Mitigating Emergent Biases in Online Learning

Peter Bartlett, Alex Berg, Aldo Pacchiano, Jakob Foerster
peter@berkeley.edu, acberg@fb.com, pacchiano@berkeley.edu, jnf@fb.com,

Description – Motivation, Related Work: Why here, why now? Online learning algorithms are the basis of a myriad of data driven systems used to perform extremely consequential decisions in finance, internet commerce, and even policing (5, 7, 11). It has recently come to the attention of policy makers and machine learning researchers (9, 13, 17) that careless use of these automated decision making systems may result in perpetuation and amplification of existing societal biases. In many of these tasks, the feedback structure of an online learner can be modeled as one-sided, since the true labels are only observed for the examples that get "accepted". For example, in the policing criminal recidivism example, the learner only observes an inmate's recidivism if an inmate has been released. In this class of problems, which we refer to as the *bank loan problem* (BLP), the learner only observes if a customer will have a chance to repay a loan if the loan is issued to begin with. In these cases, inability to design effective ways to update a model can lead to severe training imbalances that may lead to discriminatory behavior towards some individuals. As more and more consequential decisions about individuals' access to finance, health and education are automated it is important to design algorithms that can mitigate the emergence of potential imbalances that can lead to the perpetuation of the same discriminatory biases that automated systems are trying to avoid. In this project we propose efficient algorithms that properly align the incentives of reward maximization for the decision maker with that of non-discrimination.

Novelty and Innovation - What is novel, innovative, risky about your approach? Data used to train machine learning systems often contain human and societal biases that can lead to treat individuals unfavorably (*unfairly*) on the basis of characteristics such as race, gender, disabilities, etc. This has motivated researchers to investigate techniques to ensure models satisfy fairness properties (2, 3, 4, 10, 14, 15). One way to mitigate biases and prevent discrimination is via introducing appropriate constraints. The extension of this work to the online setting has been less studied, although some works do treat the problem of imposing online population or individual fairness constraints on the predictions of a model (1, 6, 8, 12). We make use of the BLP problem as a laboratory for studying the effects of online decision making in propagating and preventing fairness imbalances in the case a decision maker is also trying to maximize a reward signal. In the case the online decision maker is interested in maximizing the accuracy of its decisions, the BLP problem can be cast as a contextual bandit problem with two actions. We use it as a test bed for developing contextual bandit algorithms with provable regret guarantees and practical implementation in the seldom studied and challenging setting of function approximation with Deep Neural Networks.

Technical Contributions – What are your technical goals, in specific and measurable terms? There exist some recent algorithms that aim to study the problem of contextual bandits in the neural network setting such as NeuralUCB (16). In this work we aim to develop a simple computationally efficient alternative to NeuralUCB. We introduce the PLOT Algorithm (Pesudo Label Optimism). PLOT is designed to work in the setting of online classification. The algorithm is extremely simple. Every time the learner is to make a decision regarding a datapoint(s), it retrains a model where the datapoint(s) she is deciding on are artificially included in the training set with a positive label. We show empirically that PLOT achieves competitive regret guarantees on a variety of public datasets when compared with other existing neural bandit algorithms such as $\epsilon-$greedy and NeuralUCB. We have tested the performance of PLOT in a variety of simulated and public datasets of the UCI database. We see the regret of PLOT is competitive and oftentimes smaller from other approaches to the BLP such as $\epsilon-$greedy or NeuralUCB.

Software. The PLOT algorithm's library can be found at (`https://github.com/pacchiano/OnlineBias`). We want to open source our implementation hoping it serves as a test-bed for neural bandit algorithms in the future.

References

[1] Y. Bechavod, C. Jung, and Z. S. Wu. Metric-free individual fairness in online learning. *arXiv preprint arXiv:2002.05474*, 2020.

[2] S. Chiappa and W. S. Isaac. A causal bayesian networks viewpoint on fairness. In *IFIP International Summer School on Privacy and Identity Management*, pages 3–20. Springer, 2018.

[3] S. Corbett-Davies, E. Pierson, A. Feller, S. Goel, and A. Huq. Algorithmic decision making and the cost of fairness. In *Proceedings of the 23rd acm sigkdd international conference on knowledge discovery and data mining*, pages 797–806, 2017.

[4] C. Dwork, M. Hardt, T. Pitassi, O. Reingold, and R. Zemel. Fairness through awareness. In *Proceedings of the 3rd innovations in theoretical computer science conference*, pages 214–226, 2012.

[5] V. Eubanks. *Automating inequality: How high-tech tools profile, police, and punish the poor*. St. Martin's Press, 2018.

[6] S. Gillen, C. Jung, M. Kearns, and A. Roth. Online learning with an unknown fairness metric. *arXiv preprint arXiv:1802.06936*, 2018.

[7] M. Hoffman, L. B. Kahn, and D. Li. Discretion in hiring. *The Quarterly Journal of Economics*, 133(2): 765–800, 2018.

[8] M. Joseph, M. Kearns, J. Morgenstern, and A. Roth. Fairness in learning: Classic and contextual bandits. *arXiv preprint arXiv:1605.07139*, 2016.

[9] M. Mann and T. Matzner. Challenging algorithmic profiling: The limits of data protection and anti-discrimination in responding to emergent discrimination. *Big Data & Society*, 6(2):2053951719895805, 2019.

[10] H. Narasimhan, A. Cotter, M. Gupta, and S. Wang. Pairwise fairness for ranking and regression. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 34, pages 5248–5255, 2020.

[11] C. Perlich, B. Dalessandro, T. Raeder, O. Stitelman, and F. Provost. Machine learning for targeted display advertising: Transfer learning in action. *Machine learning*, 95(1):103–127, 2014.

[12] G. N. Rothblum and G. Yona. Probably approximately metric-fair learning. *arXiv preprint arXiv:1803.03242*, 5(2), 2018.

[13] J. Silberg and J. Manyika. Notes from the ai frontier: Tackling bias in ai (and in humans). *McKinsey Global Institute (June 2019)*, 2019.

[14] S. Verma and J. Rubin. Fairness definitions explained. In *2018 ieee/acm international workshop on software fairness (fairware)*, pages 1–7. IEEE, 2018.

[15] Y. Wu, L. Zhang, X. Wu, and H. Tong. Pc-fairness: A unified framework for measuring causality-based fairness. *arXiv preprint arXiv:1910.12586*, 2019.

[16] D. Zhou, L. Li, and Q. Gu. Neural contextual bandits with ucb-based exploration. In *International Conference on Machine Learning*, pages 11492–11502. PMLR, 2020.

[17] F. Zuiderveen Borgesius et al. Discrimination, artificial intelligence, and algorithmic decision-making. 2018.