

- [108] Adam J Mersereau, Paat Rusmevichientong, and John N Tsitsiklis. A structured multiarmed bandit problem and the greedy policy. *IEEE Transactions on Automatic Control*, 54(12):2787–2802, 2009.
- [109] Jacob Metcalfe, Emanuel Moss, Elizabeth Anne Watkins, Ranjit Singh, and Madeleine Clare Elish. Algorithmic impact assessments and accountability: The co-construction of impacts. In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, pages 735–746, 2021.
- [110] Alison D Morantz. Has devolution injured american workers? state and federal enforcement of construction safety. *The Journal of Law, Economics, & Organization*, 25(1):183–210, 2009.
- [111] Emanuel Moss, Elizabeth Anne Watkins, Jacob Metcalfe, and Madeleine Clare Elish. Governing with algorithmic impact assessments: six observations. Available at SSRN 3584818, 2020.
- [112] Deirdre K Mulligan and Kenneth A Bamberger. Procurement as policy: Administrative process for machine learning. *Berkeley Tech. LJ*, 34:773, 2019.
- [113] Mohammad Naghsvar, Tara Javidi, and Kamalika Chaudhuri. Noisy bayesian active learning. In *2012 50th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, pages 1626–1633. IEEE, 2012.
- [114] Yusuke Narita, Shota Yasui, and Kohei Yata. Debaised off-policy evaluation for recommendation systems. In *Fifteenth ACM Conference on Recommender Systems*, pages 372–379, 2021.
- [115] Xinkun Nie, Xiaoying Tian, Jonathan Taylor, and James Zou. Why adaptively collected data have negative bias and how to correct for it. In *International Conference on Artificial Intelligence and Statistics*, pages 1261–1269. PMLR, 2018.
- [116] Jacob Nogas, Tong Li, Fernando J Yanez, Arghavan Modiri, Nina Deliu, Ben Prystawski, Sofia S Villar, Anna Rafferty, and Joseph J Williams. Algorithms for adaptive experiments that trade-off statistical analysis with reward: Combining uniform random assignment and reward maximization. *arXiv preprint arXiv:2112.08507*, 2021.
- [117] Ziad Obermeyer, Brian Powers, Christine Vogeli, and Sendhil Mullainathan. Dissecting racial bias in an algorithm used to manage the health of populations. *Science*, 366(6464):447–453, 2019.
- [118] Office of Management and Budget. Requirements for payment integrity improvement. <https://www.whitehouse.gov/wp-content/uploads/2018/06/M-18-20.pdf>, 2018. Executive Office of the President. Online; Accessed Jan 10, 2022.
- [119] Pedro A Ortega and Daniel A Braun. Generalized thompson sampling for sequential decision-making and causal inference. *Complex Adaptive Systems Modeling*, 2(1):1–23, 2014.
- [120] Aldo Pacchiano, Shaun Singh, Edward Chou, Alex Berg, and Jakob Foerster. Neural pseudo-label optimism for the bank loan problem. *Advances in Neural Information Processing Systems*, 34, 2021.
- [121] Nicolas Papernot, Patrick McDaniel, Arunesh Sinha, and Michael Wellman. Towards the science of security and privacy in machine learning. *arXiv preprint arXiv:1611.03814*, 2016.
- [122] Vianney Perchet, Philippe Rigollet, Sylvain Chassang, and Erik Snowberg. Batched bandit problems. *The Annals of Statistics*, 44(2):660–681, 2016.
- [123] Reid Pryzant, Kelly Shen, Dan Jurafsky, and Stefan Wagner. Deconfounded lexicon induction for interpretable social science. In *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long Papers)*, pages 1615–1625, 2018.
- [124] Martin L. Puterman. *Markov Decision Processes—Discrete Stochastic Dynamic Programming*. John Wiley & Sons, Inc., New York, NY, 1994.
- [125] Edward Raff and Jared Sylvester. Gradient reversal against discrimination. *arXiv preprint arXiv:1807.00392*, 2018.
- [126] Inioluwa Deborah Raji, Peggy Xu, Colleen Honigsberg, and Daniel Ho. Outsider oversight: Designing a third party audit ecosystem for ai governance. In *Proceedings of the 2022 AAAI/ACM Conference on AI, Ethics, and Society, AIES '22*, page 557–571, New York, NY, USA, 2022. Association for Computing Machinery. ISBN 9781450392471. doi: 10.1145/3514094.3534181. URL <https://doi.org/10.1145/3514094.3534181>.
- [127] Alberto Reyes, Matthijs TJ Spaan, and L Enrique Sucar. An intelligent assistant for power plants based on factored MDPs. In *2009 15th International Conference on Intelligent System Applications to Power Systems*, pages 1–6. IEEE, 2009.
- [128] Julien Roy, Roger Girgis, Joshua Romoff, Pierre-Luc Bacon, and Christopher Pal. Direct behavior specification via constrained reinforcement learning. *arXiv preprint arXiv:2112.12228*, 2021.
- [129] David S Rubenstein. Acquiring ethical ai. *Fla. L. Rev.*, 73:747, 2021.
- [130] Cynthia Rudin. Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nature Machine Intelligence*, 1(5):206–215, 2019.
- [131] Jeffrey Rupert. Doi wildland fire program policy memorandum no. 2020-004. 2020.
- [132] Yoan Russac, Claire Vernade, and Olivier Cappé. Weighted linear bandits for non-stationary environments. In *NeurIPS 2019-33rd Conference on Neural Information Processing Systems*, 2019.
- [133] Daniel Russo, Benjamin Van Roy, Abbas Kazerouni, Ian Osband, and Zheng Wen. A tutorial on thompson sampling. *arXiv preprint arXiv:1707.02038*, 2017.
- [134] Natasha Sarin and Lawrence H Summers. Shrinking the tax gap: approaches and revenue potential. Technical report, National Bureau of Economic Research, 2019.
- [135] Jonathan Scholz, Martin Leivhn, Charles Isbell, and David Wingate. A physics-based model prior for object-oriented MDPs. In *International Conference on Machine Learning*, pages 1089–1097. PMLR, 2014.
- [136] Burr Settles. Active learning literature survey, 2009.
- [137] Mike Shann and Sven Seuken. Adaptive home heating under weather and price uncertainty using GPs and MDPs. In *Proceedings of the 2014 international conference on Autonomous agents and multi-agent systems*, pages 821–828, 2014.
- [138] Jaehyeok Shin, Aaditya Ramdas, and Alessandro Rinaldo. Are sample means in multi-armed bandits positively or negatively biased? *Advances in Neural Information Processing Systems*, 32:7102–7111, 2019.
- [139] Jaehyeok Shin, Aaditya Ramdas, and Alessandro Rinaldo. On the bias, risk, and consistency of sample means in multi-armed bandits. *SIAM Journal on Mathematics of Data Science*, 3(4):1278–1300, 2021.
- [140] Pranav Shyam, Wojciech Jaśkowski, and Faustino Gomez. Model-based active exploration. In *International conference on machine learning*, pages 5779–5788. PMLR, 2019.
- [141] Ayan Sinha, David F Gleich, and Karthik Ramani. Deconvolving feedback loops in recommender systems. *Advances in neural information processing systems*, 29: 3243–3251, 2016.
- [142] Dennis Soemers, Tim Brys, Kurt Driessens, Mark Winands, and Ann Nowé. Adapting to concept drift in credit card transaction data streams using contextual bandits and decision trees. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 32, 2018.
- [143] Liang Tang, Yexi Jiang, Lei Li, and Tao Li. Ensemble contextual bandits for personalized recommendation. In *Proceedings of the 8th ACM Conference on Recommender Systems*, pages 73–80, 2014.
- [144] Jean Tarbouriech and Alessandro Lazaric. Active exploration in markov decision processes. In *The 22nd International Conference on Artificial Intelligence and Statistics*, pages 974–982. PMLR, 2019.
- [145] Taxpayer Advocate Service. Improper earned income tax credit payments: Measures the IRS takes to reduce improper earned income tax credit payments are not sufficiently proactive and may unnecessarily burden taxpayers. https://www.taxpayeradvocate.irs.gov/wp-content/uploads/2020/07/ARC18_Volume1_MSP_06_ImproperEarnedIncome.pdf, 2018. 2018 Annual Report to Congress – Volume One. Online; Accessed Jan 10, 2022.
- [146] Sebastian B Thrun, Knut Möller, and A Linden. Active exploration in dynamic environments. In *NIPS*, pages 531–538, 1991.
- [147] Catherine Tucker. Privacy, algorithms, and artificial intelligence. In *The economics of artificial intelligence: An agenda*, pages 423–437. University of Chicago Press, 2018.
- [148] US Executive Office of the President. Promoting Active Management of America’s Forests, Rangelands, and Other Federal Lands To Improve Conditions and Reduce Wildfire Risk (Executive Order 13855). 2018.
- [149] Raffaele Vardavas, Pavan Katkar, Andrew M Parker, GR Aliyev, Marlon Graf, and KB Kumar. Rand’s interdisciplinary behavioral and social science agent-based model of income tax evasion. 2019.
- [150] Claire Vernade, Alexandra Carpentier, Tor Lattimore, Giovanni Zappella, Beyza Ermis, and Michael Brueckner. Linear bandits with stochastic delayed feedback. In *International Conference on Machine Learning*, pages 9712–9721. PMLR, 2020.
- [151] Alex Warofka. An independent assessment of the human rights impact of facebook in myanmar. *Facebook Newsroom*, November, 5, 2018.
- [152] Songbai Yan, Kamalika Chaudhuri, and Tara Javidi. Active learning from imperfect labelers. *Advances in Neural Information Processing Systems*, 29:2128–2136, 2016.
- [153] Jiayu Yao, Emma Brunskill, Weiwei Pan, Susan Murphy, and Finale Doshi-Velez. Power constrained bandits. In *Machine Learning for Healthcare Conference*, pages 209–259. PMLR, 2021.
- [154] Taraneh Younesian, Zilong Zhao, Amirmasoud Ghiassi, Robert Birke, and Lydia Y Chen. Qactor: Active learning on noisy labels. In *Asian Conference on Machine Learning*, pages 548–563. PMLR, 2021.
- [155] Ahmed M Youssef, Biswajeet Pradhan, and Abdallah Mohamed Hassan. Flash flood risk estimation along the st. katherine road, southern sinai, egypt using gis based morphometry and satellite imagery. *Environmental Earth Sciences*, 62(3):611–623, 2011.
- [156] Andrea Zanette, Kefan Dong, Jonathan Lee, and Emma Brunskill. Design of experiments for stochastic contextual linear bandits. *Advances in Neural Information Processing Systems*, 34, 2021.
- [157] Ruohan Zhan, Zhimei Ren, Susan Athey, and Zhengyuan Zhou. Policy learning with adaptively collected data. *arXiv preprint arXiv:2105.02344*, 2021.
- [158] Peng Zhao, Lijun Zhang, Yuan Jiang, and Zhi-Hua Zhou. A simple approach for non-stationary linear bandits. In *International Conference on Artificial Intelligence and Statistics*, pages 746–755. PMLR, 2020.
- [159] Stephan Zheng, Alexander Trott, Sunil Srinivasa, Nikhil Naik, Melvin Gruesbeck, David C Parkes, and Richard Socher. The ai economist: Improving equality and productivity with ai-driven tax policies. *arXiv preprint arXiv:2004.13332*, 2020.