Computational Journalism: Assignment 3

Pankhuri Kumar (pk2569)

ProPublica's article on Machine Bias began a conversation on *fairness* and *algorithmic accountability* that has only gotten more convoluted and technical with time. Northpointe's defense of COMPAS showed that fairness is subjective and leaves room for interpretation based on the metric we choose to optimize for.

Studies in the past few years show that it is impossible for any system to achieve *total fairness* across all outcome groups, making trade-offs inevitable. This research, however, remains limited to academia - journalists continue to report on issues of fairness and accountability without a complete understanding of the latest developments on the issue.

My final project aims to bridge this gap between the academic and journalism communities, connecting concepts like confusion matrices, calibration and false positive rates to real-world applications of fairness and bias. I hope that journalists will find themselves better equipped to comment on issues of fairness after appreciating the social science behind fairness and the mathematics that does not allow us to optimize for them simultaneously.

Through scrolly-telling and an interactive data visualization, I will make it easier for anyone to understand that:

1. given two definitions of fairness over one dataset, it is not always possible to achieve both types of fairness simultaneously (calibration vs. false positive rates)
2. given one definition of fairness over two datasets, an algorithm that is fair over one dataset may not be fair over the other (gender fairness vs. race fairness)

Sources:

1. Fairness in Machine Learning, Solon Barocas and Moritz Hardt, NIPS 2017
2. Fairness in Criminal Justice Risk Assessments: The State of the Art (arXiv:1703.09207 **[stat.ML]**)
3. Machine Bias, Julia Angwin, Jeff Larson, Surya Mattu and Lauren Kirchner, ProPublica