capacidad de razonamiento de los modelos de lenguaje. Yann argumenta que los modelos de lenguaje actuales NO son capaces de desempen ãr razonamiento logico debido a la manera en que son entrenados. ¿Estas de acuerdo con esta opinion? Justifica brevemente. ¿Como describirıas la capacidad de razonar en tus propias palabras?

Estoy de acuerdo con la afirmación de Yann LeCun de que los modelos de lenguaje actualmente no pueden llevar a cabo razonamiento lógico en el sentido profundo y complejo que asociamos con el razonamiento humano. Los LLMs, como explica LeCun, aplican una cantidad constante de computación para cada token que generan, sin considerar la complejidad de la tarea. En consecuencia, no pueden dedicar más recursos a problemas más complejos, como lo haríamos nosotros al enfrentarnos a desafíos que requieren un pensamiento más profundo y analítico.

Esto se alinea con la diferencia entre lo que en psicología se denomina procesamiento del "Sistema 1" y del "Sistema 2". El "Sistema 1" se refiere a nuestro razonamiento rápido y instintivo, mientras que el "Sistema 2" es nuestro razonamiento lento, deliberativo y lógico. Los modelos actuales de LLM operan más en línea con el "Sistema 1", generando respuestas rápidas y basadas en patrones dentro de los datos en los que fueron entrenados, pero sin la capacidad de reflexionar o razonar de manera consciente y metódica, lo que sería característico del "Sistema 2".

Además, el desarrollo de sistemas de diálogo que puedan "pensar" antes de responder sigue siendo un área de investigación activa y distinta de la operación típica de un LLM autoregresivo. La comprensión de los LLMs sobre la adecuación de una respuesta a una pregunta es una medida de ajuste o correspondencia, no una señal de razonamiento real. Por tanto, aunque pueden seleccionar respuestas que parezcan adecuadas, no tienen la capacidad de razonamiento lógico que se da por inferencia o evaluación consciente

La capacidad de razonar para mi es la habilidad de llegar a razonamiento nuevo y de una forma más compleja comparado con el inicio. Es una acción que nos permitió de avanzar como especie humana y ha traído hombre a hacer innovación científica y filosófica.

La capacidad de razonar, en mi concepción, es la facultad cognitiva que permite a los seres humanos procesar información de manera lógica para formular juicios, deducir conclusiones y tomar decisiones basadas en evidencia y conocimiento previo. Es un proceso mental que involucra evaluar argumentos, reconocer patrones, utilizar el pensamiento crítico y aplicar la creatividad para resolver problemas nuevos y complejos.
El razonamiento nos distingue como especie, ya que nos permite aprender de nuestras experiencias, adaptarnos a cambios, innovar y construir complejas estructuras de conocimiento que se reflejan en todas las esferas de nuestra vida, desde la ciencia hasta la filosofía. Además, el razonamiento no solo está limitado a conclusiones basadas en hechos concretos, sino que también abarca la capacidad de considerar posibilidades, hipótesis y conceptos abstractos que no son inmediatamente verificables. En esencia, razonar es tejer una red de ideas que se sostienen y explican mutuamente dentro de un marco lógico y coherente.

3- Contrario a los modelos de lenguaje, sistemas l ógicos como Clingo nos permiten extraer conclusiones l ógicas a partir de un problema bien definido con una serie de reglas, habilidad que se ha demostrado compleja para los grandes modelos de lenguaje. ¿A que crees que se debe esto? Haz referencia a la manera en que son entrenados y el concepto de alucinaci ón.

Los sistemas lógicos como Clingo permiten extraer conclusiones lógicas de problemas bien definidos mediante un conjunto de reglas lógicas. Este enfoque contrasta con los grandes modelos de lenguaje (LLMs), que generan lenguaje y respuestas basadas en patrones aprendidos de grandes conjuntos de datos de texto.

La dificultad de los LLMs para realizar razonamiento lógico como el de sistemas basados en reglas se debe principalmente a la forma en que están entrenados. Los LLMs aprenden de ejemplos masivos sin una comprensión inherente de las reglas lógicas o estructuras subyacentes que gobiernan esos ejemplos. No pueden razonar a priori o deducir reglas de manera lógica porque su entrenamiento se basa en correlaciones estadísticas en lugar de relaciones lógicas causales.

Además, el concepto de alucinación en LLMs, donde el modelo puede generar respuestas fluidas pero incorrectas o sin sentido, también es un reflejo de esta limitación. Los LLMs, al producir un token a la vez y seguir una distribución de probabilidad, pueden derivar rápidamente en respuestas que se alejan de lo lógicamente coherente, especialmente con secuencias largas. Esto es un claro contraste con sistemas como Clingo, que pueden garantizar la coherencia lógica dentro de las limitaciones de su conjunto de reglas bien definido.

En el caso de Clingo, que es una herramienta para la programación lógica basada en el paradigma de Answer Set Programming (ASP), las conclusiones lógicas se obtienen a partir de reglas claramente especificadas y restricciones de cardinalidad que aseguran la validez de las soluciones generadas dentro de un espacio de problema bien definido.

4-Propone una idea (no debe ser una implementaci ón concisa, basta con un concepto ambiguo) de como fomentar el razonamiento l ógico en modelos de lenguaje para obtener conclusiones similares a las que genera Clingo pero de manera autom ática.

Para fomentar el razonamiento lógico en modelos de lenguaje y lograr conclusionessimilares a las de Clingo de manera automática, se podría considerar una integración híbrida que combine la capacidad de aprendizaje profundo de los LLMs con la precisión de los sistemas de programación lógica. La idea sería implementar un marco de trabajo donde los LLMs no solo generen lenguaje natural, sino que también aprendan a reconocer y aplicar reglas lógicas durante su entrenamiento.

Una estrategia sería entrenar a los LLMs con un conjunto de datos anotados que incluyan ejemplos de razonamiento lógico formal y su correspondencia en lenguaje natural. De esta manera, el modelo podría aprender a identificar patrones lógicos en el texto y aplicar reglas de inferencia similares a las que usa Clingo. Esto requeriría un diseño cuidadoso de los ejemplos de entrenamiento para cubrir un amplio espectro de estructuras lógicas y su representación textual.

Además, se podría emplear un sistema de retroalimentación donde las salidas del LLM sean verificadas y corregidas según las reglas lógicas por un sistema como Clingo. Con el tiempo, el modelo se ajustaría para mejorar su precisión en el razonamiento lógico. Esta retroalimentación podría automatizarse utilizando técnicas de aprendizaje por refuerzo, donde las "recompensas" se basen en la precisión lógica de las salidas del modelo.

Este enfoque híbrido permitiría a los LLMs no solo entender y generar texto, sino también razonar lógicamente de manera coherente y, eventualmente, producir conclusiones lógicas con una precisión que se acerque a la que ofrece Clingo, pero con la flexibilidad y riqueza del lenguaje natural que caracteriza a los modelos de lenguaje.