

Social Impacts Literature Search Assignment

Name: Aryan Raval

Full citation:

Sun, W., Nasraoui, O., & Shafto, P. (2020). Evolution and impact of bias in human and machine learning algorithm interaction. PLOS ONE, 15(8), e0235502.
<https://doi.org/10.1371/journal.pone.0235502>

Short summary:

The article examines how unchecked information in machine learning algorithms can cause algorithmic bias, leading to inequality and limited human ability to discover relevant data. An iterated-learning framework is proposed to study three forms of algorithmic bias, with findings showing that iterated filter bias has the most significant effect on class-1-blindspot size. The framework provides an effective analysis of the impact of iterated algorithmic bias in human-algorithm interaction.

Nature of actual or potential social impact:

The unchecked information in machine learning algorithms can lead to algorithmic bias, negatively affecting society by perpetuating and exacerbating social inequalities.

Clarifying question:

How can we ensure that human and algorithmic decision-making processes are transparent, accountable, and aligned with social values and ethical principles?

Substantive critique:

The authors did not address the potential impact of using their proposed iterated-learning framework on privacy and security concerns related to collecting and analyzing users' data, which is crucial in the context of personalized user interfaces.