RESEARCH STATEMENT

Ashudeep Singh
mail@ashudeepsingh.com
Dec 31, 2020

In most applications of machine learning, humans work together with algorithms to form joint decision-making systems. Search and recommendation systems are great examples of how algorithms influence human choices and how they have a long-term influence on the entire system, including the users and the recommended items. One of the main thrusts of my research has been to develop a rigorous computational framework for search and recommendation on online platforms that rethinks how these systems foster long-term economic growth while ensuring utility, fairness, and safety for the users as well as the creators and producers of the items. In my current research, I study group and individual fairness notions for ranking in such platforms by considering exposure as the economic opportunity to be equitably allocated. These notions are then used to develop machine learning algorithms that optimize user utility while providing fairness guarantees. With the growing use of algorithm-assisted decision-making in even more consequential and sensitive domains, e.g. criminal justice, hiring, admissions, etc., this study of addressing and mitigating bias is more pertinent than ever, and I wish to continue my research on facilitating the use of fair machine learning in its applications.

My research spans the broad areas of machine learning, recommendation systems, and information retrieval. A common thread in my research work has been to build machine learning algorithms to learn from interactive user feedback in user-facing platforms while dealing with partial information, offline data, and selection bias. In my research, I have resorted to causal inference techniques to answer counterfactual questions arising in evaluation and learning, as well as used Reinforcement Learning and Control theory approaches to solve the emergent constrained optimization problems. My research process often follows these three distinct exercises: (1) Going from normative considerations to mathematical formulations and definitions, (2) devising and analyzing practical Machine Learning algorithms, and (3) mapping out the trade-offs between the desired metrics. Alongside this process, I derive the valuable joy of doing research by communicating ideas, facilitating their impact on real-world applications, and through collaboration and mentorship opportunities. One of our recent papers from a collaboration in which I mentored two undergraduate researchers at Cornell University led to a paper called "Controlling Fairness and Bias in Dynamic Learning-to-Rank" [1] that was awarded the *Best Paper Award* at ACM SIGIR 2020. In the remainder of this statement, I elaborate on specific research directions that I have focused on during my Ph.D., and later outline some future directions and my research vision.

Fairness of Exposure in Rankings [1, 3, 4, 5]

Rankings are ubiquitous in today's multi-stakeholder online economies (e.g., online marketplaces, job search, property renting, media streaming). In these systems, the items to be ranked are products, job candidates, artistic content, or other entities that transfer economic benefit. It is widely recognized that the position of an item in the ranking has a crucial influence on its exposure and economic success, greatly influencing which products get purchased, which candidates get interviewed, and which movies get streamed. Surprisingly, though, the algorithms used to learn these rankings are typically oblivious to the effect they have on the items being ranked.

In work published in KDD 2018, we develop a conceptual and computational framework to define exposure-based group fairness criteria. For example, for a particular application domain, one might consider a proportionality constraint that avoids *Disparate Exposure* by ranking items such that each group (e.g. a seller, the group of female candidates, etc.) is allocated an exposure that is proportional to the average relevance of the items in that group. The key problem such a framework tackles is that the conventional ranking procedure leads to only the top few positions in the ranking getting most of the user attention and hence leading to downstream effects such as the rich-get-richer effect, polarization, misrepresentation of real-world distributions, etc. Our framework allows specification of fairness constraints while optimizing for conventional user utility functions (e.g. DCG, Precision@k, Average Rank) over ranking policies. One key idea that opens up the space of solutions is to allow the class of ranking policies to represent the set of stochastic ranking policies (i.e. the set of distributions over permutations of items being ranked). While the utility maximization part prefers policies that are highly personalized to the users, the fairness constraint prefers stochastic ranking policies that satisfy notions of group fairness amortized over the distribution of users. Alongside this convenience in optimization, it introduces several interesting theories and mechanisms design questions. This work also provides a pathway to derive efficient machine learning algorithms to rank with other exposure allocation schemes.

In subsequent work published at NeurIPS 2019 [3], we propose the first method for merit-based ranking fairness into the Learning-to-Rank framework to propose a novel policy gradient algorithm called Fair-PG-Rank to learn ranking policies that optimize user utility while satisfying fairness constraints. Through our algorithm, we show that efficiently learning highly effective ranking policies subject to fairness guarantees is feasible. Moreover, we also show that satisfying the ranking-based fairness constraints makes the model learn to ignore biased features and that the fairness properties generalize well to unseen queries and candidate sets.

While Fair-PG-Rank [3] is an efficient Learning-to-Rank (LTR) algorithm for offline learning from batched data, LTR algorithms are often deployed in an online setting, where the ranking function adapts based on the feedback that users provide. Such dynamic LTR problems are ubiquitous in online systems, for example, news-feed rankings that adapt to the number of "likes" an article receives, online stores that adapt to the number of positive reviews for a product, or movie-recommendation systems that adapt to who has watched a movie. In all of these systems, learning and prediction are dynamically intertwined, where past feedback influences future rankings in a specific form of online learning with partial information feedback. While dynamic LTR systems are in widespread use and unquestionably useful, there are at least two issues that require careful design considerations– first, the ranking system induces a *selection bias* through the rankings it presents since the items ranked highly are more likely to collect additional feedback, and second, the algorithm might allocate exposure disparately to two groups of candidate items. These issues in turn influence future rankings and may promote misleading rich-get-richer dynamics, for example, unfairly allocated economic gain from product sales or streaming revenue, the polarization of a news platform towards a particular ideology, etc. In the work published at SIGIR 2020 [1], we propose a learning-to-rank algorithm that solves both these issues while simultaneously learning the ranking function from implicit partial feedback data from users. The algorithm takes the form of a controller that integrates unbiased estimators for fairness and utility, dynamically adapting both as more data becomes available. In addition to its rigorous theoretical foundation and convergence guarantees, we empirically show that the algorithm is highly practical and robust.

As a recognition for the importance of this line of research and its impact on the areas of Information Retrieval (IR) and Recommender Systems (RecSys), our paper "Controlling Fairness and Bias in Dynamic Learning-to-Rank" [1] was awarded the Best Paper Award at ACM SIGIR 2020. Advances in understanding the impact and formulating these notions of fairness have also led to keen interest from the industry in rethinking production ranking and recommendation algorithms. Some notable examples include the LinkedIn Talent search [8], a large-scale production recommendation system at Google [9], and Spotify Music recommendations [10].

Impact of Recommender Systems on Users [1, 2]

User engagement and well-being are important considerations when building platforms to foster long-term healthy engagement. Through the lens of exposure fairness in ranking (as discussed above), I believe that user engagement in the long term depends on the platform's ability to retain and promote diverse content, opinions, and creators. Moreover, we must realize that these algorithms may also have a profound impact on the users' long-term preferences as individuals and as a society. With algorithms built to solely optimize for short-term user engagement objectives, this impact is often overlooked and, in the worst cases, can even cause harm.

One of the problems we study in our recent work [2] is that solely optimizing for a user engagement metric in a sequential recommendation setting, may lead to trajectories that may be considered as *unhealthy* user behavior. For example, consider a user dynamics model represented by a Markov decision process (MDP) such that on exposure to unhealthy content (e.g. violent movies, junk food, misinformation, etc.), the user transitions to a state where the user returns an even higher reward for the unhealthy content. The goal of the recommender agent, in this case, is to avoid such user trajectories. Moreover, this risk may be unevenly distributed across the set of users, where only a minority subset of the users experience such unhealthy trajectories. In a way, the goal for a recommendation agent becomes to satisfy the Rawlsian principle of maximin welfare [11] which states that a fair system should be designed to maximize the position of those who will be worst-off in it. To develop a recommendation approach that maximizes the "healthiness" of worst-case user experiences, we propose a distributional notion of risk through a metric known as Conditional Value-at-Risk (CVaR). In this work [2], we adopt a Safe Reinforcement Learning (Safe RL) approach to propose a policy gradient algorithm to train a recommendation agent that adapts to the user's behavior in a sequential recommendation setting and show its efficacy and properties as compared to multi-objective RL.

Machine-assisted human decision making — A Research Vision

As machine learning-based systems growingly become an inseparable part of our lives, they have a direct and an indirect impact on our future selves, our society as well as the economy. In the future, I wish to continue to pursue this inquiry into the impacts of such systems. This section lays out some of the ongoing and relevant research directions for the future.

Diversity and Fairness in Rankings

Fairness of exposure in rankings is seemingly close to the idea of diversity in information retrieval. While both goals lead to rankings that do not follow the probability ranking principle [12], the two are fundamentally different. In particular, algorithms that optimize diversity simply use a different model of user utility, while still optimizing exclusively for user utility without any consideration of the ranked items. One difference between conventional and diversified utility measures is that they are not necessarily modular (i.e., linearly additive) in the set of ranked items, but that they can model dependencies between the items, most commonly in the form of a submodular set function [13]. So while diversity and fairness of exposure have different goals, they do appear to have mutually compatible effects on the rankings they produce. An interesting future research problem to study diversified ranking algorithms with provable fairness guarantees. One way to achieve this is by defining submodular utility measures for evaluation and optimization (e.g. [14]), however, optimizing such metrics is substantially more challenging than the modular utility metrics (e.g. DCG) considered so far, but existing connections between submodular set functions and linear programming [15] may provide a path to tackling the problem with the techniques developed in my work [4].

Position Bias, Trust Bias, Uncertainty, and Exploration

Past studies have provided evidence that position in the ranking not only affects exposure (Position Bias), but also the user's valuation of the item (i.e., the user's perception of relevance) [16]. This is commonly referred to as *Trust Bias*, where the user takes the rank of the item as evidence of its relevance. Trust bias is often intertwined with the position-based attention bias when we observe clicks. However, Trust Bias is also fundamentally different since its causal pathway is through the user's valuation of the item and not through the missingness of the data as is the case with Position Bias. This makes it an interesting causal inference challenge to solve either through minor interventions or by harnessing natural experiments in observational data.

Although correcting for position and trust bias may result in estimates that are correct on average, the variance of these estimates may still be substantially different between groups. This variance problem is well known for IPS estimators used for counterfactual evaluation and learning [17]. Furthermore, the estimates are likely to have a higher variance for historically underrepresented items because of the lack of data. This means that a single point-estimate can be far off, leading to potential unfairness. To overcome this problem, it is important to study how active exploration can be introduced into the estimation problem of [1], leading to the largely unexplored questions of fairness in online learning (e.g. [18, 19]). In this way, active exploration adds another component not only to the trade-off between short-term versus long-term user utility but also to the fairness of the system.

Studying these effects also requires a scientific inquiry into user behavior models. Since the work on exposure-based item fairness and the impact on users uses certain assumptions about the user behavior, and since all these systems operate under complex computer-human interaction conditions and interfaces, it becomes very important to examine and validate these behavioral models. This opens the door to making definitive claims about the effect of these systems on its users. In an industry setting, there is huge scope to understand this better via minor interventions [20], or harnessing natural experiments in observational data [21].

Fair and Safe Exploration in Sequential Recommendations

Given the considerable success of Reinforcement Learning (RL) in games, robotics, and physical system control, it has also become a common framework for training recommender systems that optimize user feedback throughout the sequence [22]. However, the use of RL for recommendations brings new challenges of its own, for example, using off-policy logged data to optimize the recommendation policy, exploration in the online setting, the enormous size of the action space, etc. From the perspective of fairness and safety, exploration, in particular, raises some complex, unanswered ethical questions. The hope is that, even though exploration temporarily degrades the user experience, it leads to improvements in the long run. However, there is a high chance that some groups of users share much of the burden of exploration without sufficient payoff in the longer term. This is a fascinating topic of study both from an ethical and algorithmic perspective, and it perfectly aligns with the research goal of building fair and responsible recommender systems.

Applying research to real-world problems

In machine learning, a research cycle includes applying methods to systems in the real world to validate the proposed hypotheses and measure the efficacy of algorithms when the underlying theoretical assumptions may not apply. Apart from making a tangible difference in the world, through my research work, I wish to study and understand how fair automated tools can be used and where their pitfalls are. In the past, I have worked on interesting practical applications of my research where the outcomes would either define how to build those systems right, or inform policymaking about what these systems should or should not be allowed to do. First, as part of an ongoing research project on recommendations under uncertain relevance estimates, we implemented a recommender system for paper and networking recommendations at the KDD 2020 conference. In addition to ameliorating some of the challenges that come with the virtual conference format like facilitating engagement with papers and other conference attendees, we set out to measure and understand the impact of the choice of recommendation policy on the visibility and reach of the conference papers by running a randomized controlled trial. Through the evidence collected in the experiment, we are able to show that the recommender system does not lose significant user engagement when it tries to make sure that the recommended items (i.e. conference papers in this case) are exposed more equitably by the system to the participants as compared to a *greedy* recommendation policy. In another research collaboration, we are working with the CS Ph.D. admissions committee at Cornell to understand how to improve the application review process by avoiding and mitigating biases. Current human-based decisions have their own flaws and blind spots, since screening hundreds or thousands of applications is error-prone and quite likely biased. Moreover, the high variance between reviewers makes it even more difficult to ensure that all reviewers are using consistent criteria. Machine-learned models have different strengths and weaknesses, and hence may have a role in augmenting the decision-making process. In this engagement, we use machine learning models to identify reviewer blind spots in the peer review process with an aim to answer a key research question, i.e., whether automated tools can augment the overall process to eventually improve decisions while also mitigating human errors and biases.

Overall, it is evident that when humans and algorithms come together to form complex decision-making systems, the challenge of mitigating the social, economic, and legal concerns becomes even harder. I believe that there is a need for a continuously evolving multidisciplinary effort towards regulating these systems and, in my future research, I look forward to actively contributing to this effort.

Select Publications

[1] Marco Morik*, Ashudeep Singh*, Jessica Hong, Thorsten Joachims. **"Controlling Fairness and Bias in Dynamic Learning-to-Rank"**. *In Proceedings of 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval 2020.* (* equal contribution) [**Best Paper Award**]

[2] Ashudeep Singh, Yoni Halpern, Nithum Thain, Konstantina Christakopoulou, Ed H. Chi, Jilin Chen, and Alex Beutel. **"Building Healthy Recommendation Sequences for Everyone: A Safe Reinforcement Learning Approach"**. *In FAccTRec Workshop at ACM RecSyS, 2020. Full version under submission.*

[3] Ashudeep Singh and Thorsten Joachims. **"Policy Learning for Fairness in Ranking"**. *In Proceedings of Advances in Neural Information Processing Systems (NeurIPS) 2019, Vancouver, BC, Canada.*

[4] Ashudeep Singh and Thorsten Joachims. **"Fairness of Exposure in Rankings"**. *In KDD '18: The 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (SIGKDD), August 19–23, 2018, London, United Kingdom.*

[5] Ashudeep Singh, Thorsten Joachims. **"Equality of Opportunity in Rankings"**. *At Workshop on Prioritising Online Content at NeurIPS 2017.*

[6] Ashudeep Singh, Thorsten Joachims. **"Learning Item Embeddings using Biased Feedback"**. *At Causal Inference and Machine Learning for Intelligent Decision Making Workshop at NeurIPS 2017.*

[7] Tobias Schnabel, Adith Swaminathan, Ashudeep Singh, Navin Chandak, Thorsten Joachims. **"Recommendations as Treatments: Debiasing Learning and Evaluation"** *In Proceedings of The International Conference on Machine Learning (ICML), 2016.*

Other References

[8] Sahin Cem Geyik, Stuart Ambler, and Krishnaram Kenthapadi. Fairness-aware ranking in search & recommendation systems with application to linkedin talent search. *arXiv preprint arXiv:1905.01989*, 2019.

[9] Alex Beutel, Jilin Chen, Tulsee Doshi, Hai Qian, Li Wei, Yi Wu, Lukasz Heldt, Zhe Zhao, Lichan Hong, Ed H Chi, et al. Fairness in recommendation ranking through pairwise comparisons. *arXiv preprint arXiv:1903.00780*, 2019.

[10] Rishabh Mehrotra, James McInerney, Hugues Bouchard, Mounia Lalmas, and Fernando Diaz. Towards a fair marketplace: Counterfactual evaluation of the trade-off between relevance, fairness & satisfaction in recommendation systems. In *Proceedings of the 27th ACM International Conference on Information and Knowledge Management*, pages 2243–2251. ACM, 2018.

[11] John Rawls. *Justice as fairness: A restatement.* Harvard University Press, 2001.

[12] Stephen E Robertson. The probability ranking principle in ir. *Journal of documentation*, 33(4):294–304, 1977.

[13] Yisong Yue and Carlos Guestrin. Linear submodular bandits and their application to diversified retrieval. In *Advances in Neural Information Processing Systems*, pages 2483–2491, 2011.

[14] Fernando Diaz, Bhaskar Mitra, Michael D Ekstrand, Asia J Biega, and Ben Carterette. Evaluating stochastic rankings with expected exposure. *arXiv preprint arXiv:2004.13157*, 2020.

[15] G. Calinescu, C. Chekuri, M. Pal, and J. Vondrak. Maximizing a submodular set function subject to a matroid constraint. In *Conference on Integer Programming and Combinatorial Optimization (IPCO)*, pages 182–196, 2007.

[16] T. Joachims, L. Granka, Bing Pan, H. Hembrooke, F. Radlinski, and G. Gay. Evaluating the accuracy of implicit feedback from clicks and query reformulations in web search. *ACM Transactions on Information Systems (TOIS)*, 25(2), April 2007.

[17] A. Swaminathan and T. Joachims. Batch learning from logged bandit feedback through counterfactual risk minimization. *Journal of Machine Learning Research (JMLR)*, 16:1731–1755, Sep 2015. Special Issue in Memory of Alexey Chervonenkis.

[18] Vishakha Patil, Ganesh Ghalme, Vineet Nair, and Y Narahari. Achieving fairness in the stochastic multi-armed bandit problem. *arXiv preprint arXiv:1907.10516*, 2019.

[19] Matthew Joseph, Michael Kearns, Jamie H Morgenstern, and Aaron Roth. Fairness in learning: Classic and contextual bandits. In *Advances in Neural Information Processing Systems*, pages 325–333, 2016.

[20] T. Joachims, A. Swaminathan, and T. Schnabel. Unbiased learning-to-rank with biased feedback. In *ACM Conference on Web Search and Data Mining (WSDM)*, pages 781–789, 2017.

[21] Amit Sharma, Jake M. Hofman, and Duncan J. Watts. Estimating the causal impact of recommendation systems from observational data. In *Proceedings of the Sixteenth ACM Conference on Economics and Computation*, EC '15, page 453–470, New York, NY, USA, 2015. Association for Computing Machinery.

[22] Minmin Chen, Alex Beutel, Paul Covington, Sagar Jain, Francois Belletti, and Ed H Chi. Top-k off-policy correction for a reinforce recommender system. In *Proceedings of the Twelfth ACM International Conference on Web Search and Data Mining*, pages 456–464, 2019.