University of Pennsylvania – Wharton School
Operations, Information and Decisions

OIDD 941: Bandits, RL & Transfer Learning

Tentative Syllabus – Spring 2021 Q4

Professor Hamsa Bastani (hamsab@wharton.upenn.edu)

Office Hours: TBD

Course Description

The main objective of OIDD941 is to familiarize students with recent literature in bandits, reinforcement learning and transfer learning, as well as their applications to healthcare, operations and revenue management and social good. We will emphasize open research problems throughout the class.

Meeting Times Fridays 2-5pm

Tentative Schedule & Reading List

Class 1 (March 19) Bandits

- **[UCB]** Abbasi-Yadkori, Yasin, Dávid Pál, and Csaba Szepesvári. "Improved algorithms for linear stochastic bandits." *Advances in Neural Information Processing Systems*. 2011.
- **[TS]** Agrawal, Shipra, and Navin Goyal. "Analysis of thompson sampling for the multi-armed bandit problem." *Conference on learning theory*. 2012.
- **[Nonparametric]** Slivkins, Aleksandrs. "Contextual bandits with similarity information." *Proceedings of the 24th annual Conference On Learning Theory*. JMLR Workshop and Conference Proceedings, 2011.

Class 2 (March 26) Bandits continued

- **[High Dimension]** Bastani, Hamsa, and Mohsen Bayati. "Online decision-making with high-dimensional covariates." *Forthcoming in Operations Research*. 2015.
- **[Greedy]** Bastani, Hamsa, Mohsen Bayati, and Khashayar Khosravi. "Mostly exploration-free algorithms for contextual bandits." *Management Science*. 2020.
- **[Constraints]** Badanidiyuru, Ashwinkumar, Robert Kleinberg, and Aleksandrs Slivkins. "Bandits with knapsacks." *2013 IEEE 54th Annual Symposium on Foundations of Computer Science*. IEEE, 2013.
- **[Safety]** Wu, Yifan, et al. "Conservative bandits." *International Conference on Machine Learning*. 2016.
- **[Guest Talk]** Sinha, Deeksha, et al. "Multi-armed Bandits with Cost Subsidy." *arXiv.* 2020.

Class 3 (April 2) Challenges in practice

- **[Nonstationarity]** Cheung, Wang Chi, David Simchi-Levi, and Ruihao Zhu. "Learning to optimize under non-stationarity." *The 22nd International Conference on Artificial Intelligence and Statistics.* 2019.
- **[Batching]** Perchet, Vianney, et al. "Batched bandit problems." *The Annals of Statistics.* 2016.
- **[Delayed feedback]** Joulani, Pooria, et al. "Online learning under delayed feedback." *International Conference on Machine Learning*. 2013.

- **[Constraints]** Agrawal, Shipra, and Nikhil R. Devanur. "Bandits with concave rewards and convex knapsacks." *Proceedings of the fifteenth ACM conference on Economics and computation*. 2014.

Class 4 (April 9) Reinforcement learning

- **[UCRL]** Jaksch, Thomas, Ronald Ortner, and Peter Auer. "Near-optimal Regret Bounds for Reinforcement Learning." *Journal of Machine Learning Research*. 2010.
- **[Thompson Sampling]** Osband, Ian, and Benjamin Van Roy. "Why is posterior sampling better than optimism for reinforcement learning?" *International Conference on Machine Learning.* 2017.
- **[Conservative RL]** Garcelon, Evrard, et al. "Conservative exploration in reinforcement learning." *International Conference on Artificial Intelligence and Statistics.* 2020.
- **[Safe RL]** Garcıa, Javier, and Fernando Fernández. "A comprehensive survey on safe reinforcement learning." *Journal of Machine Learning Research*. 2015.
- **[Inventory Management]** Gijsbrechts, Joren, et al. "Can deep reinforcement learning improve inventory management? performance on dual sourcing, lost sales and multi-echelon problems." *SSRN*. 2019.

Class 5 (April 16) Transfer learning

- **[Survey]** Pan, Sinno Jialin, and Qiang Yang. "A survey on transfer learning." *IEEE Transactions on knowledge and data engineering*. 2009.
- **[Sparse Transfer]** Bastani, Hamsa. "Predicting with proxies: Transfer learning in high dimension." *Management Science*. 2020.
- **[Multi-Task Learning]** Ben-David, Shai, et al. "A theory of learning from different domains." *Machine learning*. 2010.
- **[Multi-Task Learning]** Farias, Vivek, and Andrew Li. "Optimal Recovery of Tensor Slices." *Artificial Intelligence and Statistics*. 2017.

Class 6 (April 23) Student final project presentations

Grading

50% class participation (attendance, asking questions); 50% final project

Final Project

The topic of the project is entirely of your choice! You should see this as an opportunity to explore in more depth an idea or a problem that you find interesting and/or relevant.