Offline Evaluation and Optimization for Interactive Systems

Tutorial Abstract

Lihong Li
Microsoft Research
One Microsoft Way
Redmond, WA 98052
lihongli@microsoft.com

Categories and Subject Descriptors

H.3.3 [**Information Systems**]: Information Storage and Retrieval—*Information Search and Retrieval*; I.2.6 [**Artificial Intelligence**]: Learning

General Terms

Experimentation; Performance

Keywords

Offline evaluation; counterfactual analysis; advertising; information retrieval; Web search; recommender systems; interactive systems; contextual bandits

ABSTRACT

Evaluating and optimizing an **interactive** system (like search engines, recommender and advertising systems) from historical data against a predefined **online** metric is challenging, especially when that metric is computed from user feedback such as clicks and payments. The key challenge is **counterfactual** in nature: we only observe a user's feedback for actions taken by the system, but we do not know what that user would *have* reacted to a different action. The golden standard to evaluate such metrics of a user-interacting system is online A/B experiments (a.k.a. randomized controlled experiments), which can be expensive in terms of both time and engineering resources. *Offline* evaluation/optimization (sometimes referred to as off-policy learning in the literature) thus becomes critical, aiming to evaluate the same metrics *without* running (many) expensive A/B experiments on live users.

One approach to offline evaluation is to build a user model that simulates user behavior (clicks, purchases, etc.) under various contexts, and then evaluate metrics of a system with this simulator. While being straightforward and common in practice, the reliability of such model-based approaches relies heavily on how well the user model is built. Furthermore, it is often difficult to know *a priori* whether a user model is good enough to be trustable.

Recent years have seen a growing interest in another solution to the offline evaluation problem. Using statistical techniques like importance sampling and doubly robust estimation, the approach can give unbiased estimates of metrics for a wide range of problems. It enjoys other benefits as well. For example, it often allows data scientists to obtain a confidence interval for the estimate to quantify the amount of uncertainty; it does not require building user models, so is more robust and easier to apply. All these benefits make the approach particularly attractive to a wide range of problems. Successful applications have been reported in the last few years by some of the industrial leaders.

This tutorial gives a review of the basic theory and representative techniques. Applications of these techniques are illustrated through several case studies done at Microsoft and Yahoo!.

WSDM'15, February 2–6, 2015, Shanghai, China.
ACM ACM 978-1-4503-3317-7/15/02.
http://dx.doi.org/10.1145/2684822.2697040

PRESENTER'S BIO

Lihong Li is a Researcher in the Machine Learning Department at Microsoft Research-Redmond. Prior to joining Microsoft, he was a Research Scientist in the Machine Learning Group at Yahoo! Research in Silicon Valley. He obtained a PhD degree from Rutgers University in Computer Science. His main research interests are machine learning with interaction, including reinforcement learning, multi-armed bandits, online learning, active learning, and their numerous applications on the Web like recommender systems, search, and advertising. He has published over 50 research papers, and is the winner of an ICML Best Student Paper Award, a WSDM Best Paper Award, an AISTATS Notable Paper Award, and a Yahoo! Superstar Team Award. He has served as area chair or senior program committee member at ICML, NIPS, and IJCAI. **URL**: `http://research.microsoft.com/en-us/people/lihongli`.