

The ML Fairness Project

Home About

Vanguard Projects

Resources

Presentations

Reports

Community

FAQ

Archive

GET INVOLVED

Resources

Algorithmic unfairness definition

The definition articulates the range of algorithmic unfairness that can occur in products, and establishes a shared understanding and language for internal efforts related to this issue.

Metrics and Measurements for Fairness

Google strives for algorithmic fairness across products. This document starts to outline the nitty-gritty of how we can measure and bolster optimally equal experiences for users—at the level of the math, algorithms, and code.

ML Fairness Communication Best Practices

Describing best practices for communicating on topics related to the inclusion, fairness, and transparency of machine learning algorithms; guidance for creating, reviewing, and sharing documents related to ML Fairness; and when and how to maintain attorney-client privilege.

Comms/PR Best Practices for Speaking and Posting about ML Fairness

Planning to present or write about ML Fairness? As an Al-first company, Google aims to develop the benefits of machine learning for everyone. Building inclusive algorithms, datasets, and products is crucial to this mission. The PR/Comms team provides guidance on how to best deliver your message.

Model Understanding

Whether seeking to improve performance of ML systems, combat problems with ML fairness, conduct research into model behavior, or deploy ML in products in a non-trival way, gaining better understanding of models can be extremely valuable. While there is no one-size-fits-all approach and many open areas of research, there are also many useful tools that exist today.

Fairness Tools

A growing list of available tools to help train, interpret, visualize, and debug models in service of fairness.





Glassbox

The Glassbox group has developed ML algorithms to help ensure fairness.

Equality of Opportunity in Machine Learning

At the heart of the approach is the idea that individuals who qualify for a desirable outcome should have an equal chance of being correctly classified for this outcome.

ML Test Certified

Based on years of prior experience using ML at Google, in systems such as ad click prediction and the Sibyl ML platform, the TFX team have developed a set of best practices for using machine learning systems. They present these practices as a set of actionable tests, and offer a scoring system to measure how ready for production a given machine learning system is.

GET INVOLVED