MSc Projects CWI's Machine Learning Group

Best Arm Identification in Survival Models

Supervisor: Wouter M. Koolen
Keywords: Pure Exploration, Bandits, Survival Analysis, Hypothesis Testing

Background: Pure Exploration is an active area of Machine Learning and Statistics. Its central problem of Best Arm Identification has been studied intensively since (Even-Dar, Mannor, and Mansour, 2002), and worst-case optimal methods have been developed for the fixed confidence, fixed budget and simple regret settings (Bubeck, Munos, and Stoltz, 2011). After a long and respectable series of papers establishing worst-case optimality, a revolutionary new approach called Track-and-Stop was pioneered by Garivier and Kaufmann (2016) that delivers instance-optimal methods. Since then, several aspects of Track-and-Stop have been generalised and refined: tighter stopping thresholds were constructed by Kaufmann and Koolen (2021), computational efficiency was improved using saddle-point methods by Degenne, Koolen, and M´enard (2019), and problems with multiple answers were analysed by Degenne and Koolen (2019).

Academic Content: In this project the Best Arm Identification problem will be extended to active testing in survival analysis models. Two complications arise when transitioning from a standard bandit to survival analysis models: data now arrive with delay, and are right-censored. There is extensive recent work on bandits with delayed feedback, both in the regret and pure exploration settings, by Vernade, Capp´e, and Perchet (2017), Pike-Burke et al. (2018), Vernade, Carpentier, Zappella, et al. (2018), Grover et al. (2018), and Vernade, Carpentier, Lattimore, et al. (2020). Bandits with censored feedback are considered by Abernethy, Amin, and Zhu (2016), though the literature seems much less developed. Passive hypothesis testing with survival data (including optional stopping problems) has recently been considered by Grunwald et al. (2020).

The project will focus on developing the model and theory, following these steps:
1. Review literature on Best Arm Identification, in particular the Track-and-Stop methodology.
2. Review literature on nonparametric survivial models, in particular the Cox proportional hazards model.
3. Study the active testing problem (BAI) for survival models. Defining carefully the interaction protocol and metric. Gain experience by addressing the BAI problem for survival models but without delays and censoring.
4. Extend the Track-and-Stop method to active testing in survival models with delays and censoring. The project is envisaged to consist of mostly theoretical work, with only minor computational and empirical components.

References

Abernethy, J. D., K. Amin, and R. Zhu (2016). "Threshold bandits, with and without censored feedback". In: Advances In Neural Information Processing Systems 29, pp. 4889–4897.

Bubeck, S., R. Munos, and G. Stoltz (2011). "Pure Exploration in Finitely Armed and Continuous Armed Bandits". In: Theoretical Computer Science 412, 1832-1852 412, pp. 1832–1852.

Degenne, R. and W. M. Koolen (Dec. 2019). "Pure Exploration with Multiple Correct Answers". In: Advances in Neural Information Processing Systems (NeurIPS) 32, pp. 14591–14600.

Degenne, R., W. M. Koolen, and P. M´enard (Dec. 2019). "Non-Asymptotic Pure Exploration by Solving Games". In: Advances in Neural Information Processing Systems (NeurIPS) 32, pp. 14492–14501.

Even-Dar, E., S. Mannor, and Y. Mansour (2002). "PAC Bounds for Multi-armed Bandit and Markov Decision Processes". In: Computational Learning Theory, 15th Annual Conference on Computational Learning Theory, COLT 2002, Sydney, Australia, July 8-10, 2002, Proceedings. Vol. 2375. Lecture Notes in Computer Science, pp. 255–270.

Garivier, A. and E. Kaufmann (2016). "Optimal Best arm Identification with Fixed Confidence". In: Proceedings of the 29th Conference On Learning Theory (COLT).

Grover, A., T. Markov, P. Attia, N. Jin, N. Perkins, B. Cheong, M. Chen, Z. Yang, S. Harris, W. Chueh, et al. (2018). "Best arm identification in multi-armed bandits with delayed feedback". In: International Conference on Artificial Intelligence and Statistics. PMLR, pp. 833–842.

Grunwald, P., A. Ly, M. Perez-Ortiz, and J. ter Schure (2020). "The Safe Log Rank Test: Error Control under Optional Stopping, Continuation and Prior Misspecification". In: arXiv preprint arXiv:2011.06931.

Kaufmann, E. and W. M. Koolen (Nov. 2021). "Mixture Martingales Revisited with Applications to Sequential Tests and Confidence Intervals". In: Journal of Machine Learning Research 22.246, pp. 1–44.

Pike-Burke, C., S. Agrawal, C. Szepesvari, and S. Grunewalder (2018). "Bandits with delayed, aggregated anonymous feedback". In: International Conference on Machine Learning. PMLR, pp. 4105–4113.

Vernade, C., O. Capp´e, and V. Perchet (2017). "Stochastic bandit models for delayed conversions". In: arXiv preprint arXiv:1706.09186.

Vernade, C., A. Carpentier, T.Lattimore, G. Zappella, B. Ermis, and M. Brueckner (2020). "Linear bandits with   stochastic delayed feedback". In: International Conference on Machine Learning.   PMLR, pp. 9712–9721.

Vernade, C., A. Carpentier, G. Zappella, B. Ermis, and M. Brueckner (2018). "Contextual Bandits under Delayed Feedback." In: arXiv preprint arXiv:1807.02089.