

REFERENCES

- [1] Emrah Akyol, Cedric Langbort, and Tamer Basar. 2016. Price of transparency in strategic machine learning. *arXiv preprint arXiv:1610.08210* (2016).
- [2] Abhijit V Banerjee. 1992. A simple model of herd behavior. *The Quarterly Journal of Economics* 107, 3 (1992), 797–817.
- [3] Marco Barreno, Blaine Nelson, Russell Sears, Anthony D Joseph, and J Doug Tygar. 2006. Can machine learning be secure?. In *Proceedings of the 2006 ACM Symposium on Information, Computer and Communications Security*. 16–25.
- [4] Sushil Bikhchandani, David Hirshleifer, and Ivo Welch. 1992. A theory of fads, fashion, custom, and cultural change as informational cascades. *Journal of Political Economy* 100, 5 (1992), 992–1026.
- [5] Amit K Chopra and Munindar P Singh. 2018. Sociotechnical systems and ethics in the large. In *Proceedings of the 2018 AAAI/ACM Conference on AI, Ethics, and Society*. 48–53.
- [6] Danielle Keats Citron and Frank Pasquale. 2014. The scored society: Due process for automated predictions. *Washington Law Review* 89 (2014), 1.
- [7] Nilesh Dalvi, Pedro Domingos, Sumit Sanghai, and Deepak Verma. 2004. Adversarial classification. In *Proceedings of the 10th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. 99–108.
- [8] Boi Faltings and Goran Radanovic. 2017. Game theory for data science: Eliciting truthful information. *Synthesis Lectures on Artificial Intelligence and Machine Learning* 11, 2 (2017), 1–151.
- [9] Ganesh Ghalme, Vineet Nair, Itay Eilat, Inbal Talgam-Cohen, and Nir Rosenfeld. 2021. Strategic Classification in the Dark. In *Proceedings of the 38th International Conference on Machine Learning (Proceedings of Machine Learning Research, Vol. 139)*, Marina Meila and Tong Zhang (Eds.). PMLR, 3672–3681. <https://proceedings.mlr.press/v139/ghalme21a.html>
- [10] Bryce Goodman and Seth Flaxman. 2017. European Union regulations on algorithmic decision-making and a “right to explanation”. *AI Magazine* 38, 3 (2017), 50–57.
- [11] Moritz Hardt, Nimrod Megiddo, Christos Papadimitriou, and Mary Wootters. 2016. Strategic classification. In *Proceedings of the 2016 ACM Conference on Innovations in Theoretical Computer Science*. 111–122.
- [12] Jon Kleinberg and Manish Raghavan. 2020. How Do Classifiers Induce Agents to Invest Effort Strategically?. *ACM Transactions on Economics and Computation (TEAC)* 8, 4 (2020), 1–23.
- [13] Reshef Meir, Ariel D Procaccia, and Jeffrey S Rosenschein. 2012. Algorithms for strategyproof classification. *Artificial Intelligence* 186 (2012), 123–156.
- [14] Pradeep K Murukannaiah, Nirav Ajmeri, Catholijn M Jonker, and Munindar P Singh. 2020. New foundations of ethical multiagent systems. In *Proceedings of the 19th International Conference on Autonomous Agents and MultiAgent Systems (AAMAS’20)*. 1706–1710.
- [15] Cathy O’Neil. 2016. *Weapons of math destruction: How big data increases inequality and threatens democracy*. Crown.
- [16] Javier Perote and Juan Perote-Pena. 2004. Strategy-proof estimators for simple regression. *Mathematical Social Sciences* 47, 2 (2004), 153–176.
- [17] Suresh Venkatasubramanian and Mark Alfano. 2020. The philosophical basis of algorithmic recourse. In *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*. 284–293.
- [18] Sandra Wachter, Brent Mittelstadt, and Chris Russell. 2017. Counterfactual explanations without opening the black box: Automated decisions and the GDPR. *Harvard Journal of Law & Technology* 31 (2017), 841.