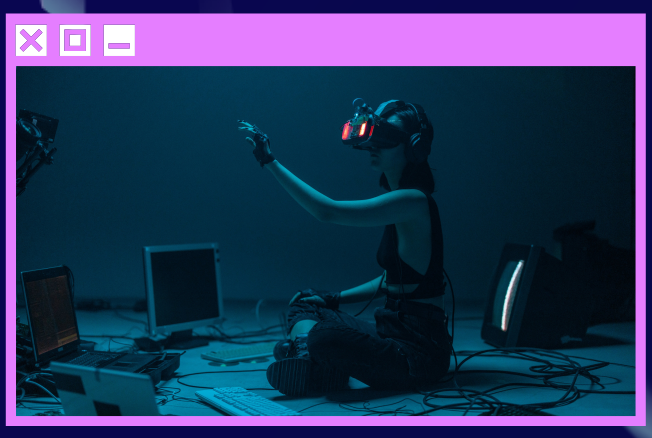
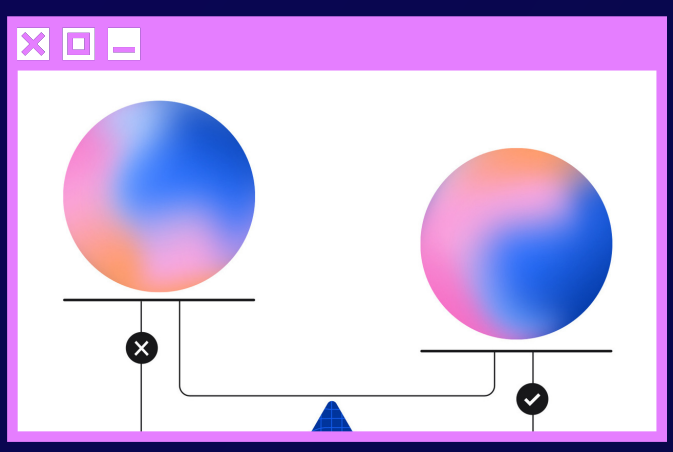
Beginner's Guide

AI FAIRNESS

INTRODUCTION TO FAIRNESS IN AI

Fairness in AI ensures decisions made by algorithms are equitable for all users, actively mitigating biases to maintain trust.



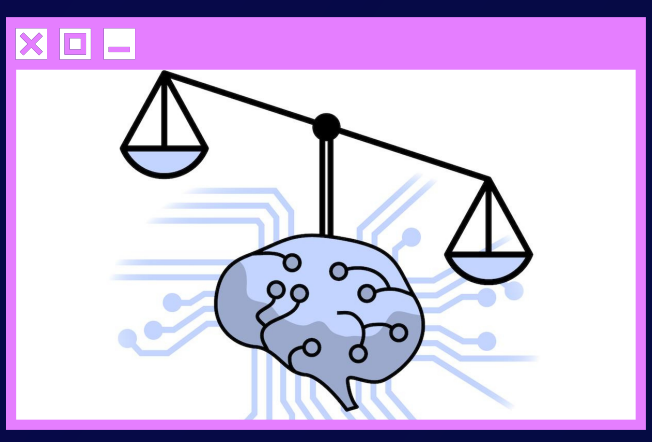


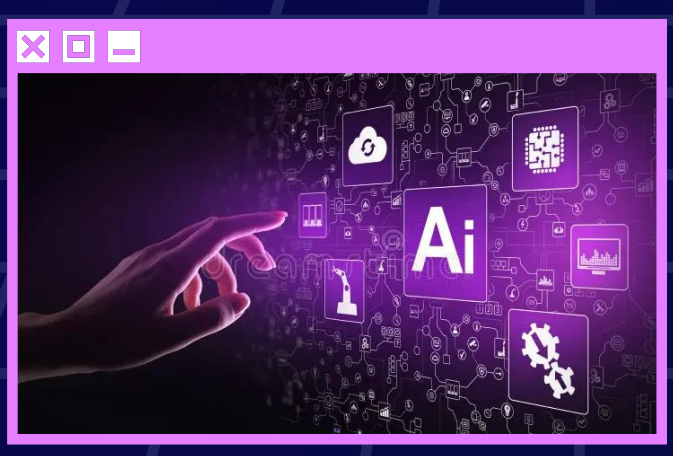
FAIRNESS METRICS EXPLAINED

Fairness metrics in AI ensure equitable decision-making across user groups, using benchmarks like Calibration and Equalized Odds to identify and mitigate biases. Learn more on this link

EXAMPLES OF FAIRNESS METRICS

Demographic Parity equalizes outcomes regardless of group qualifications, while **Equal Opportunity** and **Equality of Odds** ensure consistent model accuracy across groups, balancing fairness in outcomes and error rates.



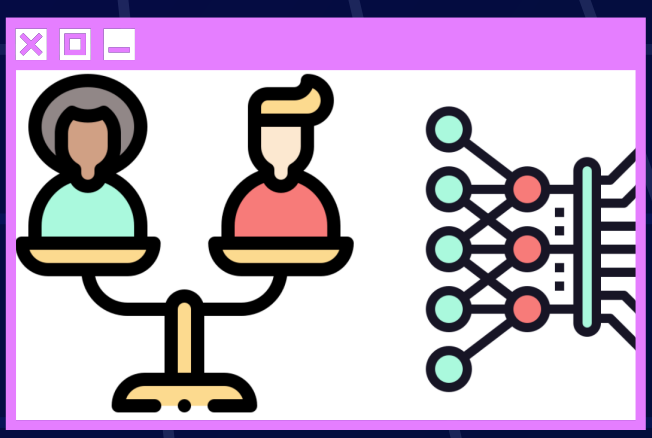


Real-World Impacts of Group Unfairness

In hiring, AI systems have perpetuated bias by favoring resumes from male candidates, notably in tech roles, due to historical data trends. Loan approval algorithms have disadvantaged minority groups by using zip codes as proxies for race, leading to lower approval rates.

Best Practices for Achieving Group Fairness

Ensure group fairness by adopting fair data practices, conducting bias audits, and integrating fairness constraints into algorithms. Continuous oversight and a diverse development team are essential for identifying and addressing biases effectively.Learn more on this link



AI FAIRNESS IN MY PROJECT

INTRODUCTION TO MY PROJECT IDEA

My project idea is to create a simple tool that helps universities manage their space better. By using pictures to check if classrooms are empty or full, it makes it easier for students and staff to find available spaces for studying or meetings, ultimately making campus life more efficient.

FAIRNESS METRICS USED IN MY PROJECT

Calibration ensures that my predictions are equally reliable across all demographic groups, providing fair and accurate probabilities for correct classification to every individual, regardless of their background.
Predictive Parity: With this AI fairness metric, I guarantee equal accuracy for positive predictions across all demographic groups, ensuring fairness in the model's positive outcomes for everyone.
Equalised Odds: I ensure equal true positive and true negative rates across demographics, guaranteeing fair and accurate classification for every user, without bias towards any group.

IDENTIFIED BIASES

I aim to address **two key biases** in my project: firstly, the inadequate representation of people with disabilities in the dataset, and secondly, the limited diversity in terms of racial and ethnic backgrounds. By doing so, my goal is to significantly improve model fairness and accuracy.

MITIGATING BIAS STRATEGIES

Mitigating biases is crucial for fair AI. My project emphasizes inclusivity by enhancing representation of people with disabilities and increasing racial/ethnic diversity in datasets. Acknowledging limited images of people with disabilities and diversity, I am dedicated to rectifying these gaps for a more equitable and accurate model.

IMPACT ON ORGANIZATIONAL LIFE

My tool transforms space management, enabling efficient scheduling and use of study rooms. It's not just about occupancy detection; it's about creating a more inclusive and productive educational environment.

CALL TO ACTION: TOWARDS A FAIRER AI

AI fairness is crucial. Every one should make the best effort when shaping technology that's equitable for all, paving the way for a just digital world.