DLI Accelerated Data Science Teaching Kit

Lecture 4.3 Challenges for Underrepresented Groups Relating to Data Ethics and Bias

Data and Algorithm Biases in AI

The emergence and growth of research on issues of ethics in AI, and in particular algorithmic fairness comes from the data used to train predictive models and in the design of objective functions.

Design of unbiased algorithms and fair socio-technical systems are key desired outcomes which depend on practitioners from the fields of data science and computing.

Data from the underrepresented groups show that they continue to be subject to systemic structural biases, often manifesting at data collection, that can skew the outcome of automated decision-making processes.

Indeed, the interdisciplinary community that has recently arisen to address these biases in algorithmic design and deployment has made great strides in identifying unfairness and working to ad- dress it from a computational perspective.

Source:Kuhlman, Caitlin, Latifa Jackson, and Rumi Chunara. "No computation without representation: Avoiding data and algorithm biases through diversity." arXiv preprint arXiv:2002.11836 (2020).

Why Improving Diversity is Essential to the Ethics in AI Community

Bias in data and algorithms are critical issues, and efforts to address these are essential as computing researchers and practitioners design models and algorithms that are being deployed in ever more real-world scenarios.

Structural inequality is a condition where one category of people are attributed an unequal status in relation to other categories of people, and this relationship is perpetuated and reinforced by a confluence of unequal relations in roles, functions, decisions, rights, and opportunities.

Impact of structural inequality on algorithmic fairness analysis

Impact of structural inequality on the computing community

Source:Kuhlman, Caitlin, Latifa Jackson, and Rumi Chunara. "No computation without representation: Avoiding data and algorithm biases through diversity." arXiv preprint arXiv:2002.11836 (2020).

Fairness in the Literature and Possible Confounding

Bias in data: due to a lack of representation of both female faces and dark-skinned faces in the training datasets used, prediction rates by these commercial systems suffered greatly for these groups.

Algorithmic Bias: at the same health risk score, black patients are considerably sicker than whites due to the way the risk score is attributed to different illnesses that occur disparately.

Missing Labels: missing data; specifically when membership labels for a protected class are unavailable.

Source:Kuhlman, Caitlin, Latifa Jackson, and Rumi Chunara. "No computation without representation: Avoiding data and algorithm biases through diversity." arXiv preprint arXiv:2002.11836 (2020).

Recommendations for Increasing Diversity Within Computing and Ethics in AI Community

Building collaborations with minority-serving institutions

Prioritizing research collaboration between the Ethics in AI community and underrepresented/interdisciplinary groups

Providing enhanced mentorship to trainees at Ethics in AI research conferences

Source:Kuhlman, Caitlin, Latifa Jackson, and Rumi Chunara. "No computation without representation: Avoiding data and algorithm biases through diversity." arXiv preprint arXiv:2002.11836 (2020).

DEEP
LEARNING
INSTITUTE

PRAIRIE VIEW
A&M UNIVERSITY

DLI Accelerated Data Science Teaching Kit

Thank You