Why the Shooting in the Dark Method Dominates Recommender Systems Practice?

David Rohde

Criteo AI Lab, France d.rohde@criteo.com

Abstract. The introduction of A/B Testing represented a great leap forward in recommender systems research. Like the randomized control trial for evaluating drug efficacy; A/B Testing has equipped recommender systems practitioners with a protocol for measuring performance as defined by actual business metrics and with minimal assumptions. Crucially A/B testing provided a way to *measure* the performance of two or more candidate systems, but it provided no guide of *what policy we should test*. The focus of this industry talk is to better understand, why the development of A/B testing was the last great leap forward in the development of reward optimizing recommender systems despite more than a decade of efforts in both industry and academia. The talk will survey: industry best practice, standard theories and tools including: collaborative filtering (MovieLens RecSys), contextual bandits, attribution, off-policy estimation, causal inference, click through rate models and will explain why we have converged on a fundamentally heuristic solution or *guess and check type method*. The talk will offer opinions about which of these theories are useful, and which are not and make a concrete proposal to make progress based on a non-standard use of deep learning tools.

Keywords: Recommender Systems · A/B Testing · Contextual Bandits · Bayesian Decision Theory

1 Introduction

A/B testing is rightly viewed as a great leap forward by recommender systems practitioners and researchers alike (Kohavi et al., 2009, 2012). Provided some basic assumptions such as the single unit treatment value assumption (SUTVA) are satisfied (Imbens and Rubin, 2010), A/B testing provides a reliable way to measure the performance of a new candidate recommender system. This gives a well founded way to measure the performance of two or more competing recommendation algorithms, but provides *no guidance about what candidate system should be tested*.

The subject of this industry talk, will be to explain why recommender systems practitioners adopt a guess and check formulation i.e. by proposing some sort of proxy method, and then *shooting in the dark* by testing differing proxies at A/B testing time. The talk will draw both upon industry experience and the academic literature and will follow closely the material in Rohde (2024).

2 Structure of the Talk

The talk will open by introducing the curious position that RecSys research finds itself in: there is a rigorous approach to measure performance of a system, but a surprisingly heuristic method for proposing algorithms to test. The talk will then summary 'best practice' both from an industrial and academic formulation.

Given that optimizing A/B testing is the target of a reward optimizing approach, a connection must be made between the very aggregate view of A/B testing, and the very granular logs that make up recommender system logs is required. This connection will be made using the formalism of Bayesian decision theory (De Finetti, 1937; Lad, 1996; Bernardo and Smith, 2009; Kadane, 2020).

2.1 Current Best Practice and Limitations

The contextual bandit (Beygelzimer and Langford, 2009; Bottou et al., 2013) is introduced as an imperfect theoretical framework for building reward optimizing recommender systems, it is noted that attribution is a key heuristic that can be employed to coerce real problems into a contextual bandit framework, but the idea that attribution can be axiomized as proposed in Singal et al. (2019) is rejected.

It is then noted that in even the most toy recommendation setting (one recommendation is shown at a time, a context with tiny cardinality, a tiny catalog of recommendations, and plausible effect sizes) then the estimation problem is highly non-identifiable. That is the likelihood function does not concentrate on a point allowing the identification of good recommendations in all contexts. This causes theoretical approaches based on the IPS or Horvitz-Thompson estimator Bottou et al. (2013); Wasserman (2012, 2022); Sims (2006, 2022)to fail in anything resembling a real system, and requires heuristics such as 'feature engineering' to be resorted to for large scale click models McMahan et al. (2013).

The issue of causality and unobserved confounding in recommender systems will be briefly addressed. It will be argued that confounding will only occur due to poor engineering practice, although poor practices are unfortunately common.

2.2 Deep, but not to go Deep: A Proposal to Unlock Reward Optimizing Recommendation

The talk, will propose that there is hope in building reward optimizing recommender systems, by the use of a non-standard use of deep learning. The proposal is described in more detail in Rohde (2024); Sakhi et al. (2020); Aouali et al. (2023b); Baha et al. (2023).

The fundamental idea is that once an attribution heuristic for clicks has been adopted it is possible to formulate recommendation tasks using an *extended contextual bandit*. Usually real systems, show multiple recommendations simultaneously and more than a simple reward is observed. For example, even if a click is considered a reward on its own, but additional information is contained in *what was clicked*. A strength of deep learning tooling, is not just to 'go deep', but also to build bespoke models suitable for the specific information that is available in a system. These could be the Probabilistic Rank and Reward model if the recommendations form a banner (Aouali et al., 2023b), or the cascade model for search results (Chuklin et al., 2022).

As mentioned before, the fundamental difficulty with the contextual bandit formulation (extended or not), is the extreme lack of identifiability in estimating the reward for each recommendations. Theoretically the correct tool for handling this problem is Bayesian inference, which can both a) handle very uneven information (high information on actions preferred by the old system, low information in other cases); b) combining the reward signal with other information derived from content or collaborative filtering. This can be done by the use of item-item, context-context and item-context similarities Sakhi et al. (2020). Fortunately, deep learning tooling enables the productionizing of Bayesian inference in industrial environments by enabling variational approximations via the re-paramterization trick. Recently this tooling has become mature and available in many production environments Cheng et al. (2016).

The talk will close by summarizing the advantages and possible limitations of the proposed approach.

3 Speaker Bio

David is a Staff Research Lead at the Criteo AI Lab. His research focuses are on causality (Rohde, 2022; Lattimore and Rohde, 2019), recommender systems (Sakhi et al., 2020, 2023; Rohde et al., 2018), contextual bandits (Aouali et al., 2023a) and Bayesian inference (Rohde et al., 2016; Sakhi et al., 2019). From a production perspective, David's main interest are in using Deep Learning for solving recommendation and bidding problems.

David has done numerous public talks, a selection of them are available here:

- Criteo Beer and Tech: Causal Inference with Bayes Rule
- Laplace's Demon: Causal Inference is Inference – A Beautifully Simple Idea that not Everyone Accepts
- A Gentle Introduction to Recommendation as Counterfactual Policy Learning (SIGCHI 2020 Tutorial:)
- Bayesian Value Based Recommendation (RecSys 2020 Tutorial)
- Cornell Invited Talk 2020: Decision Theory for Recommendation
- Bayesian Causal Inference for Real World Interactive Systems (KDD Workshop 2021): Chair of Panel at Bayesian Causal Inference for Real World Interactive Systems

Bibliography

Imad Aouali, Victor-Emmanuel Brunel, David Rohde, and Anna Korba. Exponential smoothing for off-policy learning. *arXiv preprint arXiv:2305.15877*, 2023a.

Imad Aouali, Achraf Ait Sidi Hammou, Sergey Ivanov, Otmane Sakhi, David Rohde, and Flavian Vasile. Probabilistic rank and reward: A scalable model for slate recommendation. 2023b.

Imane Baha, Rohde David, Hamlet Jesse Medina Ruiz, and Vasile Flavian. Large scale bayesian recommendation systems in production environments. In *CONSEQUENCES Workshop RecSys (2023)*, 2023.

José M Bernardo and Adrian FM Smith. *Bayesian theory*, volume 405. John Wiley & Sons, 2009.

Alina Beygelzimer and John Langford. The offset tree for learning with partial labels. In *Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 129–138, 2009.

Léon Bottou, Jonas Peters, Joaquin Quiñonero-Candela, Denis X Charles, D Max Chickering, Elon Portugaly, Dipankar Ray, Patrice Simard, and Ed Snelson. Counterfactual reasoning and learning systems: The example of computational advertising. *Journal of Machine Learning Research*, 14(11), 2013.

Heng-Tze Cheng, Levent Koc, Jeremiah Harmsen, Tal Shaked, Tushar Chandra, Hrishi Aradhye, Glen Anderson, Greg Corrado, Wei Chai, Mustafa Ispir, et al. Wide & deep learning for recommender systems. In *Proceedings of the 1st workshop on deep learning for recommender systems*, pages 7–10, 2016.

Aleksandr Chuklin, Ilya Markov, and Maarten De Rijke. *Click models for web search*. Springer Nature, 2022.

Bruno De Finetti. La prévision: ses lois logiques, ses sources subjectives. In *Annales de l'institut Henri Poincaré*, volume 7, pages 1–68, 1937.

Guido W Imbens and Donald B Rubin. Rubin causal model. In *Microeconometrics*, pages 229–241. Springer, 2010.

Joseph B Kadane. *Principles of uncertainty*. CRC press, 2020.

Ron Kohavi, Roger Longbotham, Dan Sommerfield, and Randal M Henne. Controlled experiments on the web: survey and practical guide. *Data mining and knowledge discovery*, 18:140–181, 2009.

Ron Kohavi, Alex Deng, Brian Frasca, Roger Longbotham, Toby Walker, and Ya Xu. Trustworthy online controlled experiments: Five puzzling outcomes explained. In *Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 786–794, 2012.

Frank Lad. *Operational subjective statistical methods: a mathematical, philosophical and historical introduction*. 1996.

Finnian Lattimore and David Rohde. Replacing the do-calculus with bayes rule. *arXiv preprint arXiv:1906.07125*, 2019.

H Brendan McMahan, Gary Holt, David Sculley, Michael Young, Dietmar Ebner, Julian Grady, Lan Nie, Todd Phillips, Eugene Davydov, Daniel Golovin, et al. Ad click prediction: a view from the trenches. In *Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 1222–1230, 2013.

David Rohde. Causal inference, is just inference: A beautifully simple idea that not everyone accepts. In *I (Still) Can't Believe It's Not Better! Workshop at NeurIPS 2021*, pages 75–79. PMLR, 2022.

David Rohde. Position paper: Why the shooting in the dark method dominates recommender systems practice; a call to abandon anti-utopian thinking. *arXiv preprint arXiv:2402.02152*, 2024.

David Rohde, Stephen Bonner, Travis Dunlop, Flavian Vasile, and Alexandros Karatzoglou. Recogym: A reinforcement learning environment for the problem of product recommendation in online advertising. *arXiv preprint arXiv:1808.00720*, 2018.

David Rohde et al. Semiparametric mean field variational bayes: general principles and numerical issues. *Journal of Machine Learning Research*, 17(172): 1–47, 2016.

Otmane Sakhi, Stephen Bonner, David Rohde, and Flavian Vasile. Reconsidering analytical variational bounds for output layers of deep networks. *arXiv preprint arXiv:1910.00877*, 2019.

Otmane Sakhi, Stephen Bonner, David Rohde, and Flavian Vasile. BLOB: A probabilistic model for recommendation that combines organic and bandit signals. In *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, pages 783–793, 2020.

Otmane Sakhi, David Rohde, and Nicolas Chopin. Fast slate policy optimization: Going beyond plackett-luce. *arXiv preprint arXiv:2308.01566*, 2023.

Christopher Sims. On an example of larry wasserman. *online manuscript, available from http://sims. princeton. edu/yftp/WassermanExmpl/WassermanComment. pdf*, 2(10), 2006.

Christopher Sims. Christopher sims - large parameter spaces and weighted data: A Bayesian perspective. Video, 2022.

Raghav Singal, Omar Besbes, Antoine Desir, Vineet Goyal, and Garud Iyengar. Shapley meets uniform: An axiomatic framework for attribution in online advertising. In *The World Wide Web Conference*, pages 1713–1723, 2019.

Larry Wasserman. A world without referees. *ISBA Bulletin*, 19(1):7–8, 2012.

Larry Wasserman. Problems with Bayesian causal inference. Video, 2022.