Pratik Gajane | pratik.gajane@gmail.com | +31-633-313-415 | pratikgajane.github.io

Research Statement

Machine learning algorithms are increasingly being used to make crucial decisions that affect human lives. In many of these scenarios, the feedback available for such sequential decision-making problems does not conform to the standard feedback assumed in the literature. Motivated by these aspects, the overarching vision of my research is as follows:

i) **Devise efficient algorithms for sequential decision-making with practically motivated forms of feedback.**
ii) **Devise responsible machine learning algorithms with optimal mechanisms to guarantee a desired level of privacy and/or fairness in realistic settings.**

Constructing adaptive algorithms for non-stationary problems is a key interest.
Other themes such as explaining (fair) algorithmic decisions, constrained learning, human-AI interaction, data-centric learning, and autonomous learning also feature in my research agenda. Below, I explain some of the salient themes in my research.

Sequential Decision-Making with Practically Motivated Forms of Feedback

In tasks requiring human feedback, it is often more practical to elicit preference-based feedback than absolute feedback. Motivated by such scenarios, we studied reinforcement learning from preference feedback in [27] and devised a best-in-class algorithm along with a corresponding general lower bound on the performance. Experiments on real-world information retrieval datasets show that our proposed algorithm outperforms state-of-the-art benchmark solutions.

In many real-world applications of sequential decision-making such as continuous control and robotics, the available feedback is sparse. In [5, 10], we proposed efficient algorithms to learn from sparse feedback. In [7], we proposed a method to improve sample efficiency in sparse-feedback multi-agent reinforcement learning. Experimental results on benchmark problems show that our proposed method outperforms the existing methods in sample efficiency.

Batched feedback is another form of non-standard feedback often used in real-world scenarios where it is infeasible to process feedback online due to the computational complexity or cost. In [15, 16], we considered reinforcement learning with batched feedback and provided a policy-agnostic analysis. We also provided experimental results showcasing the practical applicability of our work on real data acquired from our industrial partner KPN.

In an MSc thesis under my co-supervision in collaboration with ASML, we introduced a novel deep reinforcement learning approach for supply chain management. Our proposed solution surpasses state-of-the-art methods in reducing costs while minimally affecting service performance metrics, underscoring its significant practical impact.

Ongoing and Future Work: Continuing my prior work, I plan to explore the application of preference-based feedback in clinical decision-making processes, an area where its effectiveness has been demonstrated [29]. In this context, I am currently working on devising an effective learning algorithm for scenarios where people provide feedback in the form of a preference over multiple items. Here, I am addressing a problem where these preferences may change significantly and unpredictably over time, reflecting real-world scenarios [5].

Furthermore, in many practical scenarios (e.g., patient feedback for healthcare recommendations) feedback can be delayed and may also arrive via partial rewards that are observed with different delays. To deal with such scenarios, I will work on devising an efficient algorithm for reinforcement learning problems with temporally partitioned delayed rewards (see our initial work in [8]). The versatility of this proposed work extends beyond healthcare, making it applicable to domains like microgrid management, supply chain management and online education delivery.

For our work on autonomous exploration [10], I plan to explore connections with biologically inspired curiosity-driven learning. Here, the proposed algorithm first goes via an extended developmental period in which it learns reusable skills autonomously and then uses those skills to achieve general goals. Furthermore, I plan to expand our work in [10] from single-goal navigation to multi-goal navigation. The practical applications of this work extend to the domain of automation and robotics, such as agriculture robots and rescue robots in urban search and rescue missions.

Fairness in Machine Learning

In [22], we provided theoretical as well as empirical critiques of the fairness formalizations used in the machine learning literature. We explained how these critiques limit the suitability of the fairness formalizations to certain domains. For example, we showed that counterfactual fairness formalizations may not be suitable for healthcare and law enforcement due to the presence of hindsight bias and outcome bias. This was one of the inaugural works demonstrating an interdisciplinary approach to fair machine learning research and it continues to be used as reading material for related courses e.g., at Stanford University, at the University of Waterloo, and at the Fairness and Bias workshop at MIT Open Courses. We recently wrote an extensive survey on fairness in reinforcement learning [3], studying single-agent as well as multi-agent formulations, long-term fairness, and offline learning. Moreover, I recently led a team of presenters to deliver a tutorial on fair reinforcement learning at the 33rd International Joint Conference on Artificial Intelligence (IJCAI) 2024.

Ongoing and Future Work: In an ongoing collaboration with the University of Michigan Center for Managing

Chronic Disease, we focus on identifying and rectifying gender bias in personalized recommendations for chronic pain care. See our current work in [1], where we proposed a generalizable ML solution that can be applied in conjunction with clinicians' inputs to optimize population-level outcomes while minimizing disparities in outcomes across subgroups. Next, we are planning to incorporate counterfactual fairness notions in this problem. I plan to continue this work further with insights from healthcare experts and I am eager to explore collaboration opportunities on this front. Additionally, discussions are underway with healthcare experts at the University of Twente to explore collaborative efforts examining the impact of personalization in e-healthcare on fairness objectives.

In another ongoing collaboration with a social scientist from the University of Wisconsin, Madison, we are formalizing an egalitarian fairness notion called *equality of resources* [30]. This principle, previously studied in the social sciences literature, holds potential applications in resource allocation for crucial domains like healthcare [30] and education [31].

Privacy-Preserving Personalized Recommendations

In [24, 26], we addressed the challenge of providing local differential privacy while learning to provide personalized recommendations from human feedback. We devised a frequentist algorithm and a Bayesian algorithm that achieved best-in-class performance, offering users an optimal mechanism to balance privacy and utility. These articles advanced the state of the art as most of the previous related work in the related literature focused on global differential privacy, which is a milder privacy notion. Our experimental results show that our proposed methods are useful in recommender systems. Recently in [13], I extended this problem setting to consider dynamic environments and proposed an efficient frequentist algorithm.

Ongoing and Future Work: Firstly, the solution I proposed in [13] requires that (an upper bound on) the number of changes is known in advance. I aim to devise an optimal algorithm that does not need to know the number of changes in advance. I believe an approach similar to our earlier work in [19] can be useful here. Another prospective direction is to devise an efficient and computationally feasible Bayesian approach for this problem. This Bayesian approach will be especially beneficial in real-world applications to handle parameter uncertainty and modeling errors, and in scenarios where informative priors are available [32].

Furthermore, since our approach in [13, 24] provides a strong privacy notion, it is especially vital for fair decision-making, considering the involvement of sensitive attributes. Thus, I will use the methods developed in our work [13, 24] to devise fair reinforcement learning algorithms capable of preserving local differential privacy. Moreover, a distinguishing feature of our work is an optimal mechanism for users to control the utility-privacy trade-off. This enhances the applicability of our work in domains such as healthcare, where users' perspectives on the utility-privacy trade-off vary with context [33].

Explainable Machine Learning and Human-AI Interaction

In [9], we introduced an efficient recommendation system that utilizes both measured and self-reported user data to generate personalized healthcare recommendations. To enhance user trust, we employed Shapely Additive exPlanations [34] to illustrate the contribution of each attribute to the recommendation. Additionally, in [11], we addressed a gap in the literature by investigating sampling techniques in local surrogate models for machine learning explanations. Our solution demonstrated superior performance for high-dimensional data, offering more faithful and robust explanations compared to existing state-of-the-art methods. This was validated through rigorous evaluations on synthetic and real-life datasets.

Ongoing and Future Work: A prospective direction is to provide adaptive personalized recommendations along with their explanations, using the formulation of reinforcement learning. Here the goal is to personalize both the explanations and recommendations while adapting to changing user responses to recommendations.

I also plan to work on providing narrative explanations for fair recommendations – particularly for pain care. The significance of narrative-based explanations in enhancing the efficacy and trustworthiness of pain care recommendations has been explored in [35].

Data-Centric Learning for Energy and Healthcare

Data-centric learning has recently emerged as a crucial approach in data-rich domains such as healthcare and energy. By enhancing the robustness and accuracy of predictive models through the effective handling of evolving data patterns, this approach improves patient outcomes with precise diagnostics and personalized treatments, and optimizes energy distribution and reliability to support sustainable and efficient energy systems. In our recent work [1], we demonstrated that while uncritical utilization of patient data in ML solutions can yield highly effective personalized healthcare recommendations, it also has the potential to result in disparities in outcomes among subgroups defined by sensitive attributes like gender and race. As a remedy, we suggested a an automated feature selection algorithm that can adaptively learn to select the data features that optimize for a required combination of utility and fairness.

I also worked on a European research project (titled 'Dynamically Evolving Long-Term Autonomy' ↗) with the goal of learning behaviour policy in electrical distribution networks that exhibit non-stationary behaviours. In [18, 19, 20], we devised algorithms that achieved optimal performance without prior knowledge of the number of changes (in contrast to the previous related work). Our proposed algorithms were the first optimal solutions that are not tuned with respect to the number of changes in the environment.

Ongoing and Future Work: I plan to use the versatile tools we introduced in [18, 19, 20] in various other decision-making scenarios in non-stationary environments such as healthcare recommendations (effects of treatments on patients vary with time), robotics (changes in surroundings). In many such real-world applications, assuming prior knowledge of non-stationarity (such as the number of changes) is often unrealistic. The tools that I propose to use here from my prior work [19,20] stand out in their ability to handle dynamic environments without the need for prior knowledge of non-stationarity. This sets them apart from other solutions in the literature and makes them particularly well-suited for real-world applications.

I am also working incorporating "learning-to-defer" mechanism when ML solutions learned on one dataset are to be deployed in a new scenario. In this setup, when the algorithm is uncertain about the outcomes of its decisions in certain instances in the new scenario, it learns to adaptively defer to a decisions to a human expert or baseline policy. Such a setup will be especially useful for complex domains like healthcare and automation, where leveraging insights from human experts is critical for ensuring safe decision-making in scenarios where the consequences of an incorrect decision are significant.

Concluding Remarks

In summary, my objective is to devise machine learning algorithms that can learn from practical forms of feedback to tackle real-world challenges. Considering the societal impact of my work, I aim to devise solutions with efficient mechanisms for fairness, privacy, and explainability. I am keen on developing adaptive learning algorithms to effectively address the ubiquitous challenge of non-stationarity in real-world scenarios. I have explored a diverse array of application areas for my work, including healthcare, personalised AI, supply chain management, energy management and control systems. Looking ahead, I am eager to broaden the scope of my work even further, venturing into new domains.

References

References [1] to [28] can be found in my CV.

[29] Yanan Sui and Joel W. Burdick. Correlational dueling bandits with application to clinical treatment in large decision spaces. In *Proceedings of the 26th International Joint Conference on Artificial Intelligence*, IJCAI'17, page 2793–2799, 2017.

[30] Ronald Dworkin. *Sovereign Virtue: The Theory and Practice of Equality*. Harvard University Press, 2000. ISBN 9780674002197. URL `http://www.jstor.org/stable/j.ctv1c3pd0r`.

[31] Ben Colburn and Hugh Lazenby. Hypothetical insurance and higher education. *Journal of Philosophy of Education*, 50(4):587–604, 2016.

[32] Mohammad Ghavamzadeh, Shie Mannor, Joelle Pineau, and Aviv Tamar. Bayesian reinforcement learning: A survey. *Found. Trends Mach. Learn.*, 8(5–6):359–483, 2015.

[33] André Calero Valdez and Martina Ziefle. The users' perspective on the privacy-utility trade-offs in health recommender systems. *International Journal of Human-Computer Studies*, 121:108–121, 2019. Advances in Computer-Human Interaction for Recommender Systems.

[34] Scott M Lundberg and Su-In Lee. A unified approach to interpreting model predictions. *Advances in neural information processing systems*, 30, 2017.

[35] Ana Sofia Carvalho, Sandra Martins Pereira, António Jácomo, Susana Magalhães, Joana Araújo, Pablo Hernández-Marrero, Carlos Costa Gomes, and Michael Schatman. Ethical decision making in pain management: a conceptual framework. *Journal of Pain Research*, Volume 11:967–976, May 2018.