Egaleco Advancing Fairness in Machine Learning

Product & Policy Team Gurpreet Kaur Khalsa, Alan Kyle, Mudit Mangal & Orissa Rose



The Problem, Our Motivation

COMPUTING

Racial Bias Found in a Major Health Care Risk Algorithm

Black patients lose out on critical care when systems equate health needs with costs

A STAT INVESTIGATIO

Denied by AI: How Medicare Advantage plans use algorithms to cut off care for seniors in need

By <u>Casey Ross</u> I and <u>Bob Herman</u> I March 13, 2023

Forbes

FORBES > LEADERSHIP > DIVERSITY, EQUITY & INCLUSION

AI Bias Could Put Women's Lives At Risk - A Challenge For Regulators

Al in Healthcare can create disparate impacts

How will **YOU** build algorithms that don't perpetuate or amplify historical biases?

Fairness Audit

- Several dozen fairness metrics
- Impossible to satisfy them all
- But before all that... how do we even define fairness?

Metric #1,284. What are we measuring again? Right.

Okay, the True Positives divided by the False Positives, multiplied by the total number of Negative Predictions, plus the temperature of the room, multiplied by the negative exponential of the number of words in this sentence, should be the same for all sensitive groups.

Fairness.



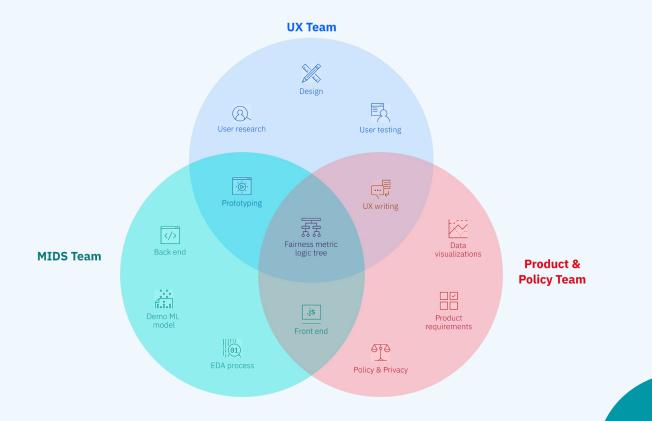
Inspired by the needs and habits of ML practitioners and fairness experts.

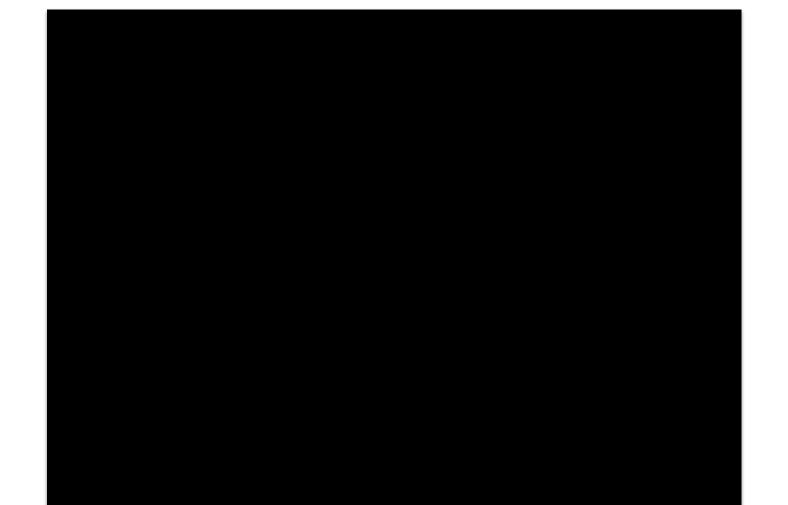
Egaleco \approx

Contextualize Fairness Visualize disparities and generate a Fairness Report

Resources for Fair ML

Interdisciplinary Collaboration



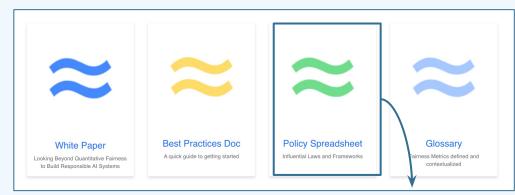


| \sim | \sim | \approx | \sim |
|--|--|--------------------|---|
| White Paper Looking Beyond Quantitative Fairness to Build Responsible AI Systems | Best Practices Doc A quick guide to getting started | Policy Spreadsheet | Glossary Fairness Metrics defined an contextualized |

- Tooling is not a panacea
- Legal fairness: anti-discrimination approach
 - Compliance ≠ holistic fairness
 - Procedural fairness can obscure harms of "neutral" systems
- Social fairness: considers context and impact
 - Promotes system transparency and explainability
 - Problematizes rigid classification



- Practitioners need support operationalizing fairness
- Resources to distribute ownership
- common vocabulary & existing resources
- Start projects with "responsibility sprints"
- Interdisciplinary, interative approach



| Name of law/regulation/principle 👳 | Enforcement or proposing = entity | Notes ÷ | Enforceability = | Jurisdiction = | Effective date | - Link |
|--|---|--|------------------|----------------|----------------|---------------------|
| California Consumer Privacy Act (as amended by the California Privacy Rights Act) | California Privacy Protection Agency + CA Attorney General | The CPRA is an amendment to the CCPA which includes language about "automated decision-making". It mentions automated decision making in the context of profiling and gives the the CA Privacy Protection Agency the ability to enforce and make more takes regaring automated decision making in the future. Note that data covered under IHPA is not covered by CCPA and CPRA. However, CCPA auto-Chich chipate the such care of a future providers to provide services to businesses subject to CCPA and CPRA. If your company conducts "profiling" as defined in the law, you already need to comply with these rules. The law states that automated decision-making in the economic situation, hailin, personal preferences, interests, reliability, behavior, location or movements." The rights granted to Californian are to: Know about the profiling rougilities and them and how it is used and shared; To delete personal information a business collects about them and how dis used and shared; Chi distation of them personal information; and hord discrimination in the service and the service and and the service and and the service and information; and hord discrimination in the service and and the service and the service and information; and hord and information; and hord discrimination in the service and the service and and and the service and the service and the service and the service a | | California | 1/1/2023 | https://www.capr |
| Confidentiality of Medical Information Act (CMIA) | California Attorney General's Office | Protects patient record confidentiality. Prohibits health care providers from disclosing medical information without first obtaining consent, and requires covered health care providers that create, maintain, store or destroy medical information in a manner that preserves the confidentiality. | Law | California | 1/1/2009 | https://leginfo.leg |
| Genetic Information Nondiscrimination Act | Equal Employment Opportunity Commission + Department of Justice | The Genetic Information Nondiscrimination Act (GNA) is a federal law that prohibits discrimination by health insures and employers based on an individual's genetic information. Note that there is a patchwork of state laws protecting Americans from genetic discriminator, though they vay widely in the scope, applicability, and amount of protection provided. GNA sets a floor of minimum protection against genetic discrimination and does not preempt state laws with stricter protections. | Law | USA | 11/21/2009 | https://www.hhs.e |
| Health Information Technology for Economic and Clinical Health (HITECH) Act | Department of Health and Human Services + Office for Civil Rights | Mandates audits of healthcare providers and data handlers to ensure compliance with HIPAA's information privacy and security requirements. | Law 🔹 | USA | 2/18/2009 | https://www.hhs. |

Conclusion

Egaleco, a novel approach to:

- Context specific guidance for Healthcare
- A new system of ranking fairness metrics
- Policy and Ethics resources



Acknowledgements



Daniel Aranki Assistant Prof. UC Berkeley



Marti Hearst Prof. UC Berkeley & Head of I School



Jared Maslin COO & Senior Director, Data Privacy



Nick Merill Research Fellow, CLTC



Irene Chen Assistant Prof, UCB & UCSF



Kit Rodalfa Research Director, Reglab, Stanford

Thank you

hello@egaleco.ai egaleco.herokuapp.com

