Unbiased Learning to Rank: Theory and Practice

Qingyao Ai
CICS, UMass Amherst
Amherst, MA, USA
aiqy@cs.umass.edu

Jiaxin Mao
DCST, Tsinghua University
Beijing, China
maojiaxin@gmail.com

Yiqun Liu
DCST, Tsinghua University
Beijing, China
yiqunliu@tsinghua.edu.cn

W. Bruce Croft
CICS, UMass Amherst
Amherst, MA, USA
croft@cs.umass.edu

ABSTRACT

Implicit feedback (e.g., user clicks) is an important source of data for modern search engines. While heavily biased [8, 9, 11, 27], it is cheap to collect and particularly useful for user-centric retrieval applications such as search ranking. To develop an *unbiased learning-to-rank* system with biased feedback, previous studies have focused on constructing probabilistic graphical models (e.g., click models) with user behavior hypothesis to extract and train ranking systems with unbiased relevance signals. Recently, a novel counterfactual learning framework that estimates and adopts examination propensity for unbiased learning to rank has attracted much attention. Despite its popularity, there is no systematic comparison of the unbiased learning-to-rank frameworks based on counterfactual learning and graphical models. In this tutorial, we aim to provide an overview of the fundamental mechanism for unbiased learning to rank. We will describe the theory behind existing frameworks, and give detailed instructions on how to conduct unbiased learning to rank in practice.

KEYWORDS

user bias, click model, counterfactual learning, unbiased learning to rank

ACM Reference Format:

Qingyao Ai, Jiaxin Mao, Yiqun Liu, and W. Bruce Croft. 2018. Unbiased Learning to Rank: Theory and Practice. In *The 27th ACM International Conference on Information and Knowledge Management (CIKM '18), October 22–26, 2018, Torino, Italy*. ACM, New York, NY, USA, 2 pages. https://doi.org/10.1145/3269206.3274274

1 INTRODUCTION

Ranking is a core problem of Information Retrieval (IR). Many IR applications such as Web search, product recommendation, and question answering are ranking problems by nature. Among many ranking paradigms, *learning to rank* is the most widely used technology in both academic research and commercial search engines [13]. The idea of learning to rank is to represent each document with a feature vector and learn a machine learning model that can rank documents based on their relevance to the query.

Although there have been studies that use unsupervised data or pseudo supervision for learning-to-rank models [1, 4], the best retrieval system is typically constructed based on supervised learning. Many of the state-of-the-art retrieval systems today make use of deep models [7, 15], which require large amounts of labeled data. Despite the development of crowdsourcing systems [5, 12], obtaining large-scale and high quality human annotations (e.g. TREC-style relevance judgments) is still expensive, if not impossible. Therefore, implicit feedback such as clicks are still the most attractive data source for the training of ranking systems.

Directly training a ranking model to optimize click data, however, is infeasible because click data are heavily biased [8, 9, 11, 27]. For example, the order of documents in a search engine result page (SERP) has a strong influence on where users click [8]. Studies of position bias show that users tend to examine and click results on the top of a SERP while ignoring those on the bottom. A naive method that treats click/non-click signals as positive/negative feedback will lead to a ranking model that optimizes the order of a search result page but not the real relevance of documents.

To leverage the full power of click data for learning to rank, IR researchers have attempted to remove the effect of user bias in the training of ranking models. One such effort is the development of click models. Click models [3, 6, 21, 22, 25] make hypotheses about how users examine the documents on a SERP, and estimate the true (unbiased) relevance feedback by optimizing the likelihood of the observed user clicks. Ranking models are then trained with the estimated relevance signals so that the overall system is unbiased [14]. Another effort to debias click data is result interleaving [2, 16, 18–20, 26]. By putting different documents at same positions and collecting clicks accordingly, we can obtain unbiased pair preferences for documents in the same result list. These preference signals can then be used to train learning-to-rank models in an online manner.

More recently, a new research direction has emerged that focuses on directly training ranking models with biased click data using counterfactual learning [10, 23, 24]. This unbiased learning-to-rank framework treats click bias as a counterfactual effect and debiases user feedback by weighting each click with their *inverse propensity weights* [17]. It uses a propensity model to quantify click biases and does not explicitly estimate the query-document relevance with training data. As theoretically proven by Joachims et al. [10], given

CIKM '18, October 22–26, 2018, Torino, Italy

ACM ISBN 978-1-4503-6014-2/18/10.
https://doi.org/10.1145/3269206.3274274

the correct bias estimation, ranking models trained with click data under this framework will converge to the same model trained with the true relevance signals.

In this tutorial, we introduce the theory behind each technique and talk about their differences in detail. We provide hands-on instructions on how to customize and apply these unbiased learning-to-rank techniques to different retrieval tasks, and this will provide important guidance for the design of unbiased ranking systems and inspire future studies on related research topics.

2 FORMAT AND SCHEDULE

This tutorial consists of a series of talks on different unbiased learning-to-rank techniques and their applications. We first introduce the motivation for unbiased learning to rank, and then discuss how previous studies tackle the problem from different angles.

The technical content of this tutorial will be organized in two parts. In the first part, we talk about previous studies on click models [3, 6, 21, 22, 25]. The idea of click models is to extract unbiased relevance signals from biased user feedback. They construct hypotheses on user behaviors and build machine learning models (e.g. probabilistic graphic models) to debias user feedback so that we can train a learning-to-rank algorithm using unbiased relevance signals. In this topic, we first introduce the concept of examination hypothesis and representative user behavior models. Then, we describe how to derive a click model based on each examination hypothesis and how to estimate the unbiased relevance signals step by step. We empirically compare different click models in a joint retrieval framework and discuss the advantages and limitations of unbiased learning to rank with click models.

In the second part, we talk about recent developments on counterfactual learning for unbiased learning to rank [10, 23, 24]. In contrast to click models, unbiased learning to rank with counterfactual learning focuses on the estimation of user examination propensity and uses an inverse perpensity weighting schema to create a learning framework in which a ranking model trained with biased user feedback can converge to the same model trained with unbiased relevance signals. In this topic, we first introduce the idea of counterfactual learning and its underlining theory. After that, we will describe how to build an unbiased learning-to-rank framework with inverse propensity weighting and how to estimate examination propensity in online systems. Finally, we discuss the connections and differences between existing unbiased learning-to-rank techniques.

3 SUPPLEMENTAL MATERIALS

The technical manuscript includes content about the click bias that have been identified in the user studies of Web search, the user behavior hypotheses and representative graphic models developed in previous studies, and the proof of theorems and equations for counterfactual learning. All supplemental materials can be found on the homepage of the authors[1].

[1]http://www.cs.umass.edu/~aiqy/

REFERENCES

[1] Qingyao Ai, Liu Yang, Jiafeng Guo, and W Bruce Croft. 2016. Analysis of the paragraph vector model for information retrieval. In *Proceedings of the 2rd ACM ICTIR*. ACM, 133–142.

[2] Olivier Chapelle, Thorsten Joachims, Filip Radlinski, and Yisong Yue. 2012. Large-scale validation and analysis of interleaved search evaluation. *ACM Transactions on Information Systems* 30, 1 (2012), 6.

[3] Olivier Chapelle and Ya Zhang. 2009. A dynamic bayesian network click model for web search ranking. In *Proceedings of the 18th WWW*. ACM, 1–10.

[4] Mostafa Dehghani, Hamed Zamani, Aliaksei Severyn, Jaap Kamps, and W. Bruce Croft. 2017. Neural Ranking Models with Weak Supervision. In *Proceedings of the 40th ACM SIGIR (SIGIR '17)*. ACM, 65–74.

[5] Anhai Doan, Raghu Ramakrishnan, and Alon Y Halevy. 2011. Crowdsourcing systems on the world-wide web. *Commun. ACM* 54, 4 (2011), 86–96.

[6] Georges E Dupret and Benjamin Piwowarski. 2008. A user browsing model to predict search engine click data from past observations.. In *Proceedings of the 31st ACM SIGIR*. ACM, 331–338.

[7] Jiafeng Guo, Yixing Fan, Qingyao Ai, and W Bruce Croft. 2016. A deep relevance matching model for ad-hoc retrieval. In *Proceedings of the 25th ACM CIKM*. ACM, 55–64.

[8] Thorsten Joachims, Laura Granka, Bing Pan, Helene Hembrooke, and Geri Gay. 2005. Accurately interpreting clickthrough data as implicit feedback. In *Proceedings of the 28th annual ACM SIGIR*. Acm, 154–161.

[9] Thorsten Joachims, Laura Granka, Bing Pan, Helene Hembrooke, Filip Radlinski, and Geri Gay. 2007. Evaluating the accuracy of implicit feedback from clicks and query reformulations in web search. *ACM Transactions on Information Systems (TOIS)* 25, 2 (2007), 7.

[10] Thorsten Joachims, Adith Swaminathan, and Tobias Schnabel. 2017. Unbiased learning-to-rank with biased feedback. In *Proceedings of the 10th ACM WSDM*. ACM, 781–789.

[11] Mark T Keane and Maeve O'Brien. 2006. Modeling Result-List Searching in the World Wide Web: The Role of Relevance Topologies and Trust Bias. In *Proceedings of the Cognitive Science Society*, Vol. 28.

[12] Aniket Kittur, Ed H Chi, and Bongwon Suh. 2008. Crowdsourcing user studies with Mechanical Turk. In *Proceedings of the SIGCHI*. ACM, 453–456.

[13] Tie-Yan Liu. 2009. Learning to rank for information retrieval. *Foundations and Trends in Information Retrieval* 3, 3 (2009), 225–331.

[14] Cheng Luo, Yukun Zheng, Jiaxin Mao, Yiqun Liu, Min Zhang, and Shaoping Ma. 2017. Training deep ranking model with weak relevance labels. In *Australasian Database Conference*. Springer, 205–216.

[15] Bhaskar Mitra, Fernando Diaz, and Nick Craswell. 2017. Learning to Match Using Local and Distributed Representations of Text for Web Search. In *Proceedings of the 26th WWW (WWW '17)*. International World Wide Web Conferences Steering Committee, Republic and Canton of Geneva, Switzerland, 1291–1299. https://doi.org/10.1145/3038912.3052579

[16] Karthik Raman and Thorsten Joachims. 2013. Learning socially optimal information systems from egoistic users. In *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*. Springer, 128–144.

[17] Paul R Rosenbaum and Donald B Rubin. 1983. The central role of the propensity score in observational studies for causal effects. *Biometrika* 70, 1 (1983), 41–55.

[18] Anne Schuth, Harrie Oosterhuis, Shimon Whiteson, and Maarten de Rijke. 2016. Multileave gradient descent for fast online learning to rank. In *Proceedings of the 9th ACM WSDM*. ACM, 457–466.

[19] Adith Swaminathan and Thorsten Joachims. 2015. Batch learning from logged bandit feedback through counterfactual risk minimization. *Journal of Machine Learning Research* 16 (2015), 1731–1755.

[20] Adith Swaminathan and Thorsten Joachims. 2015. Counterfactual risk minimization: Learning from logged bandit feedback. In *ICML*. 814–823.

[21] Chao Wang, Yiqun Liu, Meng Wang, Ke Zhou, Jian-yun Nie, and Shaoping Ma. 2015. Incorporating non-sequential behavior into click models. In *Proceedings of the 38th International ACM SIGIR Conference on Research and Development in Information Retrieval*. ACM, 283–292.

[22] Hongning Wang, ChengXiang Zhai, Anlei Dong, and Yi Chang. 2013. Content-aware click modeling. In *Proceedings of the 22nd international conference on World Wide Web*. ACM, 1365–1376.

[23] Xuanhui Wang, Michael Bendersky, Donald Metzler, and Marc Najork. 2016. Learning to rank with selection bias in personal search. In *Proceedings of the 39th ACM SIGIR*. ACM, 115–124.

[24] Xuanhui Wang, Nadav Golbandi, Michael Bendersky, Donald Metzler, and Marc Najork. 2018. Position Bias Estimation for Unbiased Learning to Rank in Personal Search. In *Proceedings of the 11th ACM WSDM (WSDM '18)*. ACM, New York, NY, USA, 610–618. https://doi.org/10.1145/3159652.3159732

[25] Wanhong Xu, Eren Manavoglu, and Erick Cantu-Paz. 2010. Temporal click model for sponsored search. In *Proceedings of the 33rd international ACM SIGIR conference on Research and development in information retrieval*. ACM, 106–113.

[26] Yisong Yue and Thorsten Joachims. 2009. Interactively optimizing information retrieval systems as a dueling bandits problem. In *Proceedings of the 26th ICML*. ACM, 1201–1208.

[27] Yisong Yue, Rajan Patel, and Hein Roehrig. 2010. Beyond position bias: Examining result attractiveness as a source of presentation bias in clickthrough data. In *Proceedings of the 19th WWW*. ACM, 1011–1018.