**Tecnológico de Costa Rica**

**Escuela de computación**

**La equidad en la inteligencia artificial:  
un desafío ineludible**

**Ensayo presentado por**  
Matias Benavides Sandoval

**Curso:** Comunicación escrita  
**Profesora:** Natalia Rodríguez Herra

Cartago, Costa Rica  
 24 de abril de 2025

La equidad en la inteligencia artificial: un desafío ineludible

La inteligencia artificial (IA) se ha convertido en una transformación que ha impactado desde el ámbito médico y la aplicación de justicia hasta la concesión de créditos y la gestión financiera de una manera muy rápida “…tecnología de vanguardia reciente, que apenas se discute su utilización, el enfoque ético que debe tener y sus posibles repercusiones económicas, políticas y culturales. La IA está en desarrollo. Apenas florecen sus amplias y diversas aplicaciones, su utilidad y sus riesgos.” (Guzmán Gómez, 2024, p.17) Su capacidad para procesar cantidades de datos gigantescos y de aprender de patrones complejos promete una mayor eficiencia y personalización de muchos servicios (Zesergio MELO, 2024), esto puede conllevar riesgos cuando esos datos contienen prejuicios históricos o carecen de representatividad. Por ello, resulta esencial analizar cómo la IA afecta a la sociedad y por qué debe abordarse su equidad desde múltiples perspectivas. El uso de la IA nos planea preguntas sobre el cómo esta misma puede llegar a afectar a la sociedad moderna.

Garantizar la equidad en los sistemas de IA significa crear programas informáticos que impidan decisiones discriminatorias y promuevan la transparencia hacia el público. Como cuando un algoritmo evalúa solicitudes de crédito o sugiere diagnósticos médicos “está transformando el ámbito de la atención sanitaria al incrementar la precisión diagnóstica, personalizar los tratamientos y optimizar los resultados clínicos” (Zesergio MELO, 2024, 49), no basta con únicamente garantizar precisión: debe ser justo en su trato a todas las poblaciones. La falta de criterios claros y transparencia en los mismos en la detección y corrección de sesgos puede llevar a el rechazo injustificado de préstamos a personas de comunidades vulnerables o en diagnósticos menos precisos para minorías, lo que nos da la alerta de establecer reglas y métodos efectivos para evitar casos de este estilo y reducir al máximo el sesgo a diferentes grupos.

La justificación de este enfoque es simple: la IA ya se maneja en sectores como la salud, la justicia y las finanzas, donde un fallo algorítmico es crítico y puede llegar a ocasionar perjuicios graves a quienes dependen de estas decisiones automatizadas. “La falta de transparencia en los algoritmos de IA puede generar desconfianza y preocupación entre los estudiantes, profesores y otras partes interesadas” (Maldonado Zuñiga et all. 2024, p.73) Para minimizar estos riesgos, es imprescindible definir marcos regulatorios, protocolos y prácticas de diseño que ayuden a una inteligencia artificial responsable “falta de infraestructura tecnológica, la escasez de personal capacitado y la necesidad de políticas claras que regulen el uso de la IA”.( Maldonado Zuñiga et all, 2024. p.72) Entender las posibles raíces del problema, evaluar sus consecuencias y proponer soluciones son las bases sobres las que debe sustentarse la aproximación a la equidad algorítmica, desde el por qué surgen los sesgos en la IA, qué efectos tienen en la toma de decisiones automatizada y, finalmente, qué podemos hacer para construir sistemas más justos y transparentes

**I. Los algoritmos pueden ser injustos y aumentar las desigualdades**

Aunque prometen objetividad, los algoritmos pueden reproducir desigualdades cuando se entrenan con datos que las reflejan. En muchos casos, los datos incluyen patrones de exclusión social o económica. Al aprender de estas fuentes, un modelo de IA termina reforzando las mismas barreras que pretendía superar

A esto hay que agregar la falta de diversidad en equipos de desarrollo: cuando todos los involucrados en la creación de sistemas algorítmicos tienen visiones similares en cuanto a género, etnia o formación, es fácil pasar por alto sesgos. La falta de perspectivas diferentes fomenta la creación de tecnológicas mas “opacas” y cerradas ya que no hay diversidad de opiniones ni mucho menos una que los haga plantear ciertas decisiones de diseño en los algoritmos o fuentes de los datos.

Lo que a su vez puede llegar a generar que estos algoritmos creados de manera cerrada no sean transparentes ni mucho menos claros para el público lo que genera desconfianza en los mismos desde todos los sectores “un error en un sistema de IA podría reducir la confianza tanto del público como de los profesionales de la salud, afectando negativamente la aceptación de estas tecnologías y frenando su desarrollo.” (Zesergio MELO 2024 p.52)

La transparencia y la rendición de cuentas representan pilares éticos fundamentales al implementar tecnologías de IA en el ámbito de la salud. Con el aumento de su uso en entornos clínicos, es esencial que tanto los procesos de decisión como los datos que sustentan estas herramientas sean accesibles y explicables para todos los actores relevantes. Las instituciones deben garantizar que el funcionamiento de los algoritmos de IA sea interpretable y que se establezcan mecanismos de responsabilidad claramente definidos ante sus resultados. Esta claridad no solo facilita decisiones clínicas más fundamentadas, sino que también refuerza la credibilidad de la IA, fomentando la confianza tanto en los profesionales de la salud como en los pacientes. (Zesergio MELO 2024)

Las consecuencias ya se han estado notando y son reales y preocupantes. En el ámbito financiero, bancos han visto cómo sus sistemas automáticos rechazaban préstamos de forma desproporcionada a solicitantes de determinadas zonas o con ciertos apellidos, replicando prejuicios socioeconómicos y perpetuando la exclusión. De igual modo, en salud se han visto como herramientas de apoyo al diagnóstico que muestran menor sensibilidad al identificar enfermedades en pacientes de piel más oscura o de origen cultural distinto, debido a la insuficiente representación de estos grupos en los datos de entrenamiento.

Este panorama evidencia la necesidad urgente de implementar métodos de auditoría continua, promover la diversidad en los equipos de desarrollo y fomentar la transparencia de los algoritmos. Solo así podremos aprovechar el potencial de la IA sin comprometer la justicia ni los derechos de las personas más vulnerables.

Bibliografía:

Guzmán Gómez (2024) Inteligencia artificial y equidad de género. Una perspectiva histórica de los sesgos culturales y su impacto en la relación humana con las tecnologías de la información y comunicación. *Sintaxis revista científica del centro de la investigación para la comunicación aplicada*, 8, 14-30

Zesergio MELO (2024) Inteligencia artificial en salud: desafíos éticos para lograr la aplicación de las tecnologías a la salud del paciente. *Dignidad humana e inteligencia artificial: aplicaciones éticas para una armonía futura* 18, 49-62

Maldonado Zuñiga et all, (2024) Desafíos universitarios: equidad, transparencia, responsabilidad y oportunidades en el uso de la Inteligencia Artificial. *Serie Científica de la Universidad de las Ciencias Informáticas* 17 (5) 70-83