Katherine Michelle Ortiz Burgos
Código: 506221041
Profesor: Oscar Alexander Mendez Aguirre
Fundación Universitaria Konrad Lorenz

Coded Bias

I. Resumen

Este artículo explora los desafíos éticos y técnicos de la Inteligencia Artificial (IA) en el diseño de interfaces de usuario (UI/UX), tomando como base el documental ”Coded Bias”. Se puede observar cómo el documental resalta problemas como el reconocimiento facial y la discriminación, y se resaltan estos temas con un caso real de implementación de IA en UI/UX. A través de un análisis ético y técnico, se reflexiona sobre cómo abordar estos desafíos de manera ética y efectiva en el diseño de IA.

II. Introducción

El cruce entre la ética y la tecnología, especialmente en el ámbito de la IA aplicada al diseño de interfaces, es un área con constante aumento de interés. El documental ”Coded Bias” arroja luz sobre estos temas al explorar cómo la IA puede perpetuar sesgos y discriminación, especialmente en el reconocimiento facial y la vigilancia. Este artículo analiza cómo estos problemas se reflejan en un caso real de aplicación de IA en diseño de interfaces, y cómo los diseñadores pueden abordarlos de manera responsable.

III. Desarrollo

En ”Coded Bias” se presentan desafíos importantes en el diseño de UI/UX relacionados con la IA. Por ejemplo, se muestra cómo el reconocimiento facial puede ser impreciso, especialmente para personas de color, debido a la falta de diversidad en los conjuntos de datos. Además, resalta cómo la IA puede usarse en vigilancia masiva, lo que puede llevar a la discriminación racial y violaciones de los derechos civiles.

En el documental, también existen numerosos casos en la vida real que reflejan estos desafíos éticos y técnicos. Por ejemplo, el sistema de seguridad en Londres, que utiliza tecnología de reconocimiento facial, ha sido criticado por su falta de precisión en la identificación de personas de color, lo que genera preocupaciones sobre la discriminación y la violación de derechos. Del mismo modo, el uso de IA en aplicaciones gubernamentales, como la emisión de cédulas de identidad, puede resultar en exclusiones injustas y negación de servicios a ciertos grupos étnicos o raciales. Estos casos reales ilustran cómo los problemas presentados en este documental se manifiestan en la práctica y subrayan la importancia de abordar estos desafíos de manera ética y efectiva en el diseño de interfaces de IA.

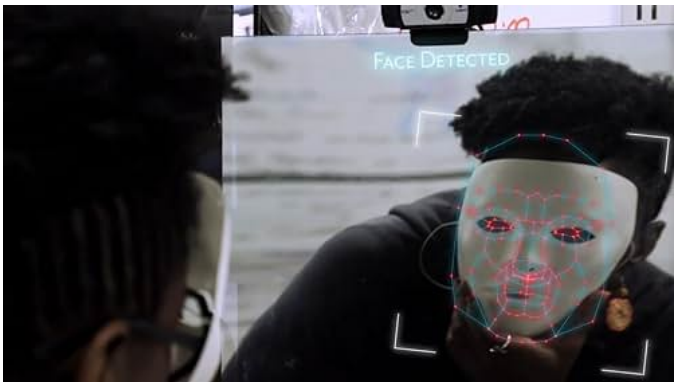


Imagen 1. Coded Bias.

IV. Análisis Ético

Al evaluar los casos presentados en ”Coded Bias” y los casos reales seleccionados, surge una serie de preocupaciones éticas importantes. Por ejemplo, la falta de precisión en los sistemas de reconocimiento facial puede resultar en discriminación y exclusión injusta de ciertos grupos. Del mismo modo, el uso de IA puede dar lugar a violaciones de derechos civiles y negación de servicios básicos a personas inocentes. Esto destaca la necesidad de una evaluación crítica y una supervisión cuidadosa de los sistemas de IA para garantizar que sean justos, equitativos y respetuosos de los derechos individuales.

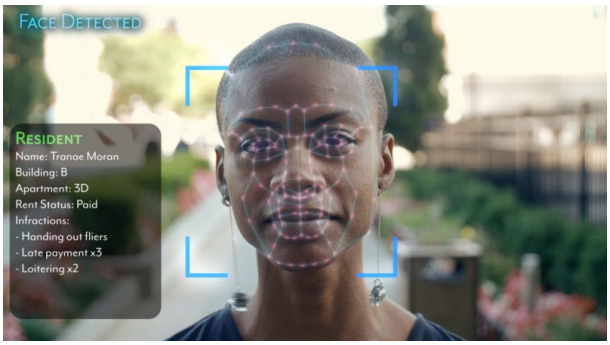
1. Caso de Amazon: En ”Coded Bias”, se expone el caso de Amazon, donde se reveló que su algoritmo de contratación estaba sesgado contra las mujeres. A pesar de que el sistema estaba diseñado para eliminar prejuicios, terminó discriminando a las candidatas mujeres. Este ejemplo destaca cómo los algoritmos de IA pueden heredar sesgos presentes en los datos utilizados para entrenarlos, lo que lleva a decisiones injustas y discriminatorias.
2. Cédulas y Reconocimiento Facial: Otro caso real, es el ejemplo de cómo el uso de tecnología de reconocimiento facial en aplicaciones gubernamentales, como la emisión de cédulas de identidad, puede tener consecuencias negativas para ciertos grupos de personas. En algunas ocasiones, se ha demostrado que estos sistemas tienen dificultades para reconocer con precisión a personas de ciertos

grupos étnicos o raciales, lo que puede resultar en la exclusión o la negación de servicios a estas poblaciones.

3. Perjuicio Injusto de la IA: Otro aspecto crítico explorado en el documental es la capacidad de la IA para perjudicar injustamente a personas inocentes debido a sus fallas inherentes. Los sistemas de IA, aunque se presentan como objetivos y neutrales, pueden estar sujetos a errores y sesgos que pueden tener consecuencias devastadoras para individuos y comunidades. Por ejemplo, un sistema de reconocimiento facial defectuoso puede identificar erróneamente a una persona como sospechosa de un delito, lo que podría resultar en su arresto injustificado o en la pérdida de oportunidades laborales. Esta situación resalta la necesidad de una evaluación crítica y una supervisión cuidadosa de los sistemas de IA para mitigar los riesgos de perjuicio injusto y asegurar la equidad y la justicia en su aplicación[4].

Imagen 3. Sesgos Inherentes.

4. Seguridad en Londres: Otro ejemplo citado en el documental es el sistema de seguridad en Londres, que emplea tecnología de reconocimiento facial para identificar a posibles delincuentes. Sin embargo, se demostró que este sistema era menos preciso en personas de color, lo que llevó a una discriminación sistemática contra estas comunidades. Esta situación representa cómo los algoritmos de IA pueden tener resultados desproporcionadamente negativos en ciertos grupos demográficos, lo que plantea preocupaciones éticas y legales sobre el uso de esta tecnología.
5. Necesidad de control en la IA: En ”Coded Bias”, se examina el tema del control y la vigilancia a través del ejemplo de Japón, donde se han implementado sistemas de IA para monitorear el comportamiento de los ciudadanos. Estos sistemas utilizan tecnologías de reconocimiento facial y análisis de comportamiento para identificar y clasificar a las personas en función de su nivel de ”sociabilidad”. Aunque se presenta como una medida para garantizar la seguridad pública, este enfoque plantea preocupaciones sobre la invasión de la privacidad y el potencial de discriminación y control social[2].
6. Caso de IBM: Se revela cómo los algoritmos de selección de currículums utilizados por algunas empresas, como IBM, pueden estar sesgados hacia ciertos grupos demográficos. Por ejemplo, se expone el caso de una empresa que rechazaba sistemáticamente los currículums de candidatos que habían asistido a una institución educativa exclusivamente para mujeres. Esta discriminación implícita refleja cómo los algoritmos pueden perpetuar sesgos históricos y culturales, lo que lleva a una exclusión injusta. Además, se ilustra cómo la tecnología de reconocimiento facial se ha utilizado en entornos laborales para monitorear y evaluar el desempeño de los empleados. Por ejemplo, se menciona el caso de profesores que fueron despedidos de sus empleos debido a la lectura facial incorrecta de las emociones de los estudiantes[3].



figureControl, Vigilancia.

V. Problemas de control en el desarrollo de la IA

1. La falta de claridad en el desarrollo de algoritmos de IA puede llevar a un problema de control, ya que las personas afectadas por las decisiones de estos algoritmos no tienen conocimiento sobre cómo se toman esas decisiones ni quién es responsable de ellas. Esto puede generar desconfianza y frustración entre los usuarios, quienes pueden sentir que están siendo controlados por sistemas opacos y poco regulados.
2. El problema de la vigilancia masiva está estrechamente relacionado con el control, ya que implica la recopilación y el monitoreo de información personal de manera extensiva y a menudo sin consentimiento. Esto puede ser percibido como un intento de ejercer control sobre la población, lo que plantea preocupaciones sobre la invasión de la privacidad y el abuso de poder por parte de quienes tienen acceso a los datos recopilados.
3. La inequidad en el acceso y uso de la tecnología de IA puede perpetuar el control por parte de aquellos que tienen privilegios y recursos para desarrollar y utilizar estos sistemas. Las comunidades marginadas pueden ser objeto de un control injusto o discriminatorio si se ven excluidas de las oportunidades y servicios que ofrecen los avances en IA. Esto puede agravar las disparidades sociales y económicas existentes, perpetuando un ciclo de control y desigualdad.
4. La falta de diversidad en el desarrollo de IA puede contribuir a un problema de control al perpetuar sesgos

y perspectivas limitadas en los sistemas de IA. Si los equipos de desarrollo no representan adecuadamente la diversidad de la sociedad, es más probable que pasen por alto ciertos problemas éticos y técnicos, lo que puede llevar a la creación de sistemas de IA que ejercen un control injusto o discriminatorio sobre ciertos grupos de personas[5].

VI. Conclusiones

1. El análisis ético de la inteligencia artificial es fundamental para identificar y abordar preocupaciones sobre la equidad, la justicia y los derechos individuales. Es crucial realizar evaluaciones éticas continuas para garantizar que los sistemas de IA sean diseñados y utilizados de manera responsable.
2. El uso de la inteligencia artificial debe realizarse con un enfoque en la protección de la privacidad y los derechos individuales. Las medidas de seguridad y las regulaciones sólidas son necesarias para proteger la información personal y garantizar que el uso de la IA respete los derechos fundamentales de las personas.
3. Es fundamental considerar la diversidad y la equidad en todas las etapas del desarrollo de la inteligencia artificial. La inclusión de perspectivas diversas puede ayudar a identificar y mitigar sesgos inherentes en los datos y algoritmos, promoviendo así sistemas más justos y equitativos.

VII. Referencias

1. Kantayya, S. (2021, abril 5). Prejuicio cifrado.
2. BBC News Mundo. (2017, diciembre 26). China, el Estado que todo lo ve: así es la red de videovigilancia más grande y más sofisticada del mundo. BBC. https://www.bbc.com/mundo/noticias-internacional-42398920
3. ¿Qué es el sesgo de la IA? (2024, abril 17). Ibm.com. https://www.ibm.com/es-es/topics/ai-bias
4. Moore, P. V. (s/f). Inteligencia artificial en el entorno laboral. Desafíos para los trabajadores. OpenMind. Recuperado el 18 de mayo de 2024, de https://www.bbvaopenmind.com/articulos/inteligencia-artificial-en-entorno-laboral-desafios-para-trabajadores/
5. Pike, S. (2017, agosto 31). Por qué no basta con el aprendizaje automático. Kaspersky. https://www.kaspersky.es/blog/ai-fails/14276/