



La IA y un modelo que reproduce desigualdades

Lilian San Martín Medina
Directora de Gestión de Operaciones Académicas
Presidenta Comité de Igualdad de Género, Facultad de Ingeniería UNAB

Plataformas como ChatGPT y otros modelos de Inteligencia Artificial (IA) procesan grandes volúmenes de datos, aprenden y se ajustan según la información con la que son entrenados. Esto implica que esos datos, en muchos casos, reflejan las desigualdades preexistentes en la sociedad y, al ser replicados, perpetúan su existencia.

Los sesgos de género en la inteligencia artificial no son accidentales, sino una manifestación de las construcciones sociales que se filtran en los sistemas que diseñamos. Las mujeres, por ejemplo, son más frecuentemente asociadas con roles tradicionales como la enfermería o la educación, mientras que los hombres se vinculan con posiciones de poder y liderazgo. Esta representación refleja las estructuras de poder predominantes en la sociedad.

La integración de la IA en procesos de toma de decisiones y en la vida cotidiana hace que este tipo de controversias sea aún más relevante. El riesgo no es solo que los sistemas reproduzcan estereotipos, sino que los refuercen de tal manera que afecten a millones de personas en aspectos clave como la salud, la educación o el empleo. Por ejemplo, en el campo de la salud, la IA tiende a enfocarse en los síntomas típicos de los hombres, lo que puede llevar a diagnósticos incorrectos o tratamientos inadecuados para las mujeres. Además, el uso de voces femeninas en asistentes virtuales refuerza la idea de que las mujeres son más aptas para realizar tareas de servicio, mientras que los modelos de lenguaje tienden a asociar la profesión de "enfermería" con mujeres y "ciencia" con hombres, perpetuando roles de género que limitan las posibilidades de las personas en función de su sexo.

El desafío para abordar estos sesgos es considerable. No se trata únicamente de eliminar los datos problemáticos, sino de replantear el diseño de los algoritmos para que no solo reconozcan estos estereotipos, sino que los cuestionen y los corrijan activamente. Para lograrlo, es preciso que los desarrolladores

de IA sean conscientes de los prejuicios inherentes en los datos con los que entrenan los modelos. Esto es especialmente relevante en el contexto de plataformas como ChatGPT Latam, donde los sesgos de género pueden verse exacerbados por las particularidades culturales y sociales de la región.

Corregir estos sesgos requiere más que ajustes técnicos. Se necesita un compromiso profundo con la creación de tecnologías que respeten la equidad y reflejen una verdadera diversidad. Los esfuerzos por mejorar la inclusión en la IA deben ir acompañados de la creación de equipos de desarrollo diversos, que comprendan los matices sociales y culturales de los problemas que intentan resolver. La implementación de metodologías que promuevan la inclusión y la transparencia en el proceso de desarrollo es clave para garantizar que las soluciones sean efectivas y no reproduzcan los mismos problemas que intentan solucionar.

Tal como muestra el Informe de la Brecha Global de Género de 2023, solo el 30% de las personas que trabajan en IA son mujeres. Esto evidencia la falta de representatividad de género en el desarrollo de estas tecnologías, lo que contribuye a la perpetuación de sesgos de género en los modelos. Por otro lado, la brecha digital de género, especialmente en países de bajos ingresos, provoca un déficit de datos, lo que agrava el sesgo de género en la IA.

El impacto de estos sesgos es tangible desde los inicios de la masificación de esta tecnología. Ya en 2018, Amazon abandonó una herramienta de IA para reclutar personal, debido a que favorecía los currículums de hombres sobre los de mujeres. Además, los sistemas de reconocimiento facial han mostrado dificultades para identificar correctamente a las mujeres, en particular a las mujeres de color, lo que puede resultar en errores de identificación con consecuencias graves, especialmente en el ámbito de la seguridad pública. Estos casos demuestran que los sesgos de género en la IA no son solo un inconveniente técnico, sino un problema con implicaciones sociales significativas.

La solución pasa por un esfuerzo conjunto de los desarrolladores, reguladores y la sociedad en general, para construir responsablemente una inteligencia artificial que no solo sea eficiente, sino que también respete y avance en equidad.