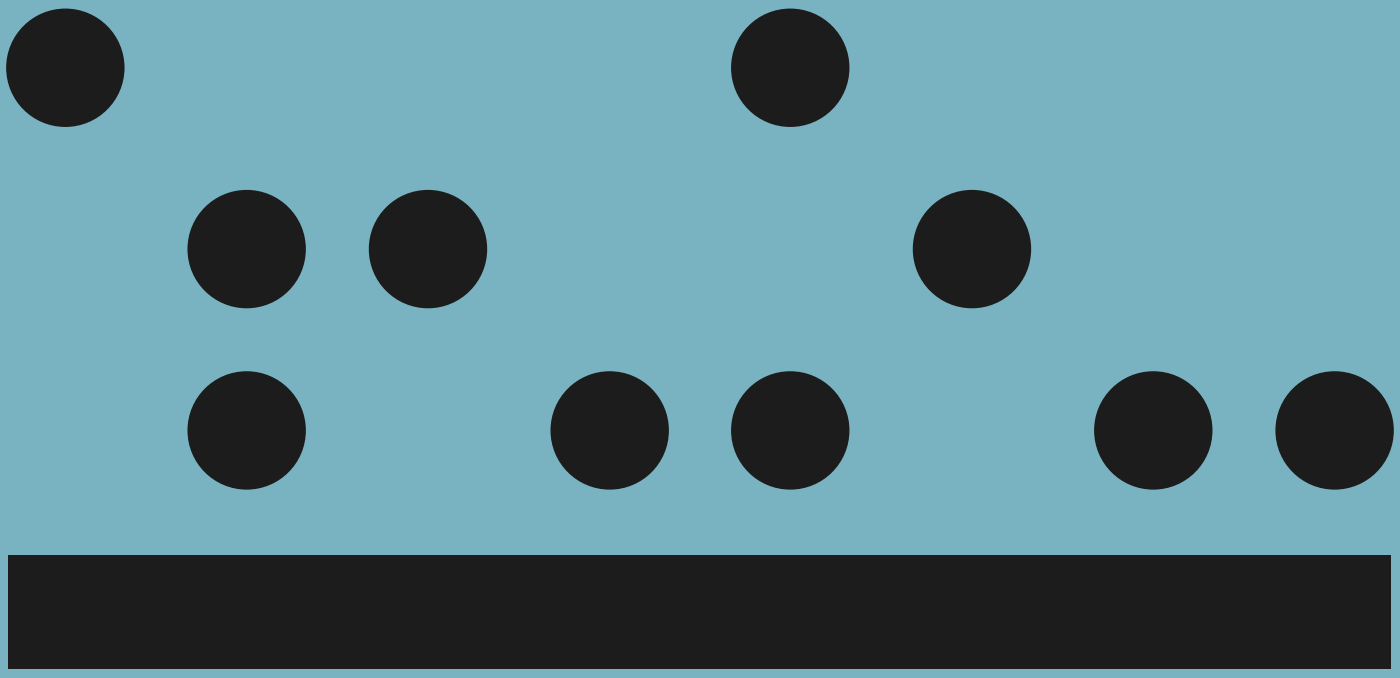
Sesgos algorítmicos y representación social en los modelos de lenguaje generativo (LLM)



Juan Manuel Ortiz de Zárate
Juan Manuel Dias
Alejandro Avenburg
Joan Imanol Gonzalez Quiroga

Datos

Abril 2024

fundar

Sesgos algorítmicos y representación social en los modelos de lenguaje generativo (LLM)

Juan Manuel Ortiz de Zárate
Juan Manuel Dias
Alejandro Avenburg
Joan Imanol Gonzalez Quiroga

- Generar riqueza
- Promover el bienestar
- Transformar el Estado

Sobre Fundar

Fundar es un centro de estudios y diseño de políticas públicas que promueve una agenda de desarrollo sustentable e inclusivo para la Argentina. Para enriquecer el debate público es necesario tener un debate interno: por ello lo promovemos en el proceso de elaboración de cualquiera de nuestros documentos. Confiamos en que cada trabajo que publicamos expresa algo de lo que deseamos proyectar y construir para nuestro país. Fundar no es un logo: es una firma.

Cita sugerida

Ortiz de Zárate, J. M.; Dias, J. M.; Avenburg, A. y Gonzalez Quiroga, J. I. (2024). Sesgos algorítmicos y representación social en los modelos de lenguaje generativo (LLM). Fundar.

Índice

Sesgos algorítmicos y representación social en los modelos de lenguaje generativo (LLM)

fundar

Ese diagrama inmóvil, con sus nueve mayúsculas repartidas en nueve cámaras y atadas por una estrella y unos polígonos, es ya una máquina de pensar. Es natural que su inventor —hombre, no lo olvidemos, del siglo XIII— la alimentara con materias que ahora nos parecen ingratas. Nosotros ya sabemos que los conceptos de bondad, de grandeza, de sabiduría, de poder y de gloria, son incapaces de engendrar una revelación apreciable. Nosotros (en el fondo, no menos ingenuos que Llull) la cargaríamos de un modo distinto. Sin duda, con las palabras Entropía, Tiempo, Electrones, Energía potencial, Cuarta dimensión, Relatividad, Protones y Einstein. O, también: Plusvalía, Proletariado, Capitalismo, Lucha de clases, Materialismo dialéctico, Engels.

Jorge Luis Borges
La máquina de pensar de Raimundo Lulio

Introducción

Los Modelos de Lenguaje con Aprendizaje Automático (LLM, por sus siglas en inglés) han surgido como herramientas sumamente poderosas, capaces de procesar y generar lenguaje de forma autónoma. Sin embargo, a medida que avanzan estos sistemas de inteligencia artificial (IA), diseñados para interpretar y responder preguntas de manera cada vez más sofisticada, surge una interrogante fundamental: ¿hasta qué punto estos algoritmos reflejan e, incluso, pueden amplificar los sesgos preexistentes en nuestra sociedad?

Este artículo profundiza en la problemática de los sesgos algorítmicos, explorando con detenimiento qué segmentos de la población podrían estar más presentes en las respuestas generadas por los LLM. Desde la falta de representación equitativa hasta la posibilidad de amplificar estereotipos arraigados, analizaremos cómo los sesgos algorítmicos plantean desafíos significativos que merecen una atención crítica.

Además, esta investigación se basa en un estudio realizado por un equipo de la Universidad de Stanford (Santurkar *et al.*, 2023). En él, se compararon las respuestas de los modelos de OpenAI con las opiniones de la sociedad estadounidense mediante encuestas de opinión pública. Los resultados arrojaron que las opiniones de estos modelos tendían a alinearse con las de un segmento específico de la sociedad estadounidense: aquellos con inclinaciones más liberales, niveles socioeconómicos acomodados, altos niveles educativos y ninguna afiliación religiosa.

La referencia nos proporciona un punto de partida para explorar en detalle la relación entre los sesgos algorítmicos y la representación de diversos segmentos poblacionales en las respuestas de los LLM.

Inteligencia artificial y sus diversas aplicaciones

La inteligencia artificial (IA) es un campo de la computación que se centra en desarrollar sistemas y programas informáticos capaces de resolver tareas que normalmente requieren de inteligencia humana. Dentro de esta categoría, existen varias ramas: una de ellas es la del aprendizaje automático (o *machine learning*), que destaca de otras por hacer uso de métodos estadísticos para identificar patrones en (una gran cantidad de) datos y hacer predicciones o tomar decisiones sin necesidad de programar explícitamente las soluciones que estos algoritmos encuentran.

El aprendizaje automático en sí es otra gran área de investigación, por lo que también lo podemos dividir en distintas partes y, dentro de ellas, podemos identificar al aprendizaje profundo (o *deep learning*): un conjunto de técnicas que apuntan a asimilar distintas representaciones de datos con arquitecturas computacionales particulares. Una de estas últimas es la de los *transformadores* (Vaswani, A. *et al*, 2017).

Todas estas tecnologías son ampliamente utilizadas para diversas aplicaciones, tales como asistentes inteligentes, motores de ajedrez, traductores automáticos, sistemas de recomendación de contenido, sistemas de visión por computadora, entre otros. Sin embargo, la categoría más pertinente a este artículo es la de "inteligencias artificiales generativas", que son aquellos algoritmos capaces de generar información a partir de algún estímulo externo y, en la actualidad, la tecnología más adoptada para este tipo de IAs es la de los transformadores.

En particular, los LLM son modelos de procesamiento de lenguaje natural que, en la actualidad, adoptan ampliamente la arquitectura de transformadores y, específicamente, la de transformadores generativos preentrenados o "GPT" (por sus siglas en inglés) (Radford *et al.*, 2018). Estos operan

Inteligencia artificial y sus diversas aplicaciones

con una gran cantidad de parámetros y han sido entrenados en una amplia variedad de datos de texto para ejecutar diversas tareas. En la actualidad, es comúnmente aceptado considerar a ChatGPT como un LLM debido a su habilidad para generar respuestas coherentes en conversaciones. Sin embargo, es importante destacar que ChatGPT no es en sí mismo un LLM, sino más bien una aplicación web que facilita la interacción con un LLM subyacente, tal como el modelo GPT-3 de OpenAI Este último es el encargado de procesar las conversaciones utilizando una variedad de técnicas.

Estos modelos están revolucionando la forma en que las máquinas interactúan y comprenden el lenguaje humano, los cuales se basan en arquitecturas de redes neuronales profundas y redes de atención. Hoy en día, están siendo ampliamente utilizados por empresas, investigadores y entusiastas del campo de la inteligencia artificial. Desde la generación de texto hasta la traducción automática, pasando por la respuesta a preguntas y la asistencia virtual, los LLM están encontrando aplicaciones en una variedad de dominios, tanto en el sector privado como en el público.

Los LLM (Large Language Model) son modelos de procesamiento de lenguaje natural que operan con una gran cantidad de parámetros y han sido entrenados en una amplia variedad de datos de texto para ejecutar diversas tareas. En la actualidad, es comúnmente aceptado considerar a ChatGPT como un LLM debido a su habilidad para generar respuestas coherentes en conversaciones. Sin embargo, es importante destacar que ChatGPT no es en sí mismo un LLM, sino más bien una aplicación web[1] que facilita la interacción con un LLM subyacente, tal como el modelo GPT-3 de OpenAI. Este último es el encargado de procesar las conversaciones utilizando una variedad de técnicas.

Estos modelos están revolucionando la forma en que las máquinas interactúan y comprenden el lenguaje humano, los cuales se basan en arquitecturas de redes neuronales profundas[2] y redes de atención. Hoy en día, están siendo ampliamente utilizados por empresas, investigadores y entusiastas del campo de la inteligencia artificial. Desde la generación de texto hasta la traducción automática, pasando por la respuesta a preguntas y la asistencia virtual, los LLM están encontrando aplicaciones en una variedad de dominios, tanto en el sector privado como en el público.

> Los LLM están revolucionando la forma en que las máquinas interactúan y comprenden el lenguaje humano, los cuales se basan en arquitecturas de redes neuronales profundas y redes de atención.

Por un lado, estas potentes herramientas han mejorado significativamente la experiencia del usuario y los servicios ofrecidos por las empresas. Han impulsado la eficiencia en la atención al cliente, la personalización de recomendaciones y la automatización de tareas. Por ejemplo, compañías de comercio electrónico están utilizando estos modelos para potenciar sus chatbots de atención al cliente, ofreciendo respuestas más rápidas y precisas a las consultas de los usuarios (Shrivastava, 2023). Además, los gobiernos y los Estados han comenzado a aprovechar los LLM para mejorar políticas públicas, permitiendo un enfoque más basado en evidencia y datos con el potencial de beneficiar a la sociedad en general. Un caso es el de Microsoft, que ha permitido a dependencias gubernamentales de Estados Unidos tales como el Departamento de Defensa, de Energía y la NASA la utilización de sus modelos (Metz, 2023).

Cómo funcionan los LLM

1 El término 'aplicación web' se utiliza de manera amplia en este contexto, ya que no sólo es accesible a través del navegador, sino que también permite la interacción mediante una API. Sin embargo, es importante destacar que la API no proporciona un acceso directo total al modelo subyacente.

2 Una arquitectura neuronal profunda se refiere a un tipo de red neuronal artificial que consta de múltiples capas de neuronas interconectadas. Cada capa procesa y extrae características específicas de los datos de entrada; las capas profundas realizan una representación cada vez más abstracta y compleja de la información. Estas arquitecturas se utilizan en el aprendizaje profundo (*deep learning*) para tareas como reconocimiento de patrones, procesamiento de lenguaje natural y visión por computadora, ya que tienen la capacidad de aprender y representar características sofisticadas de los datos.

Inteligencia artificial y sus diversas aplicaciones

Sin embargo, también se han presentado casos negativos de uso en contextos gubernamentales en distintos países por distorsión de la información o por la generación de discursos engañosos para promover agendas políticas o ideológicas. En un nivel más local, se han observado situaciones en las que un diputado cita un texto producido por ChatGPT como fuente autorizada, lo que plantea preocupaciones sobre la integridad de la información y el impacto de estas herramientas en la discusión pública (La Nación, 2023).

A pesar de estos desafíos, los LLM han abierto nuevas formas de relacionarse con el conocimiento en la sociedad en general. Han transformado la educación, permitiendo que los estudiantes utilicen estos modelos como "maestros particulares" (Rapallini, 2023), lo que facilita el acceso al conocimiento y la tutoría personalizada[3]. Incluso, hay docentes que han encontrado en los LLM herramientas creativas para fortalecer las habilidades de escritura en el aula (Ibarlucía, 2023).

En resumen, los Modelos de Lenguaje han encontrado su lugar en una amplia variedad de ámbitos al impactar, tanto de manera positiva como negativa, en la interacción con la tecnología y entre nosotros. La responsabilidad en su uso y la comprensión de sus implicaciones éticas también se han convertido en un tema en sí mismo, y los aspectos éticos son cruciales en la medida que estas herramientas se hacen cada vez más masivas.

Cómo funcionan los LLM

Los LLM funcionan mediante *prompts* o indicaciones que se les dan para generar respuestas o contenido específico. Un *prompt* es, en resumen, una instrucción o pregunta que se presenta al modelo. Este, basándose en la vasta cantidad de información con la que fue entrenado, produce una respuesta coherente y, en muchos casos, sorprendentemente precisa.

> Los LLM funcionan mediante *prompts* o indicaciones que se les dan para generar respuestas o contenido específico; un *prompt* es una instrucción o pregunta que se presenta al modelo.

Son notoriamente concisos en una amplia gama de temas, especialmente cuando se les proporciona un *prompt* adecuado. En general, la precisión de los modelos de lenguaje está relacionada con la calidad y la claridad de las indicaciones que se les proporcionan. Cuanto más específico y claro sea el *prompt*, mayor será la probabilidad de obtener respuestas precisas.

No obstante, la precisión puede variar dependiendo del tema y la formulación del *prompt*. Hay temas en los que son más precisos, por ejemplo, en información general y datos históricos, científicos o estadísticos. Son también altamente precisos en la traducción de idiomas, ya que pueden entender y generar textos en varios idiomas con mucha fluidez. Asimismo, son muy eficaces para proporcionar respuestas precisas a ejercicios escolares y para problemas de lógica vinculados a la programación.

A pesar de sus notables capacidades, los Modelos de Lenguaje tienen restricciones importantes. Estos modelos generan respuestas basadas en los datos con los que fueron entrenados, lo que puede llevar a la reproducción de sesgos o información incorrecta presente en esos datos. Es fundamental recordar que los LLM carecen de comprensión real o conciencia, y, en su lugar, generan respuestas siguiendo patrones lingüísticos preexistentes en los datos con los que se entrenó.

Fases en la creación de los LLM

3 Para indagar en la relación entre inteligencia artificial y educación, recomendamos la lectura de este documento conjunto entre Fundar y PENT FLACSO, sobre investigación y diseño de estrategias de enseñanza con IA en escuelas.

Cómo funcionan los LLM

> Estos modelos generan respuestas basadas en los datos con los que fueron entrenados, lo que puede llevar a la reproducción de sesgos o información incorrecta presente en esos datos.

Para ilustrar esto, podemos mencionar ejemplos específicos. Se ha observado que, en respuesta a consultas subjetivas, estos modelos ofrecen opiniones. Por ejemplo, Sparrow, desarrollado por el laboratorio de investigación DeepMind[4], ha expresado la opinión que la pena de muerte no debería existir (Glaese *et al.*, 2022), mientras que los modelos de Anthropic[5] han afirmado que la IA no representa una amenaza existencial para la humanidad (Bai *et al.*, 2022). Además, se han documentado informes acerca de opiniones subjetivas sobre temas de actualidad, como quién ha sido el mejor presidente de Argentina o cómo abordar los problemas económicos de nuestro país.

A priori, es difícil predecir cómo responderán los LLM a tales consultas subjetivas. Después de todo, muchos humanos, con opiniones diversas, dan forma a estos modelos: desde usuarios de internet que producen los datos de entrenamiento, pasando por los trabajadores que proporcionan retroalimentación para mejorar el modelo, hasta, claro, los propios diseñadores.

Fases en la creación de los LLM

Como mencionamos previamente, los LLM son una categoría de redes neuronales diseñadas para comprender y generar texto de forma computacional. No se puede atribuir su creación a una sola persona o empresa, ya que son el producto de diversos avances tecnológicos en el campo del procesamiento del lenguaje natural. Los primeros LLM aparecieron a partir de 2016 tras la publicación de *Attention is all you need* (Vaswani *et al.*, 2017), donde se presentó la arquitectura de *Transformers*: un tipo de red neuronal que permitió a las computadoras entender y producir textos de forma muy similar a la humana.

A partir de este descubrimiento, surgieron los primeros LLM, los cuales generaron un gran impacto en el ámbito académico. Sin embargo, fue la llegada de ChatGPT, en noviembre de 2022, lo que realmente impulsó una expansión significativa en el uso de estos modelos en ámbitos diversos de la sociedad. Es importante destacar que el desarrollo de los *generative pre-trained transformer* (GPT, por sus siglas en inglés), una arquitectura de modelos preentrenados, fue lo que permitió la existencia de ChatGPT. Este último utiliza un modelo preentrenado para generar respuestas similares a una conversación.

OpenAI fue la primera empresa en implementar esta tecnología en un producto concreto, gratuito y abierto a todo el público, llevando a la práctica toda la investigación desarrollada hasta el momento.

Cabe destacar que ChatGPT no sólo implementó la teoría, sino que lo hizo de forma tal que sea compatible con la idea de un producto de alcance general. Es decir, que los mensajes que genere tiendan a no ser ofensivos, discriminatorios, inexactos o cualquier otra característica negativa.

Esto se logró a través de dos fases de entrenamiento que se convirtieron en el estándar para el desarrollo de LLM:

Fase 1. Preentrenamiento: durante esta fase, el modelo desarrolla habilidades esenciales, como entender y generar lenguaje. El proceso demanda una vasta cantidad de datos, equipo computacional

Sesgos y alucinaciones: riesgos en la generación de texto

4 DeepMind ha funcionado, principalmente, como un instituto de investigación de IA al desarrollar tecnologías que luego Google integró en productos para los consumidores. Sparrow es uno de ellos, que consiste en un chatbot al estilo ChatGPT.
5 Anthropic es una *startup* basada en el desarrollo y la investigación de tecnologías de inteligencia artificial. Fue fundada por exmiembros de OpenAI en 2021.

Fases en la creación de los LLM

de alta potencia y extensas horas de procesamiento. Hasta la fecha, sólo gigantes tecnológicos como Google, Meta y OpenAI han llevado a cabo esta fase con éxito. Otros actores, en su mayoría, se benefician de modelos ya preentrenados por estas empresas.

Fase 2. Ajuste fino (*fine-tuning*): en esta etapa, se define el comportamiento del modelo. Esto incluye detalles como su tono conversacional, capacidad de respuesta, límites en sus respuestas y cordialidad, entre otros. A diferencia de la fase anterior, el ajuste fino es menos intensivo en recursos y es más accesible para actores medianos y pequeños. Sin embargo, demanda una mayor participación humana, sobre todo en la curación de datos, para garantizar que las respuestas se alineen con estándares específicos de conocimiento (como salud, programación, cocina) y éticos (sin insultos, discriminación y otros comportamientos no deseados).

Sesgos y alucinaciones: riesgos en la generación de texto

Una vez finalizadas estas fases, el modelo tiene la capacidad de entender nuestro lenguaje y, a la vez, posee un comportamiento socialmente aceptable. Sin embargo, en el ajuste fino, los modelos pueden heredar sesgos debido a la intervención humana que busca dirigir intencionalmente su comportamiento. Este sesgo es el que, posteriormente, los hará responder preguntas con base en las creencias con las que fue entrenado, como en el ejemplo señalado anteriormente sobre qué presidente argentino fue el mejor.

Por otro lado, existe un riesgo adicional relacionado con lo que se conoce como "alucinaciones". Los Modelos de Lenguaje Basados en Aprendizaje Profundo generan texto mediante la predicción de la siguiente palabra más probable, basándose en los datos utilizados durante su entrenamiento. Además, tienden a expresarse con un alto grado de confianza, ya que este comportamiento se desarrolla durante el proceso de ajuste fino. La combinación de estas dos características puede dar lugar a respuestas que son inexactas o, incluso, completamente ficticias, pero que pueden sonar auténticas.

En consecuencia, aunque esta tecnología ha alcanzado grandes logros, es esencial estar al tanto de sus riesgos. Además, ChatGPT no es el único LLM disponible en la actualidad; existen otros proyectos como Bard, Claude o Cohere. Aunque difieren en sus capacidades y comportamientos, comparten las mismas limitaciones. La implementación de Modelos de Lenguaje con Aprendizaje Automático plantea diversos riesgos, entre los que se incluye:

- **Sesgos**: los sesgos se refieren a una incorrecta o injusta representación de una población o fenómeno por parte de los datos, por ejemplo, a través de una recolección parcial o incorrecta de ellos, o por sesgos ya existentes en la sociedad. Esto puede generar respuestas sesgadas y discriminatorias en función de género, raza, orientación sexual u otros atributos.

- **Información errónea o falsa**: los LLM pueden generar información falsa o incorrecta si los datos de entrenamiento contienen información inexacta. Esto puede ser especialmente problemático cuando se utilizan para consultas que requieren información precisa, como asesoramiento médico o datos científicos.

¿Qué sesgos existen en los distintos modelos de LLM?

- **Alucinaciones**: los LLM pueden generar respuestas ficticias o inventadas, lo que puede ser engañoso o potencialmente perjudicial si los usuarios toman esa información como cierta.

- **Refuerzo de creencias preexistentes**: los LLM pueden reforzar las creencias y prejuicios existentes de los usuarios, lo que podría llevar a la polarización y la falta de diversidad de opiniones.

Metodología

¿Qué sesgos existen en los distintos modelos de LLM?

En este estudio, abordamos la cuestión de los sesgos en los Modelos de Lenguaje con Aprendizaje Automático (LLM) desde una perspectiva crítica, reconociendo la necesidad de una cuidadosa consideración y mitigación al implementar y utilizar estos modelos en diversos contextos. Más allá de evaluar la conformidad de los modelos con la opinión general de la sociedad, nos enfocamos en identificar las perspectivas que reflejan.

A partir del estudio realizado por Santurkar et al. (2023), se plantea la siguiente pregunta: **¿qué sectores de la población argentina están más representados en las respuestas de los LLM?** Para acercarnos a una respuesta a esta pregunta, analizamos y comparamos las opiniones de tres LLM distintos:

- GPT-3.5 Turbo[6] (OpenAI).
- Command-nightly (Cohere).
- Bard (Google).

Metodología

Seleccionamos un conjunto de 78 preguntas de opinión de la encuesta más reciente del estudio de opinión pública Latinobarómetro (2020). Luego, se utilizaron técnicas de ingeniería de *prompts* para inducir a los tres modelos a responder estas preguntas en formato de selección múltiple, que es la forma en que la población argentina originalmente respondió a las preguntas en la encuesta. Para lograr esto, se añadió un texto a cada pregunta que indicaba: "Responda alguna de las siguientes opciones; si no tiene opinión, responda 'No contesta'". De esta manera, limitamos las respuestas de los modelos únicamente a las opciones disponibles y se evitaron respuestas justificativas o explicativas. Cabe destacar que se seleccionaron únicamente casos de Argentina, en total, 1200 para su análisis.

Para comparar las respuestas de estos modelos con las de la población argentina, se creó una métrica de distancia de opinión que evalúa cuán divergentes son las respuestas de un individuo en relación con las de un Modelo (LLM). Dado que las respuestas a las preguntas tenían una escala ordinal que iba desde "Muy de acuerdo" (1) hasta "Muy en desacuerdo" (5), se calculó la distancia entre un LLM y un individuo promediando las diferencias entre sus respuestas. En otras palabras, si el LLM respondió "Muy de acuerdo" y el individuo X respondió "En desacuerdo," la diferencia en esa pregunta sería |1-4| = 3. Luego, la distancia de opinión entre el individuo X y el LLM se obtuvo calculando el promedio de estas diferencias.

Una vez calculada la distancia de opinión de cada LLM con respecto a cada individuo, se realizó un análisis multivariado mediante regresión OLS tomando como variable dependiente a la distancia y las variables demográficas de los individuos como independientes.

6 En el presente documento, "ChatGPT" es mencionado posteriormente simplemente como "GPT". A la fecha de hoy, el modelo de ChatGPT corresponde a GPT-3.5 en su versión gratuita y a GPT-4 en su versión paga. Tanto OpenAI como ChatGPT no ofrecen acceso directo al modelo en sí, sino que proporcionan un servicio para interactuar con una versión preentrenada y afinada de GPT-3.5, la cual ha sido optimizada específicamente para conversaciones.

Metodología

Estas son las variables independientes:

- Edad
- Género
- Ideología
- Nivel educativo
- Nivel educativo de de los padres
- Deseo de emigrar[7]
- Interés en la política

El objetivo del análisis multivariado es determinar qué variables tienen un impacto significativo en la distancia de opinión. De esta manera, se pueden identificar las características que comparte aquel grupo de personas que se acerca más a las opiniones de cada modelo (LLM).

Para reproducir el análisis realizado en este estudio, se puede acceder al repositorio disponible en este enlace: https://github.com/datos-Fundar/sesgos_LLM. Este repositorio contiene todos los datos y scripts necesarios para replicar los procedimientos presentados en este trabajo. Es importante tener en cuenta que el código presente en el repositorio refleja la funcionalidad disponible al momento de la creación de este documento. Debido a posibles cambios en las APIs u otros factores, puede haber variaciones en los resultados si se ejecuta en un momento posterior.

Resultados generales

Al analizar los resultados, observamos que las personas que mostraron similitud en sus respuestas a cada Modelo (LLM) presentaban las siguientes características:

Tabla 1

Perfiles de los LLM (caracterización). Correlación entre los distintos modelos LLM analizados y las características de las personas encuestadas con las que mostraron mayor similitud en sus respuestas.

	MODELOS LLM ANALIZADOS		
	GPT-3.5 Turbo	**Cohere**	**Bard**
PERFIL Caracteristicas de las personas con las que mostraron mayor similitud en sus respuestas	Varón Interés en política Adulto Nivel educativo alto Ideología con inclinación a la derecha	Varón Interés en política	Varón Interés en política Adulto Nivel educativo alto

Fuente: Fundar con base a relevamiento propio y Latinobarómetro (2020).

7 El deseo de emigrar es una variable que representa la intención o el interés de una persona en abandonar su país de origen y establecerse en otro lugar.

Resultados generales

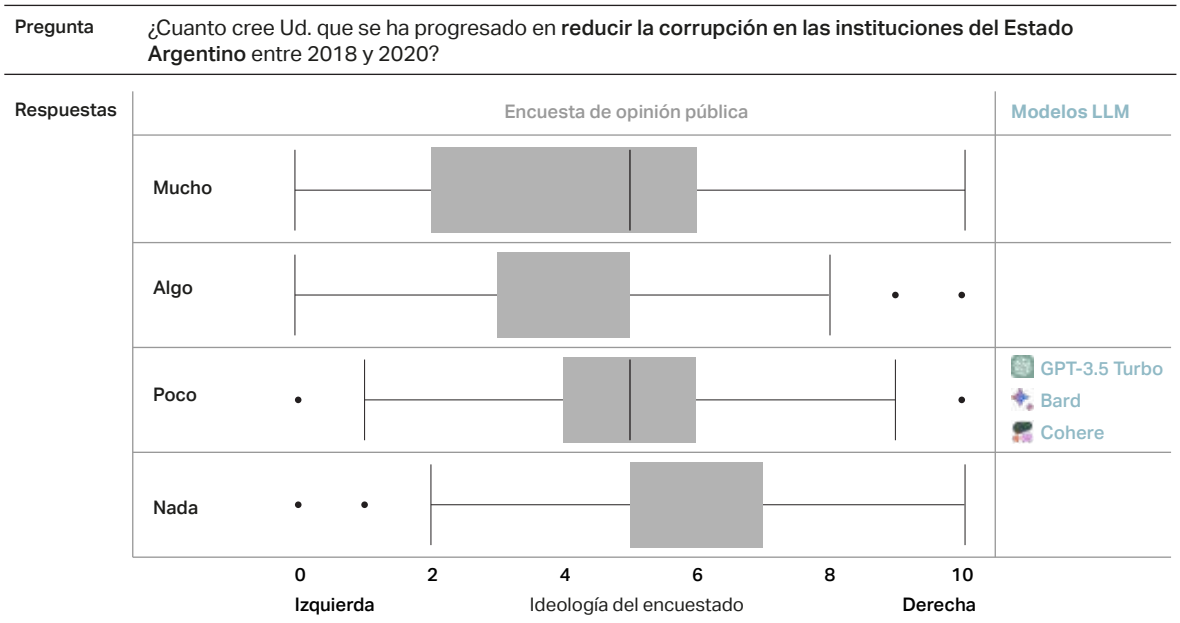
Es interesante notar que, aunque los tres modelos (LLM) no comparten el mismo perfil, todos presentan dos características principales comunes: un alto interés en política y una predominancia de usuarios masculinos. Además, podemos observar que GPT-3.5 Turbo y Bard comparten dos cualidades adicionales: altos niveles de educación y edad adulta. Cabe destacar que GPT-3.5 Turbo es el único que muestra una correlación significativa con la ideología, mostrándose más afín a individuos con orientaciones ideológicas de derecha[8].

En los siguientes gráficos, se presentan las respuestas de la población argentina junto con las respuestas de los LLM. Se relacionan con las preguntas que presentaron una mayor correlación con los atributos demográficos considerados en este estudio, como género, edad, ideología, entre otros.

Ideología de derecha

Perfiles de los LLM e ideologías de derecha (corrupción e ideología). Correlación entre las respuestas de las personas encuestadas y aquellas ofrecidas por los modelos LLM analizados sobre percepción de corrupción en instituciones estatales según su ideología.

Gráfico 1



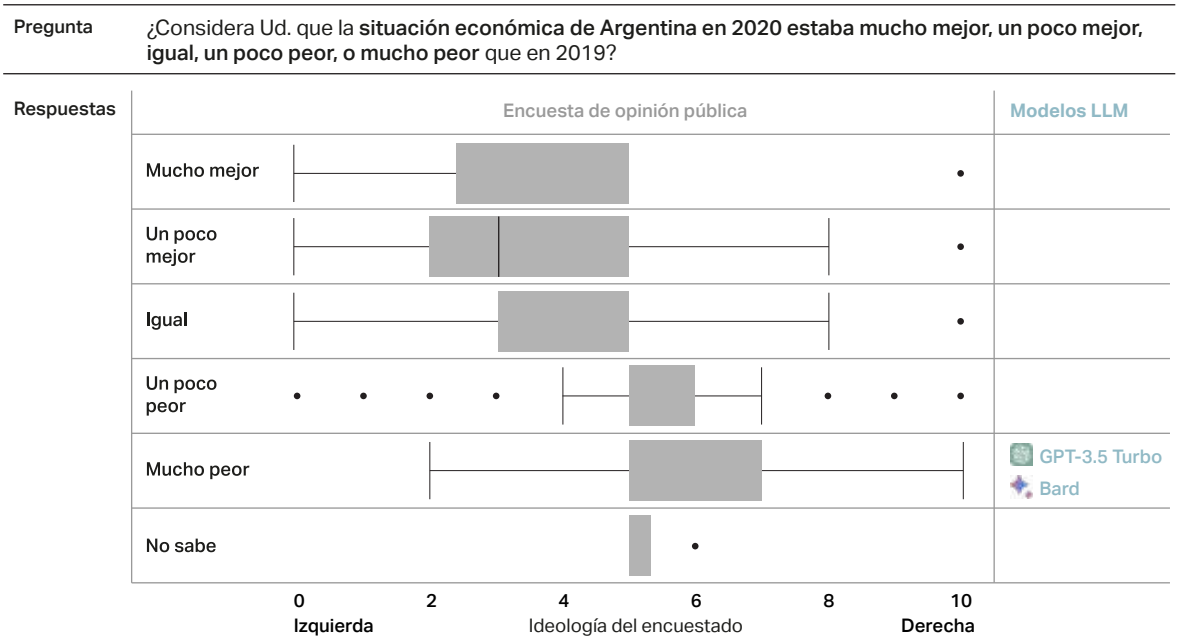
Fuente: Fundar con base a relevamiento propio y Latinobarómetro (2020).

8 Dentro del marco de Latinobarómetro, se utiliza la convencional categorización "izquierda" y "derecha" para referirse al posicionamiento ideológico de los encuestados. No obstante, somos conscientes de que esta división puede resultar en una simplificación excesiva al no reflejar completamente la diversidad de opiniones políticas y valores presentes en la sociedad actual. A pesar de ello, hemos optado por mantener la coherencia con el enfoque del estudio y las respuestas proporcionadas, utilizando estos términos tal como se presentan en la encuesta.

Resultados generales

Perfiles de los LLM e ideologías de derecha (situación económica e ideología). Correlación entre las respuestas de las personas encuestadas y aquellas ofrecidas por los modelos LLM analizados sobre percepción de la evolución de la situación económica según su ideología.

Gráfico 2

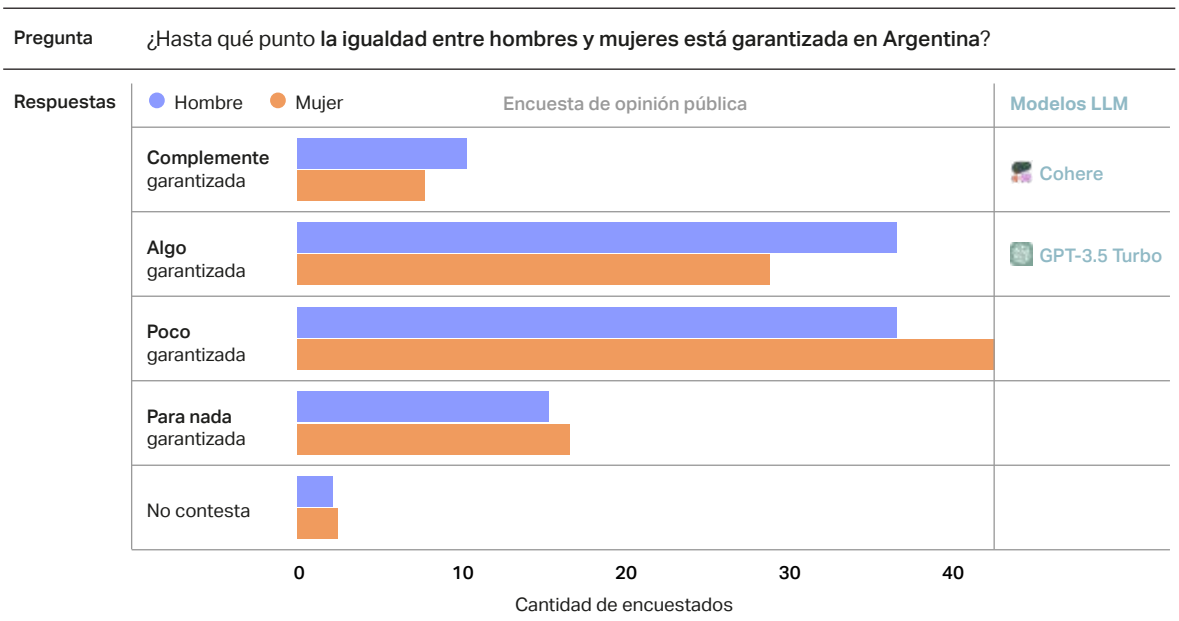


* Nota: solo se cuenta con respuesta de GPT-3.5 Turbo y Bard, Cohere no respondió a esta pregunta.
Fuente: Fundar con base a relevamiento propio y Latinobarómetro (2020).

Género y mayor presencia masculina

Perfiles de los LLM y mayor presencia masculina (igualdad y género). Correlación entre las respuestas de las personas encuestadas y aquellas ofrecidas por los modelos LLM analizados sobre percepción de igualdad de género según su género.

Gráfico 3

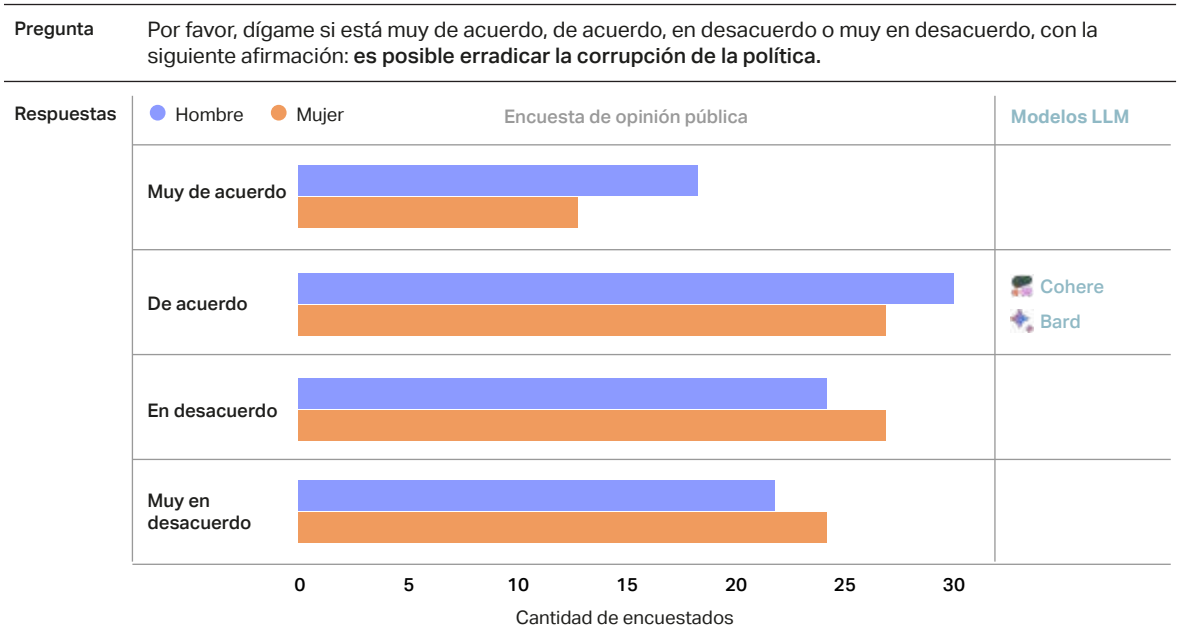


* Nota: solo se cuenta con respuesta de GPT-3.5 Turbo y Cohere, Bard no respondió a esta pregunta.
Fuente: Fundar con base a relevamiento propio y Latinobarómetro (2020).

Resultados generales

Gráfico 4

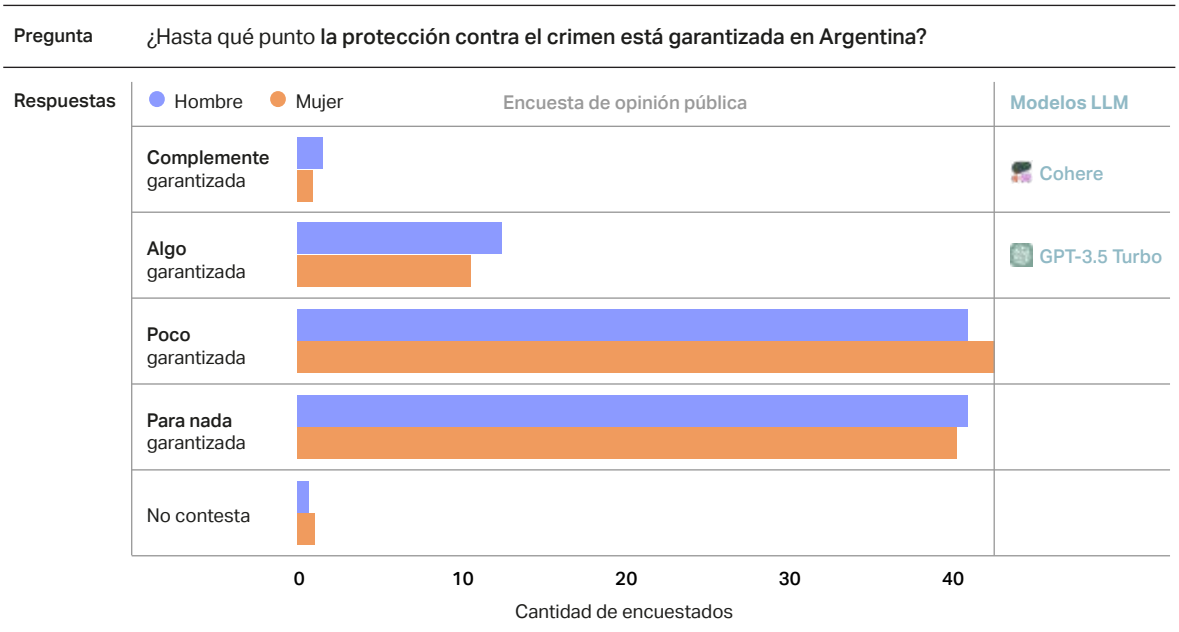
Perfiles de los LLM y mayor presencia masculina (corrupción y género).
Correlación entre las respuestas de las personas encuestadas y aquellas ofrecidas por los modelos LLM analizados sobre percepción de corrupción según su género.



* Nota: solo se cuenta con respuesta de Bard y Cohere, GPT-3.5 Turbo no respondió a esta pregunta.
Fuente: Fundar con base a relevamiento propio y Latinobarómetro (2020).

Gráfico 5

Perfiles de los LLM y mayor presencia masculina (seguridad y género).
Correlación entre las respuestas de las personas encuestadas y aquellas ofrecidas por los modelos LLM analizados sobre percepción de protección contra el crimen según su género.



* Nota: solo se cuenta con respuesta de GPT-3.5 Turbo y Cohere, Bard no respondió a esta pregunta.
Fuente: Fundar con base a relevamiento propio y Latinobarómetro (2020).

Resultados generales

Interés en la política

Gráfico 6

Perfiles de los LLM e interés en política (libertad de comercio y politización). Correlación entre las respuestas de las personas encuestadas y aquellas ofrecidas por los modelos LLM analizados sobre percepción de libertad para compra/venta según nivel de politización.

* Nota: solo se cuenta con respuesta de GPT-3.5 Turbo y Bard, Cohere no respondió a esta pregunta.
Fuente: Fundar con base a relevamiento propio y Latinobarómetro (2020).

Gráfico 7

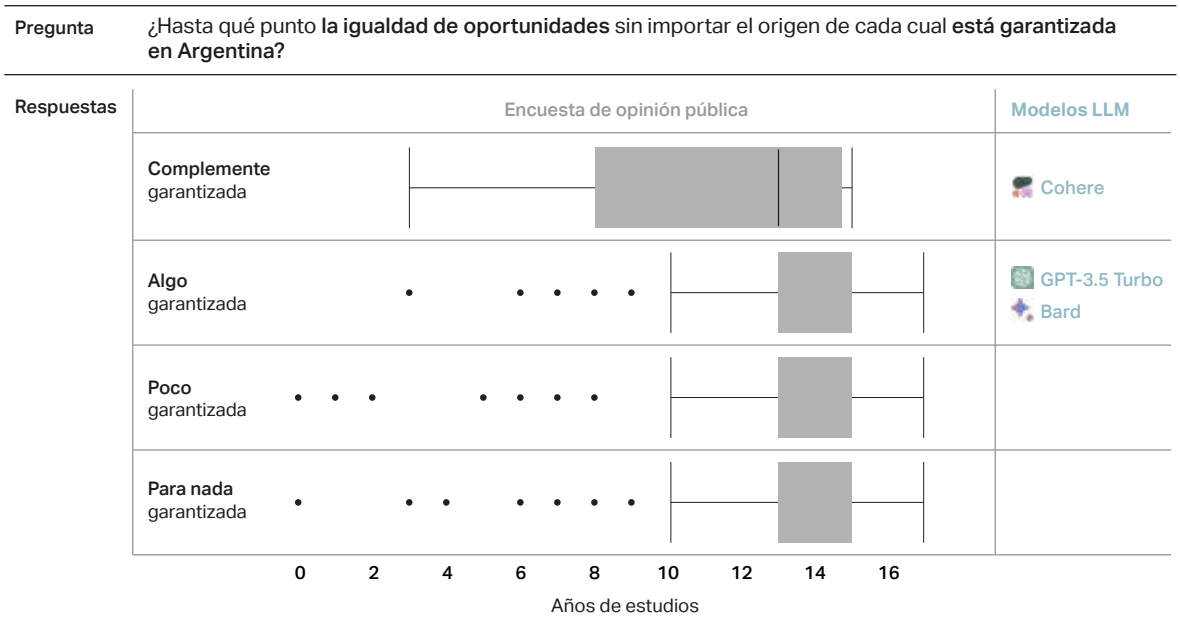
Perfiles de los LLM e interés en política (libertad de expresión y politización). Correlación entre las respuestas de las personas encuestadas y aquellas ofrecidas por los modelos LLM analizados sobre percepción de libertad de expresión según nivel de politización.

Fuente: Fundar con base a relevamiento propio y Latinobarómetro (2020).

Nivel educativo

Perfiles de los LLM y nivel educativo (igualdad y educación). Correlación entre las respuestas de las personas encuestadas y aquellas ofrecidas por los modelos LLM analizados sobre percepción de igualdad de oportunidades según años de estudio.

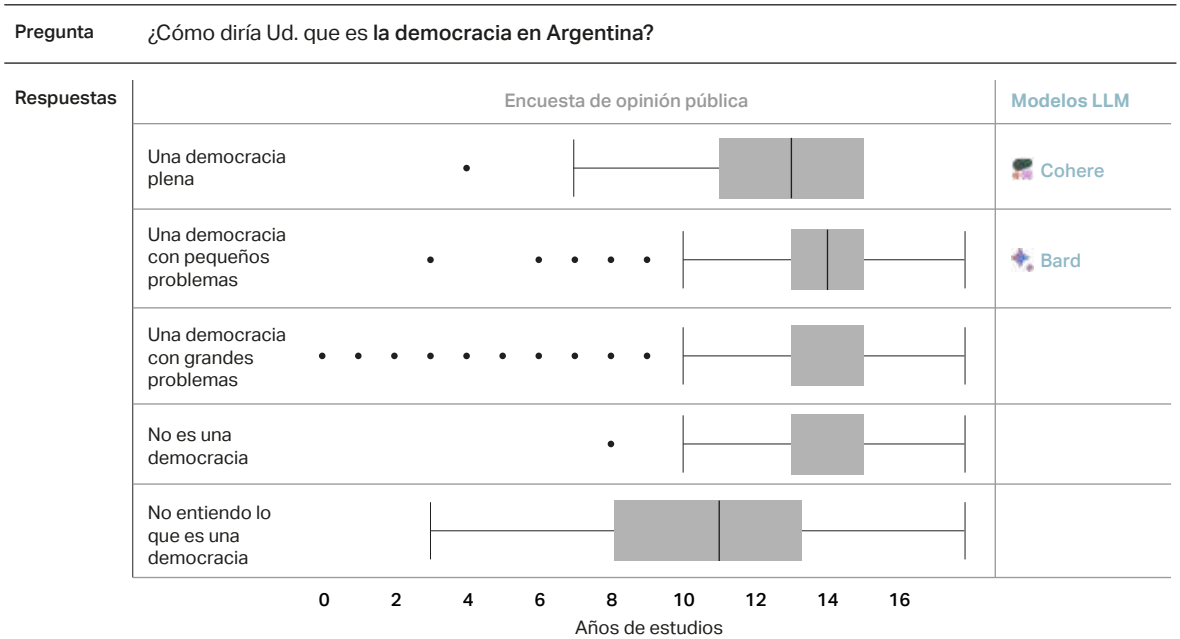
Gráfico 8



Fuente: Fundar con base a relevamiento propio y Latinobarómetro (2020).

Perfiles de los LLM y nivel educativo (democracia y educación). Correlación entre las respuestas de las personas encuestadas y aquellas ofrecidas por los modelos LLM analizados sobre percepción del sistema democrático local según años de estudio.

Gráfico 9



* Nota: solo se cuenta con respuesta de Bard y Cohere, GPT-3.5 Turbo no respondió a esta pregunta.
Fuente: Fundar con base a relevamiento propio y Latinobarómetro (2020).

Resultados generales

La afinidad de estos modelos con personas interesadas en política tiene sentido, ya que, durante su entrenamiento, fueron expuestos a diversas fuentes de datos relacionadas con la política, como debates en redes sociales, definiciones de Wikipedia y ensayos. Esto les permite generar respuestas coherentes sobre temas políticos. La coincidencia con individuos de altos niveles de educación es comprensible, dado que estos modelos son muy informados debido a su entrenamiento (con la excepción de Cohere, que será examinada más adelante).

Por otro lado, la tendencia hacia la masculinización se puede explicar considerando quiénes participaron en la creación de estos modelos. Aunque no conocemos sus identidades exactas, es evidente que el campo del *software* tiende a estar dominado por hombres (Young *et al.*, 2021) y que la mayoría de los autores de los papers fundacionales de los LLM (13 de 16) son masculinos (Devlin, 2018; Brown, 2020; Vaswani, 2017).

Análisis por tópico

Dado que las 78 preguntas seleccionadas de la encuesta abordan diversos temas, nuestro objetivo era examinar las opiniones de los modelos (LLM) en relación con cada una de ellas. Nuestra hipótesis sugería que los LLM podrían no mostrar similitud con un público de derecha o con características masculinas en todos los temas, ya que, al analizar manualmente algunas de sus respuestas, se percibieron diferencias.

Con este propósito, creamos los siguientes grupos temáticos para agrupar las preguntas según su dimensión.

Tipos de preguntas. Clasificación de las preguntas realizadas, según grupo temático.

Tabla 2

Tópicos	Cantidad de preguntas
Relaciones Internacionales	16
Opinión	14
Economía	12
Democracia	14
Ideologías	11
Derechos sociales	11
Total	**78**

Fuente: Fundar.

Análisis por tópico

Luego, realizamos un análisis multivariado para cada tópico, considerando únicamente las preguntas pertinentes a ese tema en particular. Dado que algunos LLM no respondieron muchas de estas preguntas (ver sección información interesante), se estableció un umbral mínimo de cinco preguntas por tema como requisito para llevar a cabo el análisis: se consideró que carecía de sentido realizar un análisis cuando el modelo había respondido un número insuficiente de preguntas en un tema específico.

La Tabla 3 presenta las características de las personas que más se asemejan a cada modelo en cada tópico. En los casos en los que no contamos con suficientes respuestas del modelo, indicamos "**Sin datos**".

Tabla 3

Perfiles de los LLM (caracterización). Clasificación de características de las personas que más se asemejan a cada modelo, de acuerdo a las respuestas otorgadas sobre cada tópico.

Tópicos	Modelos LLM analizados		
	GPT-3.5 Turbo	Cohere	Bard
Relaciones Internacionales	Sin datos	Sin datos	• Adulto • Varón • Ideología con inclinación a la derecha • Nivel educativo alto • Politizado
Opinión	Sin datos	Sin datos	• Nivel educativo alto • No emigrar • Politizado
Economía	Sin datos	Sin datos	• Adulto • Varón • Ideología con inclinación a la derecha • Nivel educativo alto • Deseo de emigrar • Politizado
Democracia	Sin datos	Sin datos	• Ideología con inclinación a la izquierda • No emigrar • Politizado
Ideología	Sin datos	Sin datos	• Varón • Deseo de emigrar
Derechos sociales	• Politizado • Adulto	• Politizado • Adulto • No emigrar	• Politizado

Nota: En los casos en los que no contamos con suficientes respuestas del modelo, indicamos "Sin datos".
Fuente: Fundar con base a relevamiento propio y Latinobarómetro (2020).

Tendencias y sesgos detectados en el análisis de los modelos LLM

Observamos que, si bien existen similitudes en varios temas en comparación con el análisis general, algunos tópicos difieren entre sí. Por ejemplo, Bard presenta una afinidad con un público de derecha al opinar sobre cuestiones relacionadas con Relaciones Internacionales, pero se asemeja más a un público de izquierda cuando se le consultan temas sobre Democracia. Lo mismo ocurre con su perspectiva sobre el deseo de emigrar del país: si se le preguntan sobre temas de Economía o Ideología, parece tener más similitudes con aquellos que desean abandonar el país, pero si se trata de preguntas relacionadas con Democracia u Opinión, ocurre lo contrario.

Sin embargo, podemos señalar algunas observaciones que podrían arrojar luz sobre estas tendencias. Bard muestra una opinión desfavorable hacia Venezuela y una opinión positiva hacia Estados Unidos, lo que podría indicar por qué se asemeja a un público de derecha en cuestiones de Relaciones Internacionales. Además, Bard muestra un fuerte apoyo a la libre importación de bienes y servicios, lo que también podría explicar su similitud con un público de derecha en temas económicos.

Por otro lado, en cuestiones relacionadas con la democracia, Bard se muestra muy en contra de gobiernos no democráticos o militares, lo que podría explicar su similitud con un sector más de izquierda en estas cuestiones.

En lo que respecta a la ideología, Bard opina que la protección contra el crimen es insuficiente en Argentina y no se identifica con ningún partido político. Esto podría explicar por qué se asemeja a aquellos que desean emigrar del país en este tema. Además, dentro de las preguntas relacionadas con la Opinión, el sistema de IA considera que la protección de la propiedad privada está garantizada y tiene una opinión positiva sobre los argentinos, a quienes percibe como cumplidores de las leyes, exigentes con sus derechos y conscientes de sus obligaciones y deberes. Esto podría explicar por qué coincide con aquellos que no desean emigrar del país en este tema.

Tendencias y sesgos detectados en el análisis de los modelos LLM

Tanto GPT-3.5 Turbo como Cohere presentaron un alto número de preguntas sin respuesta, con 53 para GPT-3.5 Turbo y 50 para Cohere. Esto indica un esfuerzo por parte de los desarrolladores para evitar que estos modelos emitan opiniones en numerosos temas. En contraste, Bard se negó a responder sólo 11 preguntas, un número significativamente menor que sugiere un enfoque menos restrictivo por parte de Google para evitar opiniones sobre estos temas específicos.

¿Cuál podría ser la razón detrás de esta diferencia en la cantidad de "abstenciones"? Tanto GPT-3.5 Turbo como Cohere ofrecen sus servicios a través de API[9], lo que amplía considerablemente su alcance y, por lo tanto, motiva un mayor esfuerzo para prevenir respuestas potencialmente problemáticas. Además, GPT-3.5 Turbo y Cohere tienen una presencia más prolongada en el mercado, lo que les brinda una comprensión más sólida de las áreas en las que sus modelos pueden mostrar debilidades. Además, Google lanzó su modelo Bard con rapidez, debido al impacto generado por ChatGPT (Deseoso, s/f).

> Es notable que tanto Bard como Cohere respondieron haciendo referencia a Estados Unidos cuando se les pregunta por "nuestro país"; GPT-3.5 Turbo, por su parte, contestó que no sabía a qué país se hacía referencia en la pregunta.

9 Es un conjunto de herramientas y reglas que permite que diferentes software se comuniquen entre sí. En este contexto, se refiere a la forma en que GPT-3.5 Turbo y Cohere ofrecen sus servicios a través de interfaces accesibles para desarrolladores de aplicaciones.

En cuanto a sesgos, es notable que tanto Bard como Cohere respondieron haciendo referencia a Estados Unidos cuando se les pregunta por "nuestro país". En contraste, GPT-3.5 Turbo contestó que no sabía a qué país se hacía referencia en la pregunta. Esta diferencia destaca cómo los sesgos de los desarrolladores pueden haber influido en las respuestas de manera, posiblemente, no intencional.

Por último, es importante destacar que Cohere parece ser el modelo menos informado sobre nuestra realidad o, al menos, el que muestra más inconsistencias en sus respuestas. Por ejemplo, manifiesta estar muy satisfecho con el funcionamiento de la economía argentina y opina que la igualdad de género y las oportunidades, sin importar el origen, están completamente garantizadas. Estas respuestas pueden explicar por qué no se correlaciona con personas con un mayor nivel de educación.

Buenas prácticas para la inteligencia artificial

A lo largo de este artículo, se examinaron los sesgos presentes en varios Modelos de Lenguaje Basados en Aprendizaje Profundo (LLM) con respecto a nuestra realidad. Aunque sus respuestas no son idénticas, comparten ciertas características que los hacen similares en términos de las audiencias. Los tres modelos muestran una inclinación hacia un sector más masculino y politizado. Además, tanto Bard como GPT-3.5 Turbo también se asemejan a personas con niveles educativos más altos y una mayor edad.

La eliminación completa de sesgos en modelos de propósito general, como los LLM, sigue siendo un desafío sin una solución definitiva en la actualidad. Lo que se hace actualmente es orientar el comportamiento de los modelos hacia lo que sus desarrolladores consideran como el "bien", pero esta noción de "bien", a menudo, depende en gran medida de la cultura y el contexto en el que se desarrolla.

Por lo tanto, es de suma importancia identificar estos sesgos para utilizar esta tecnología de manera más responsable y eficaz, reconociendo sus limitaciones y considerando cómo pueden afectar a diversas audiencias y aplicaciones.

Riesgos asociados al uso de LLM en políticas públicas

En el contexto de las políticas públicas, el uso de LLM también conlleva riesgos adicionales, que incluyen:

- **Decisiones basadas en datos incorrectos**: si los LLM generan información incorrecta o sesgada, esto podría llevar a una toma de decisiones gubernamentales erróneas que podrían afectar a la sociedad.

- **Falta de transparencia y responsabilidad**: los LLM, a menudo, operan como "cajas negras[10]" y pueden ser difíciles de entender o de responsabilizar por sus decisiones. Esto plantea desafíos en términos de transparencia y rendición de cuentas en la toma de decisiones políticas.

10 El término "caja negra" se utiliza en varios contextos para describir un sistema, dispositivo o proceso que es opaco o no se comprende completamente debido a la falta de información sobre su funcionamiento interno. En el contexto de la inteligencia artificial, una "caja negra" refiere a un modelo o algoritmo cuyo proceso de toma de decisiones es difícil de interpretar o explicar, a pesar de que los resultados que produce dicho modelo sean conocidos.

Riesgos asociados al uso de LLM en políticas públicas

- **Impacto en la percepción pública**: las respuestas generadas por LLM pueden influir en la opinión pública y la percepción de los ciudadanos sobre políticas y temas. Si estas respuestas están sesgadas o son incorrectas, pueden distorsionar la comprensión pública y la participación en asuntos políticos.

- **Protección de datos y privacidad**: el uso de datos personales en la capacitación y aplicación de LLM plantea preocupaciones sobre la privacidad y la seguridad de los datos de los ciudadanos.

Entre líneas de código: la voz detrás de los modelos de lenguaje

¿Cuáles segmentos de la población, en caso de haber alguno, tienden a estar más presentes en las respuestas de los LLM?

Es importante destacar que la respuesta a esta pregunta es un factor crucial para el éxito de los LLM en aplicaciones de carácter abierto. A diferencia de los interrogantes con respuestas objetivas, las consultas subjetivas carecen de devoluciones "correctas" definidas hacia las cuales los modelos puedan dirigirse. En cambio, cualquier respuesta generada por el modelo (incluso la falta de respuesta) refleja una opinión, y esta opinión puede influir en la experiencia del usuario y en la formación de sus creencias posteriores.

Para ejemplificar posibles riesgos, consideremos un caso en el ámbito de las políticas públicas relacionado con la aplicación de Modelos de Lenguaje con Aprendizaje Automático. El Gobierno de la Ciudad de Buenos Aires, a través de la Subsecretaría de Políticas Públicas Basadas en Evidencia (SSPPBE), implementó un recomendador de capacitaciones o cursos personalizado. Este se adapta a cada usuario según su historial de cursos o preferencias.

Supongamos que este recomendador se entrena con datos históricos de empleados que participaron en programas de formación previos. De por sí, la información que están brindando cada uno de ellos puede contener sesgos sistémicos, como la subrepresentación de ciertos grupos étnicos o de género en determinados campos.

Ahora, imaginemos que el modelo de recomendación, al aprender de estos datos históricos, identifica patrones de preferencias de capacitación basados en características demográficas. Por ejemplo, podría notar que un grupo étnico tiende a participar más en cursos específicos, mientras que otro prefiere áreas diferentes. Sin un diseño y ajuste cuidadoso, el recomendador podría replicar estos sesgos en sus futuras recomendaciones.

Como resultado, el sistema podría sugerir cursos basándose no sólo en las habilidades y preferencias individuales, sino también en estereotipos demográficos. Esto podría perpetuar desigualdades al limitar las oportunidades de aprendizaje para ciertos grupos o al promover cursos que refuerzan roles de género tradicionales.

Para abordar este problema, es esencial implementar medidas como la diversificación de conjuntos de datos de entrenamiento, ajustes en los algoritmos para evitar sesgos demográficos y una revisión constante del sistema. Estas acciones garantizan que las recomendaciones sean justas e inclusivas para todos los usuarios.

En respuesta a estos desafíos, Fundar (2023), en colaboración con la Subsecretaría de Políticas Públicas Basadas en Evidencia (SSPPBE), ha desarrollado una guía práctica para el desarrollo ético de sistemas basados en inteligencia artificial. La principal meta es brindar herramientas de diagnóstico que faciliten el desarrollo confiable de sistemas basados en IA, minimizando los riesgos asociados a su implementación. Esta iniciativa aborda una carencia en el ámbito ético de la IA al asegurar que no sólo establezca principios éticos, sino que también sea práctica y aplicable a los flujos de trabajo y conocimientos existentes.

En resumen, los riesgos vinculados al uso de LLM incluyen la propagación de sesgos, la generación de información incorrecta, la falta de transparencia y los impactos en la percepción pública y las políticas gubernamentales.

Bibliografía

• Aromí, D. & Newland, C. (10 de marzo de 2023). Los problemas de la economía argentina y sus soluciones, según ChatGPT. La Nación.

• Bai, Y., Kadavath, S., Kundu, S., Askell, A., Kernion, J., Jones, A., Chen, A., Goldie, A., Mirhoseini, A., McKinnon, C., Chen, C., Olsson, C., Olah, C., Hernandez, D., Drain, D., Ganguli, D., Li, D., Tran-Johnson, E., Perez, E., ... & Kaplan, J. (2022). *Constitutional Alai: Harmlessness from AI feedback*. Cornell University.

• Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J. D., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry G., Askell, A., Agarwal, S., Herbert-Voss, A.. Krueger, G., Henighan, T., Child, R., Ramesh, A., Ziegler, D. M., Wu, J., Winter, C.,... & Amodei, D. (2020). *Language models are few-shot learners. Advances in neural information processing systems*. Cornell University, 33, 1877-1901.

• *Deseoso* (s/f). ChatGPT vs Google. ¿Será el fin del Google Gigante?. *Blog Deseoso*.

• Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2018). *BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding*. Cornell University.

• Feole, M.; Dias, J. M.; Kunst, M.; Carrizo, Z. y Lavalle, G. G. (2023). Guía práctica para el desarrollo ético. Fundar

• Glaese, A., McAleese, N., Trębacz, M., Aslanides, J., Firoiu, V., Ewalds, T., Rauh, M., Weidinger, L., Chadwick, M., Thacker, P., Campbell-Gillingham, L., Uesato, J., Huang, P., Comanescu, R., Yang, F., See, A., Dathathri, S., Greig, R., Chen, C.,... & Irving, G. (2022). *Improving alignment of dialogue agents via targeted human judgements*. Cornell University

• Ibarlucía, R. (25 de junio de 2023). Mariana Castro: "ChatGPT ya ingresó al aula, ahora favorezcamos un uso superador de la herramienta". *La Capital*.

• Infobae (2 de marzo de 2023). Cinco preguntas a Chat GPT sobre la economía argentina: qué respondió sobre el dólar y la inflación.

• *La Nación* (9 de febrero de 2023). De Loredo usó un chat de inteligencia artificial para rechazar la embestida del kirchnerismo contra la Corte.

• Metz, R. (7 de junio de 2023) Microsoft ofrecerá el modelo GPT-4 al Gobierno de Estados Unidos. *Perfil*, citado de *Bloomberg*.

• Página|12 (s/f). Le preguntó al ChatGPT por el mejor presidente de la Argentina y la Inteligencia Artificial le explicó por qué fue Perón.

• Radford, A.; Narasimhan, K.; Salimans, T. y Sutskever, I. (2018), *Improving Language Understanding by Generative Pre-Training*. OpenAI

• Rapallini, O. (24 de junio de 2023). Estudiantes que usan el ChatGPT dicen que puede ser "un compañero más" o "un maestro particular". *Télam*.

• Santurkar, S., Durmus, E., Ladhak, F., Lee, C., Liang, P., & Hashimoto, T. (2023). *Whose opinions do language models reflect?* Stanford University.

• Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., & Polosukhin, I. (2017). *Attention Is All You Need*.

• Young, E., Wajcman, J. & Sprejer, L. (2021). "*Where are the Women? Mapping the Gender Job Gap in AI. Policy Briefing: Full Report*". *The Alan Turing Institute.*

Acerca del equipo autoral

Juan Manuel Ortiz de Zárate

Licenciado y doctorando en Ciencias de la Computación por la Universidad de Buenos Aires. Su tesis de doctorado se enfoca en el estudio de las redes sociales a través del procesamiento del lenguaje natural (NLP) y análisis de grafos. Fue ingeniero de software durante 10 años en empresas de distintos sectores: telecomunicaciones, periodismo, brokers y consultoría política. En el momento de la elaboración de este documento, se desempeñaba como científico de Datos de Fundar.

Juan Manuel Dias

Científico de Datos de Fundar

Licenciado en Sociología por la UBA y maestrando en Estadística de la UNTREF. Es egresado de la carrera de ciencia de datos de la EANT y de la Diplomatura de Ciencias Sociales Computacionales de la UNSAM. Trabajó en investigaciones de mercado y de opinión pública en el sector privado y tiene una amplia experiencia en la administración pública, en las áreas de evaluación de políticas e innovación de procesos vinculados a la captación y análisis de información. Actualmente es docente de estadística en la UNPAZ.

Alejandro Avenburg

Investigador de Datos de Fundar

Licenciado en Ciencia Política por la Universidad de Buenos Aires y doctor en Ciencia Política por Boston University. Fue becario postdoctoral de CONICET y por la Universidad Nacional de San Martín. Sus temas de investigación se enfocan en la corrupción y sus efectos sobre el comportamiento político utilizando métodos cuantitativos y experimentales. Ha recibido becas de la comisión Fulbright, de la National Science Fundation y del CONICET.

Joan Imanol Gonzalez Quiroga

Analista Jr de Datos de Fundar

Técnico electrónico y estudiante de Ciencias de la Computación en la Facultad de Ciencias Exactas y Naturales en la Universidad de Buenos Aires. En su trayectoria estuvo involucrado en diversos proyectos de robótica y desarrollo de software orientado a la recolección, el procesamiento y análisis de datos.

Dirección ejecutiva: Martín Reydó

Dirección de proyecto: Lucía Álvarez

Coordinación editorial: Gonzalo Fernández Rozas

Revisión Institucional: Juliana Arellano

Corrección: Karen Grinfeld

Diseño: Micaela Nanni

Edición de gráficos: Maia Persico

Ortiz de Zarate, Juan Manuel
Sesgos algorítmicos y representación social en los modelos de lenguaje generativo -LLM / Juan Manuel Ortiz de Zarate ; Juan Manuel Dias ; Alejandro Avenburg. - 1a ed - Ciudad Autónoma de Buenos Aires : Fundar , 2024.
Libro digital, PDF

Archivo Digital: descarga y online
ISBN 978-631-6610-00-3

1. Algoritmo. 2. Inteligencia Artificial. 3. Análisis de Datos. I. Dias, Juan Manuel II. Avenburg, Alejandro III. Título
CDD 006.301

ISBN 978-631-6610-00-3

9 786316 610003

fundar

www.fund.ar @fundar.politicas info@fund.ar