KOALA: Knowledge of Optimization and Learning Algorithms for Healthcare

Kai Wang

Georgia Institute of Technology
Computational Science and Engineering
kwang692@gatech.edu

Abstract

The Knowledge of Optimization And Learning Algorithms (KOALA) group studies how to integrate optimization, machine learning, and generative modeling to enable data-driven decision-making under uncertainty. We study decision-focused learning, embedding optimization as a differentiable layer to train models end-to-end for decision quality. We design scalable reinforcement learning algorithms for population and personalized healthcare, and develop efficient bilevel optimization methods for nested and multi-agent decision-making. These directions form a unified framework linking optimization and learning for impactful AI in healthcare. Through collaborations with hospitals and NGOs, our group designs and deploys algorithms for pediatric, diabetes, maternal, and mental health applications. Looking ahead, we aim to unite these foundations with generative AI to build theoretically grounded and socially responsible algorithms that advance trustworthy, real-world AI for health and beyond.

1 Decision-Focused Learning

We study decision-focused learning: integrating optimization as a differentiable layer within machine learning pipelines so models can be trained end-to-end on task-level decision quality. Our work develops methods for backpropagating through optimization problems—both convex and non-convex (Wang et al. 2020a,b; Shah et al. 2022; Diaz et al. 2024; Wang et al. 2022; Johnson-Yu et al. 2023; Zhao et al. 2025)—and for sequential decision-making (Wang et al. 2021, 2023a). These techniques establish a principled framework in which predictions are shaped by their downstream impact, rather than by surrogate losses alone. Beyond theory, our decision-focused learning algorithms have also achieved measurable impact in healthcare at scale. In partnership with the Indian NGO ARMMAN, we co-designed and field-tested a decision-focused algorithm to optimize a maternal and child health program that delivers weekly health information to pregnant women. The algorithm was deployed program-wide, improving retention by 31% for more than 350,000 participants since 2022 (Verma et al. 2022). This work demonstrates how aligning model training with downstream decisions can translate into better delivery of healthcare resources.

2 Sequential Decision Making for Healthcare

We develop efficient, tractable algorithms for bandits and reinforcement learning that support sequential decision-making in healthcare under uncertainty and at population scale. This work includes multi-armed bandit formulations designed to operate under computational constraints for large cohorts (Xu et al. 2021; Wang et al. 2023b; Hung, Hsieh, and Wang 2025; Zhang and Wang 2025; Pandey, Wang, and Garcia 2025), as well as diffusion-policy methods in reinforcement learning that leverage generative models to accelerate and stabilize learning (Ma et al. 2025; Chen et al. 2025; Hu et al. 2025). These methods form the core of our group's ongoing collaboration with Massachusetts General Hospital, where we design adaptive interventions for diabetes and chronic disease using bandits, reinforcement learning, and generative models subject to resource constraints. Our group is also partnering with Children's Healthcare of Atlanta to deploy a bedside monitoring system with intervention recommendations for pediatric care.

3 Foundations of Bilevel Structure

We study bilevel optimization—optimization problems nested within other optimization problems—as a foundation for modern machine learning and decision-making. Our recent work develops efficient fully first-order algorithms, together with complexity analysis, for bilevel problems with linear constraints (Kornowski et al. 2024) and for Stackelberg models with multiple followers (Niu, Wang, and Ziani 2025), generalizing prior analyses of fully first-order methods. By designing scalable algorithms and establishing provable guarantees, our research connects classical optimization principles with emerging challenges in machine learning, offering a unifying framework for data-driven decision-making with bilevel structures.

4 Future Directions

These directions define a research program that unites optimization, learning, and generative AI to produce algorithms that are theoretically principled, practically deployable, and societally impactful. As AI capabilities accelerate, our aim is to keep them grounded in data scarcity, cultural context, and resource constraints—so that advances translate into measurable improvements in health and well-being for all.

Acknowledgements

This work was supported by the National Science Foundation under Grant No. IIS-2403240, the National Heart, Lung, and Blood Institute of the National Institutes of Health under Award Number R01HL184139, and Schmidt Sciences AI2050 Fellowship.

References

Chen, T.; Ma, H.; Li, N.; Wang, K.; and Dai, B. 2025. One-Step Flow Policy Mirror Descent. *arXiv preprint arXiv:2507.23675*.

Diaz, P. R.; Kong, L.; Wang, K.; Alvarez-Melis, D.; and Tambe, M. 2024. What is the Right Notion of Distance between Predict-then-Optimize Tasks? *CoRR*.

Hu, Y.; Chen, T.; Li, N.; Wang, K.; and Dai, B. 2025. Primal-Dual Spectral Representation for Off-policy Evaluation. *AISTATS*.

Hung, Y.-H.; Hsieh, P.-C.; and Wang, K. 2025. Non-Stationary Restless Multi-Armed Bandits with Provable Guarantee. *arXiv preprint arXiv:2508.10804*.

Johnson-Yu, S.; Finocchiaro, J.; Wang, K.; Vorobeychik, Y.; Sinha, A.; Taneja, A.; and Tambe, M. 2023. Characterizing and improving the robustness of predict-then-optimize frameworks. In *GameSec*. Springer.

Kornowski, G.; Padmanabhan, S.; Wang, K.; Zhang, Z.; and Sra, S. 2024. First-order methods for linearly constrained bilevel optimization. *Advances in Neural Information Processing Systems*, 37: 141417–141460.

Ma, H.; Chen, T.; Wang, K.; Li, N.; and Dai, B. 2025. Efficient Online Reinforcement Learning for Diffusion Policy. *ICML*.

Niu, A.; Wang, K.; and Ziani, J. 2025. Finding a Multiple Follower Stackelberg Equilibrium: A Fully First-Order Method. *arXiv preprint arXiv:2509.08161*.

Pandey, H. S.; Wang, K.; and Garcia, G.-G. P. 2025. Neural Index Policies for Restless Multi-Action Bandits with Heterogeneous Budgets. *arXiv preprint arXiv:2510.22069*.

Shah, S.; Wang, K.; Wilder, B.; Perrault, A.; and Tambe, M. 2022. Decision-Focused Learning without Differentiable Optimization: Learning Locally Optimized Decision Losses. *Advances in Neural Information Processing Systems*.

Verma, S.; Mate, A.; Wang, K.; Taneja, A.; and Tambe, M. 2022. Case Study: Applying Decision Focused Learning in the Real World. *NeurIPS Workshop on Trustworthy and Socially Responsible Machine Learning (TSRML)*.

Wang, K.; Perrault, A.; Mate, A.; and Tambe, M. 2020a. Scalable Game-Focused Learning of Adversary Models: Data-to-Decisions in Network Security Games. In *AAMAS*.

Wang, K.; Shah, S.; Chen, H.; Perrault, A.; Doshi-Velez, F.; and Tambe, M. 2021. Learning MDPs from Features: Predict-Then-Optimize for Sequential Decision Making by Reinforcement Learning. *Advances in Neural Information Processing Systems*.

Wang, K.; Verma, S.; Mate, A.; Shah, S.; Taneja, A.; Madhiwalla, N.; Hegde, A.; and Tambe, M. 2023a. Decision-Focused Learning in Restless Multi-Armed Bandits with Application to Maternal and Child Care Domain. In *AAAI*.

Wang, K.; Wilder, B.; Perrault, A.; and Tambe, M. 2020b. Automatically learning compact quality-aware surrogates for optimization problems. *Advances in Neural Information Processing Systems*.

Wang, K.; Xu, L.; Perrault, A.; Reiter, M. K.; and Tambe, M. 2022. Coordinating followers to reach better equilibria: End-to-end gradient descent for stackelberg games. In *AAAI*.

Wang, K.; Xu, L.; Taneja, A.; and Tambe, M. 2023b. Optimistic Whittle Index Policy: Online Learning for Restless Bandits. In *AAAI*.

Xu, L.; Bondi, E.; Fang, F.; Perrault, A.; Wang, K.; and Tambe, M. 2021. Dual-mandate patrols: Multi-armed bandits for green security. In *AAAI*.

Zhang, H.; and Wang, K. 2025. Networked Restless Multi-Arm Bandits with Reinforcement Learning. *AAAI Planning and Reinforcement Learning Workshop*.

Zhao, Z.; Yeh, C.; Kong, L.; and Wang, K. 2025. Diffusion-DFL: Decision-focused Diffusion Models for Stochastic Optimization. *arXiv preprint arXiv:2510.11590*.