

Sesgos algorítmicos en educación: análisis y soluciones

Rubén Comas-Forgas

Universidad de las Islas Baleares, España

Resumen

El proyecto CHARLIE “*Challenging Bias in Big Data Used for AI and Machine Learning*”, auspiciado por la Unión Europea persigue el diseño, desarrollo y validación de un conjunto de recursos y materiales de carácter tanto teórico como práctico para el análisis crítico de los sesgos algorítmicos entre el alumnado universitario en formación y los docentes de Educación Superior. Tiene como objetivo primordial la optimización de las capacidades de aprendizaje digital y el fortalecimiento de competencias sociales y éticas entre el alumnado y profesorado de cualquier área de conocimiento.

El sesgo algorítmico surge cuando los algoritmos generan resultados que presentan prejuicios sistemáticos, los cuales tienen su origen en suposiciones erróneas que son intrínsecas al proceso de aprendizaje automático de las aplicaciones y programas que los soportan. Estos sesgos pueden ser muy variados y producirse por diversas causas, como por ejemplo: el resultado de conjuntos de datos sesgados utilizados durante la fase de entrenamiento, preconcepciones por parte de los desarrolladores, o el uso de una muestra inadecuada durante las etapas de diseño y toma de decisiones en la creación de algoritmos (Akgun & Greenhow, 2022; Florida State University, 2022).

Los sesgos algorítmicos en la educación pueden tener repercusiones profundas, ya que los sistemas educativos cada vez más dependen de la inteligencia artificial para la toma de decisiones. Cuando los algoritmos están sesgados, pueden perpetuar desigualdades y discriminaciones preexistentes. Por ejemplo, un algoritmo que recomienda recursos educativos o admisiones a programas podría favorecer inadvertidamente a estudiantes de ciertos grupos sociodemográficos sobre otros, basándose en los datos con los que fue entrenado. Esto no solo afecta las oportunidades de los estudiantes afectados sino que también refuerza estereotipos y brechas educativas. La confianza en datos sesgados para evaluar el rendimiento o personalizar la enseñanza puede desviar recursos de quienes podrían necesitarlos más, perpetuando ciclos de desventaja.

Por tanto, es crucial identificar y corregir sesgos algorítmicos para asegurar una educación equitativa y justa para todos. En esta ponencia exploraremos las características intrínsecas, los riesgos inherentes y las estrategias potenciales de mitigación ante los sesgos algorítmicos dentro del ámbito educativo.

Palabras clave: *sesgo algorítmico, equidad educativa, inteligencia artificial, mitigación de sesgos, analítica de datos en educación.*

Agradecimientos: Esta ponencia es el resultado del proyecto *CHARLIE - Challenging Bias in Big Data used for AI and Machine Learning* (Referencia: 2022-1-ES01-KA220-HED-000085257) financiado por la Unión Europea.

Algorithmic biases in education: analysis and solutions

Abstract

The CHARLIE project "Challenging Bias in Big Data Used for AI and Machine Learning," sponsored by the European Union, aims at the design, development, and validation of a set of both theoretical and practical resources and materials for the critical analysis of algorithmic biases among university students in training and higher education teachers. Its primary objective is to optimize digital learning capabilities and strengthen social and ethical competencies among students and faculty in any field of knowledge.

Algorithmic bias occurs when algorithms generate results that exhibit systematic prejudices, which originate from incorrect assumptions inherent in the machine learning process of the applications and programs that support them. These biases can be varied and caused by various factors, such as the result of biased data sets used during the training phase, preconceptions on the part of developers, or the use of an inadequate sample during the design and decision-making stages in the creation of algorithms (Akgun & Greenhow, 2022; Florida State University, 2022).

Algorithmic biases in education can have profound repercussions, as educational systems increasingly rely on artificial intelligence for decision-making. When algorithms are biased, they can perpetuate existing inequalities and discrimination. For instance, an algorithm that recommends educational resources or admissions to programs might inadvertently favor students from certain sociodemographic groups over others, based on the data with which it was trained. This not only affects the opportunities of the impacted students but also reinforces stereotypes and educational gaps. Relying on biased data to assess performance or personalize teaching can divert resources from those who may need them most, perpetuating cycles of disadvantage.

Therefore, it is crucial to identify and correct algorithmic biases to ensure equitable and just education for all. In this presentation, we will explore the intrinsic characteristics, inherent risks, and potential mitigation strategies against algorithmic biases within the educational domain.

Keywords: algorithmic bias, educational equity, artificial intelligence, bias mitigation, educational data analytics.

Acknowledgments: This presentation is the outcome of the *CHARLIE project - Challenging Bias in Big Data used for AI and Machine Learning* (Reference: 2022-1-ES01-KA220-HED-000085257) funded by the European Union.

Referencias

Akgun, S., & Greenhow, C. (2022). Artificial intelligence in education: Addressing ethical challenges in K-12 settings. *AI Ethics* 2, 431 – 440. <https://doi.org/10.1007/s43681-021-00096-7>

Florida State University. (2022). *Algorithm Bias*. Consultado el 05/11/2023 desde <https://guides.lib.fsu.edu/algorithm>

Kordzadeh, N., & Ghasemaghaei, M. (2022). Algorithmic bias: review, synthesis, and future research directions. *European Journal of Information Systems*, 31(3), 388-409. <https://doi.org/10.1080/0960085X.2021.1927212>