





















## 7. REFERENCES

- [1] Stop-and-frisk in New York City. [https://en.wikipedia.org/wiki/Stop-and-frisk\\_in\\_New\\_York\\_City](https://en.wikipedia.org/wiki/Stop-and-frisk_in_New_York_City).
- [2] <https://www.documentcloud.org/documents/2702103-Sample-Risk-Assessment-COMPAS-CORE.html>, 2016.
- [3] J. Angwin and J. Larson. Bias in Criminal Risk Scores Is Mathematically Inevitable, Researchers Say. <https://www.propublica.org/article/bias-in-criminal-risk-scores-is-mathematically-inevitable-researchers-say>.
- [4] J. Angwin, J. Larson, S. Mattu, and L. Kirchner. Machine Bias: There’s Software Used Across the Country to Predict Future Criminals. And it’s Biased Against Blacks. <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>, 2016.
- [5] S. Barocas and A. D. Selbst. Big Data’s Disparate Impact. *California Law Review*, 2016.
- [6] C. M. Bishop. *Pattern Recognition and Machine Learning*. Springer, 2006.
- [7] A. Chouldechova. Fair Prediction with Disparate Impact: A Study of Bias in Recidivism Prediction Instruments. *arXiv preprint, arXiv:1610.07524*, 2016.
- [8] K. Crawford. Artificial Intelligence’s White Guy Problem. <https://www.nytimes.com/2016/06/26/opinion/sunday/artificial-intelligences-white-guy-problem.html>.
- [9] C. Dwork, M. Hardt, T. Pitassi, and O. Reingold. Fairness Through Awareness. In *ITCSC*, 2012.
- [10] M. Feldman, S. A. Friedler, J. Moeller, C. Scheidegger, and S. Venkatasubramanian. Certifying and Removing Disparate Impact. In *KDD*, 2015.
- [11] A. W. Flores, C. T. Lowenkamp, and K. Bechtel. False Positives, False Negatives, and False Analyses: A Rejoinder to “Machine Bias: There’s Software Used Across the Country to Predict Future Criminals. And it’s Biased Against Blacks.”. 2016.
- [12] S. Goel, J. M. Rao, and R. Shroff. Precinct or Prejudice? Understanding Racial Disparities in New York City’s Stop-and-Frisk Policy. *Annals of Applied Statistics*, 2015.
- [13] G. Goh, A. Cotter, M. Gupta, and M. Friedlander. Satisfying Real-world Goals with Dataset Constraints. In *NIPS*, 2016.
- [14] J. M. Greg Ridgeway. Doubly Robust Internal Benchmarking and False Discovery Rates for Detecting Racial Bias in Police Stops. *Journal of the American Statistical Association*, 2009.
- [15] M. Hardt, E. Price, and N. Srebro. Equality of Opportunity in Supervised Learning. In *NIPS*, 2016.
- [16] F. Kamiran and T. Calders. Classification with No Discrimination by Preferential Sampling. In *BENELEARN*, 2010.
- [17] T. Kamishima, S. Akaho, H. Asoh, and J. Sakuma. Fairness-aware Classifier with Prejudice Remover Regularizer. In *PADM*, 2011.
- [18] J. Kleinberg, S. Mullainathan, and M. Raghavan. Inherent Trade-Offs in the Fair Determination of Risk Scores. In *ITCS*, 2017.
- [19] J. Larson, S. Mattu, L. Kirchner, and J. Angwin. <https://github.com/propublica/compas-analysis>, 2016.
- [20] J. Larson, S. Mattu, L. Kirchner, and J. Angwin. How We Analyzed the COMPAS Recidivism Algorithm. <https://www.propublica.org/article/how-we-analyzed-the-compas-recidivism-algorithm>, 2016.
- [21] B. T. Luong, S. Ruggieri, and F. Turini. kNN as an Implementation of Situation Testing for Discrimination Discovery and Prevention. In *KDD*, 2011.
- [22] C. Muñoz, M. Smith, and D. Patil. Big Data: A Report on Algorithmic Systems, Opportunity, and Civil Rights. *Executive Office of the President. The White House.*, 2016.
- [23] D. Pedreschi, S. Ruggieri, and F. Turini. Discrimination-aware Data Mining. In *KDD*, 2008.
- [24] J. Podesta, P. Pritzker, E. Moniz, J. Holdren, and J. Zients. Big Data: Seizing Opportunities, Preserving Values. *Executive Office of the President. The White House.*, 2014.
- [25] A. Romei and S. Ruggieri. A Multidisciplinary Survey on Discrimination Analysis. *KER*, 2014.
- [26] X. Shen, S. Diamond, Y. Gu, and S. Boyd. Disciplined Convex-Concave Programming. *arXiv:1604.02639*, 2016.
- [27] L. Sweeney. Discrimination in Online Ad Delivery. *ACM Queue*, 2013.
- [28] M. B. Zafar, I. V. Martinez, M. G. Rodriguez, and K. P. Gummadi. Fairness Constraints: Mechanisms for Fair Classification. In *AISTATS*, 2017.
- [29] R. Zemel, Y. Wu, K. Swersky, T. Pitassi, and C. Dwork. Learning Fair Representations. In *ICML*, 2013.