

Recent Advances in Unbiased Learning to Rank from Position-Biased Click Feedback

Harrie Oosterhuis

June 9, 2021

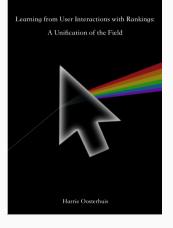
Radboud University harrie.oosterhuis@ru.nl https://twitter.com/HarrieOos

Overview of this Talk



Papers in this talk:

- Policy-Aware Unbiased Learning to Rank for Top-k Rankings
 Harrie Oosterhuis and Maarten de Rijke - SIGIR 2020
- When Inverse Propensity Scoring does not Work: Affine Corrections for Unbiased Learning to Rank Ali Vardasbi, Harrie Oosterhuis and Maarten de Rijke -CIKM 2020
- Unifying Online and Counterfactual Learning to Rank Harrie Oosterhuis and Maarten de Rijke - WSDM 2021



Introduction: Counterfactual Learning to Rank



Goal:

• Optimize a ranking model that matches the user preferences between items, based on historically logged user clicks.



Goal:

• Optimize a **ranking model** that matches the **user preferences** between items, based on **historically logged user clicks**.

Problem:

• Clicks are **biased indicators** of preference (Craswell et al., 2008): factors - other than relevance - also influence click behavior.



Goal:

• Optimize a **ranking model** that matches the **user preferences** between items, based on **historically logged user clicks**.

Problem:

• Clicks are **biased indicators** of preference (Craswell et al., 2008): factors - other than relevance - also influence click behavior.

Existing solution:

• Weight clicks to correct for position bias (Joachims et al., 2017; Wang et al., 2016).

Web Results 1 - 10 of about 676,000	for "digital camera" cheapest (0.36 seconds
An and a second se	
echoarden Alizan	
. Cottigest camble Buy careful a "See Mni Digital Cambra for 519.99. Available: \$10	Cameras
IT First \$150 or 10 or 000 property Commentation 2020 portion 102	In Stock
/wtar Digital Czegora - UP Sontal cameras	Cheaper Prices at Calber
, caineras uk "termitir digital cainera, discolat digital caineras , digital caineras Fronted a chaolisis tease coole e traditi of grant cain og chaopest , succard	Find the best prices and deals.
State of the second	Compare products, shops and review www.Calibex.com
And the second	
Yakumo Digital Camera - Cheapest digital cameras 20/2 2011	Digital Cameras - Save Save of Cameras, Accessories & Mi
	Find the Lowest Price - Smarter.com
25 rege im 00 : 00 utiline over staniera , yakumo uk , cheapest schremory I digitakamenas kerkos core , set es 124001 to tot yakumo html - 31(c- Feb 16, 2005 -	www.propriet.com
Cached - Single Logos	Eres Digital Cameras
	Canon, Kodak, Nikon cameras free! Free w/ offer signup, 18+ only.
Best Deals on Uldiful Cameras and Accessor opt 20 24 Gameras	free w oner signup. 18+ only.
 Section of a full the contract with discourse for the data water is a full call. Section of the data water is a full call of the data water is a full call call of the data water is a full call of the data	
www.bristolcaments.co.uk/- 34k - Feb 15, 2005 - Glichel - Similar, 2053	Digital Cameras Compare Prices on Digital Cameras
Chelog at Dio we Cameras and the Dom Full S3	Read Reviews & Shop at Pricegrabbs
A MICHAELENCI MOLANDANDANDING PROPRIORIDINATIONIC	www.pricegrabber.com
www.lei-eveb.co.uk/cb.	MANDON COMMITTEE ACCOUNTS
M	Factory Refurbished Cameras - Direct From Kodak, Nikon, HP & Mor
rieapest colour LCD digit to amera di bigchina com	www.refurbdepot.com
heanest colour LCD orginal camera Product Description. Charapest digital camera	
ww.ebigchina.com/ebcps/4/ed/1243997.html=94ik - Cached - Similar pages	Unbiased pro and owner reviews plus
Incluint Charmenter and bett heital compress deale	100s of merchant quotes on cameras www.digitalcamera-hg.com
Desiblet - Cheepiest computer and boots pital campras deals Desis Lett Canon Powerber, SO110 digital camera - Canon PMAA (P3000 Photo 10/2017, Deals Lett Mart Society 6405 51 MP Digital Camera w 44 Optical Zoom	
70.2097): Deals List: Nikoli Godpix 5400 5.1 MP Digital Camera w/4x Optical Zoom	Casio Digital Camera Find a Great Deal - Compare Prices
www.dewistst.com/ - 10th - Cliched - Similar pages	& Read Reviews from 100s of Stores
Greapest digital costern InfoJour search	

Image source: http://www.mediative.com/

Existing Policy-Oblivious Estimator



For an item d, a rank k (a.k.a. position), **decompose the click probability** according to rank-based examination model (Craswell et al., 2008):

$$P(C = 1 \mid k, d) = \overbrace{P(E = 1 \mid k)}^{\text{examination}} \overbrace{P(C = 1 \mid E = 1, d)}^{\text{relevance}}.$$



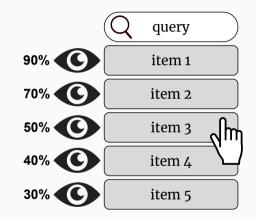
For an item d, a rank k (a.k.a. position), decompose the click probability according to rank-based examination model (Craswell et al., 2008):

$$P(C = 1 \mid k, d) = \underbrace{P(E = 1 \mid k)}_{examination} \underbrace{P(C = 1 \mid E = 1, d)}_{examination} \underbrace{P(C = 1 \mid E = 1, d)}_{examination}.$$

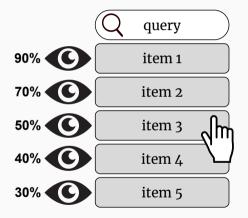
Existing work corrects for position bias by **Inverse Propensity Scoring** (Joachims et al., 2017; Wang et al., 2016). Given N displayed rankings:

$$\mathsf{relevance}(d) \approx \frac{1}{N} \sum_{i=1}^{N} \frac{c_i(d)}{P(E=1 \mid k_i(d))}.$$









Google and Contracting and State and	
Web Results 1 - 10 of about 676,000	for "digital camera" cheapest (0.36 seconds)
and an	
ne se constante a service de la constante de l	Sponsored Links
And the second s	Compare products - hops and reviews www.com/accommons.com Bightal Compare Bon Bien of the main Accessions & More Field the Science Price - Commercian wave again of com Compare and the Compare. Compare again of com Compare again of com- cases again of com- tained a science - Network - Compare Free wide science - Science Steel Free wide science - Science Steel Free wide science -
es DC and Do r Jose are and "A thin cheapent that are all find an instructioner and a div 3 div - Find 5. 200 ppg (2 div - Saint 1 200) Checks of Do ray Cameran and the Tym Find 30 Station for Prior 400 (200 phg / Annual the Company) in an and the cheanter are stated.	Digital Cameras Compare Prices on Digital Cameras Read Reviews & Shop at Pricegrabber www.pricegrabber.com
www.jewees.co.uv/cr.goveesignacameer.com consequences.co.uv/cr.goveesignacameer.com	Factory Returbished Cameras - Direct From Kodak, Nikon, HP & More www.returbidepst.com
Cheanest colore L Co digital castera Product Description. Chempest digital camera	
Construction of the second and a second	Disited Common Options Unbiased pro and owner reviews plus 100s of merchant quotes on cameraal www.digitalcamera-hq.com Casio Digital Camera
www.dexialat.com/-1016-Cachel-Simiarpages	Ease Digital Camera Find a Great Deal - Compare Prices & Read Reviews from 100s of Stores www.ShopCartUSA.com
An analysis depict comera. Starch reputs for "Cheapest digital camera". 1. Buy cheap 10. Cheap 35 Social camera - compressioned	



Estimating position bias can be done via:

• Randomization:

- Swapping the positions of item pairs (Joachims et al., 2017).
- A/B testing (Agarwal et al., 2019b).
- Expectation-Maximization:
 - Bias estimation is easy with an accurate relevance model, and vice versa (Wang et al., 2018a).
- Dual Learning Objective (Ai et al., 2018)

For this talk, we will assume the exact bias is known.

Part I: Top-k Ranking



query item 1 item 2 item 3 item 4 item 5

Top-k ranking:

- setting where only \boldsymbol{K} items can be displayed,
- very prevalent in search and recommendation.

Item-Selection Bias

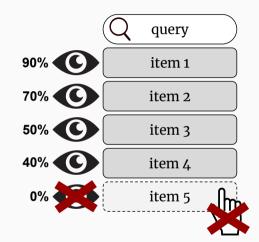


Items that are not displayed cannot be examined:

$$k > K \to P(E = 1 \mid k, d) = 0.$$

Existing approach does **not work in top-k rankings**:

• No clicks to weight!



The Novel Policy-Aware Estimator



If displayed rankings are sampled from a **stochastic policy** π , the click probability can be **conditioned** on the **policy**:

$$P(C = 1 \mid \pi, d) = \sum_{k=1}^{K} \underbrace{\pi(k \mid d)}_{k=1} \underbrace{P(E = 1 \mid k)}_{P(E = 1 \mid k)} \underbrace{P(C = 1 \mid E = 1, d)}_{P(C = 1 \mid E = 1, d)}.$$



If displayed rankings are sampled from a **stochastic policy** π , the click probability can be **conditioned** on the **policy**:

$$P(C = 1 \mid \pi, d) = \sum_{k=1}^{K} \underbrace{\pi(k \mid d)}_{k=1} \underbrace{P(E = 1 \mid k)}_{P(E = 1 \mid k)} \underbrace{P(C = 1 \mid E = 1, d)}_{P(C = 1 \mid E = 1, d)}.$$

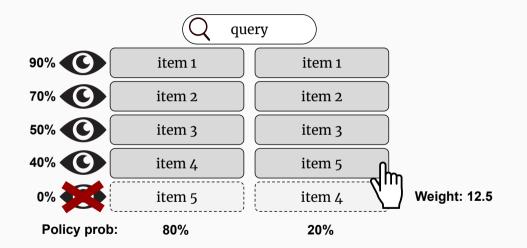
Our **Policy-Aware Estimator** weights conditioned on the policy:

$$\mathsf{relevance}(d) \approx \frac{1}{N} \sum_{i=1}^{N} \frac{c_i(d)}{P(E=1 \mid \pi, d)} = \frac{1}{N} \sum_{i=1}^{N} \frac{c_i(d)}{\sum_{k=1}^{K} \pi(k \mid d) P(E=1 \mid k, d)}$$

Unbiased if every item has a non-zero chance of being displayed in the top-k.

The Novel Policy-Aware Estimator: Visualized





Experiments



Semi-synthetic setup based on commercial LTR datasets:

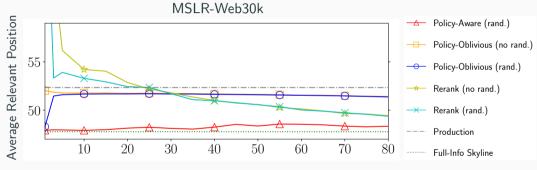
• Yahoo Webscope and MSLR-Web30k.

Displayed top-k rankings based on a pretrained 'production' ranker:

- without randomization, and
- with randomization: random (remaining) item placed on position K.

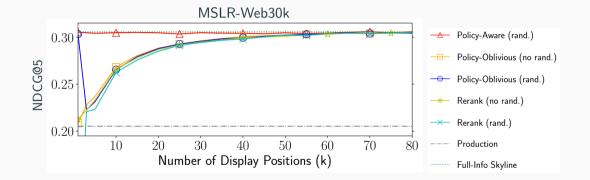
 10^8 clicks based on dataset labels, with added **noise** and **position bias**.

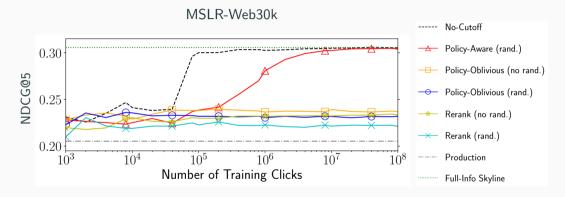




Number of Display Positions (K)







Training from clicks on top-5 rankings.

Conclusion: Part I



Policy-Aware Unbiased Learning to Rank for Top-k Rankings Harrie Oosterhuis and Maarten de Rijke - SIGIR 2020

Main takeaways:

- Existing Counterfactual LTR cannot correct item-selection bias.
- Novel **Policy-Aware** estimator can under mild randomization:
 - by basing propensities on the logging policy instead of individual rankings.

Related work at WWW'20 by Ovaisi et al. (2020) that uses the Heckman's two-stage method.

Part II: Trust Bias



So far, we have assumed the rank-based position bias model (Craswell et al., 2008):

$$P(C = 1 \mid k, d) = \overbrace{P(E = 1 \mid k, d)}^{\text{examination}} \overbrace{P(C = 1 \mid E = 1, d)}^{\text{relevance}}.$$

This assumes that - once examined - each rank is treated similarly by users.

• This ignores the **trust** that users have in ranking systems.

Users are more likely to click on **examined non-relevant** items that are ranked higher (Joachims et al., 2005).



Agarwal et al. (2019a) propose modelling **perceived relevance** and that items at **higher ranks** are **more likely** to be perceived as relevant.

Probability of clicking conditioned on relevance R_i , examination E and rank k:

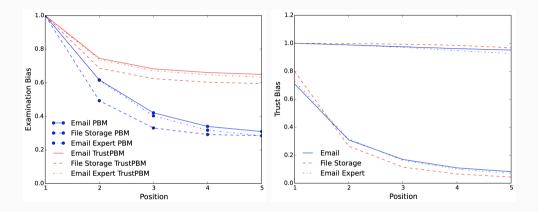
$$\epsilon_k^+ = P(C = 1 \mid R = 1, E = 1, k), \quad \epsilon_k^- = P(C = 1 \mid R = 0, E = 1, k).$$

The probability of a click on item d at rank k:

$$P(C = 1 \mid k, d) = \underbrace{P(E = 1 \mid k, d)}_{\text{examination}} \left(\underbrace{\epsilon_k^+ P(R = 1 \mid d)}_{\text{actually relevant}} + \underbrace{\epsilon_k^- P(R = 0 \mid d)}_{\text{incorrectly perceived relevant}} \right).$$

Emperical Estimate of Trust Bias





Agarwal et al. (2019a) infer these parameters from **real-world user behavior** and show their model is better at predicting user behavior.



We introduce the following compact notation (Vardasbi et al., 2020):

$$P(C=1 \mid k, d) = \alpha_k P(R=1 \mid d) + \beta_k,$$

where

$$\underline{\alpha_k = P(E=1 \mid k, d)(\epsilon_k^+ - \epsilon_k^-)},$$

 $\underbrace{\beta_k = P(E=1 \mid k, d)\epsilon_k^-}_{\mathbf{k}} \ .$

correlation between clicks and relevance

click-through-rate from user trust

We prove it is impossible to correct for trust bias with IPS estimation.



The trust bias model is an affine transformation from relevance to click probabilities:

$$P(C = 1 | k, d) = \alpha_k P(R = 1 | d) + \beta_k.$$

This affine transformation can be **inversed**:

$$P(R = 1 \mid d) = \frac{P(C = 1 \mid k, d) - \beta_k}{\alpha_k}.$$

Based on this observation, we propose the unbiased affine estimator:

$$\mathsf{relevance}(d) \approx \frac{1}{N} \sum_{i=1}^{N} \frac{c_i(d) - \beta_{k(d)}}{\alpha_{k(d)}}.$$

Part III: Unifying Online and Counterfactual LTR



Unbiased Learning to Rank:

• Learning from clicks while correcting for interaction biases.

Online Learning to Rank:

• Correct for bias by **randomizing** results through **online interventions**.

Counterfactual Learning to Rank:

• Infer a model of bias, use it to correct when learning from historical click data.

Radboud University

Position Bias:

- Users are more likely to examine higher ranked results (Craswell et al., 2008).
- Solution: Inverse Propensity Scoring (Joachims et al., 2017).

Item-Selection Bias:

- Users cannot examine items that are not displayed (Ovaisi et al., 2020).
- Solution: Policy-Aware Propensities (Oosterhuis and de Rijke, 2020).

Trust Bias:

- Users are more likely to **incorrectly presume relevance** of higher ranked results (Agarwal et al., 2019a).
- Solution: Apply inverse affine transformation (Vardasbi et al., 2020).

Intervention-Oblivious Estimator



Starting assumption: clicks follow an **affine model**, for item *d* displayed at rank *k*:

$$P(C = 1 \mid d, k) = \alpha_k P(R = 1 \mid d) + \beta_k.$$



Starting assumption: clicks follow an **affine model**, for item *d* displayed at rank *k*:

$$P(C = 1 | d, k) = \alpha_k P(R = 1 | d) + \beta_k.$$

We **condition** the click probability on the **logging policy** π :

$$P(C = 1 \mid d, \pi) = \sum_{k=1}^{K} \pi(k \mid d) (\alpha_k P(R = 1 \mid d) + \beta_k)$$

= $\mathbb{E}_k[\alpha_k \mid d, \pi] P(R = 1 \mid d) + \mathbb{E}_k[\beta_k \mid d, \pi]$



Starting assumption: clicks follow an **affine model**, for item *d* displayed at rank *k*:

$$P(C = 1 | d, k) = \alpha_k P(R = 1 | d) + \beta_k.$$

We **condition** the click probability on the **logging policy** π :

$$P(C = 1 \mid d, \pi) = \sum_{k=1}^{K} \pi(k \mid d) (\alpha_k P(R = 1 \mid d) + \beta_k)$$

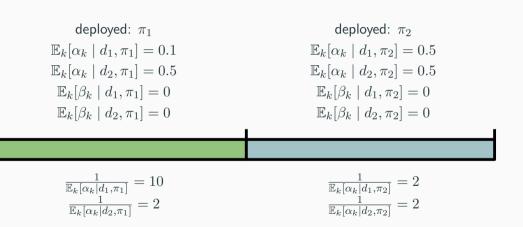
= $\mathbb{E}_k[\alpha_k \mid d, \pi] P(R = 1 \mid d) + \mathbb{E}_k[\beta_k \mid d, \pi]$

The intervention-oblivious estimator is based on the inverse of this transformation:

$$P(R=1 \mid d) = \frac{P(C=1 \mid d, \pi) - \mathbb{E}_k[\beta_k \mid d, \pi]}{\mathbb{E}_k[\alpha_k \mid d, \pi]}.$$

Intervention-Oblivious Estimator: Visualization





25

Intervention-Aware Estimator



Due to **interventions** the logging policy is **updated** during data-gathering. Let Π contain **all logging policies** for each timestep *t*:

$$\Pi = \{\pi_1, \pi_2, \ldots\}.$$



Due to **interventions** the logging policy is **updated** during data-gathering. Let Π contain **all logging policies** for each timestep *t*:

$$\Pi = \{\pi_1, \pi_2, \ldots\}.$$

We can **condition** the click probability on the **set** Π :

$$P(C = 1 \mid d, \Pi) = \frac{1}{|\Pi|} \sum_{\pi_t \in \Pi} \sum_{k=1}^K \pi_t(k \mid d) (\alpha_k P(R = 1 \mid d) + \beta_k)$$

= $\mathbb{E}_k[\alpha_k \mid d, \Pi] P(R = 1 \mid d) + \mathbb{E}_k[\beta_k \mid d, \Pi].$



Due to **interventions** the logging policy is **updated** during data-gathering. Let Π contain **all logging policies** for each timestep *t*:

$$\Pi = \{\pi_1, \pi_2, \ldots\}.$$

We can **condition** the click probability on the **set** Π :

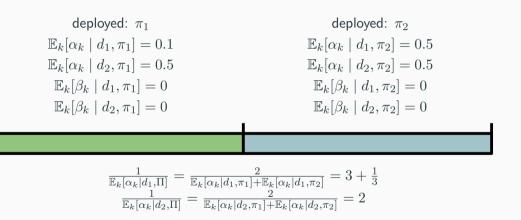
$$P(C = 1 \mid d, \Pi) = \frac{1}{|\Pi|} \sum_{\pi_t \in \Pi} \sum_{k=1}^{K} \pi_t(k \mid d) (\alpha_k P(R = 1 \mid d) + \beta_k)$$

= $\mathbb{E}_k[\alpha_k \mid d, \Pi] P(R = 1 \mid d) + \mathbb{E}_k[\beta_k \mid d, \Pi].$

The intervention-aware estimator is based on the inverse:

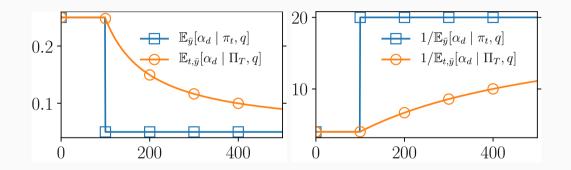
$$P(R=1 \mid d) = \frac{P(C=1 \mid d, \Pi) - \mathbb{E}_k[\beta_k \mid d, \Pi]}{\mathbb{E}_k[\alpha_k \mid d, \Pi]}.$$





27





Example of an intervention at t = 100 and how propensities change as the total number of timesteps increases.

Experiments and Results



Semi-synthetic experiments on the Yahoo! Webscope dataset (Chapelle and Chang, 2011).

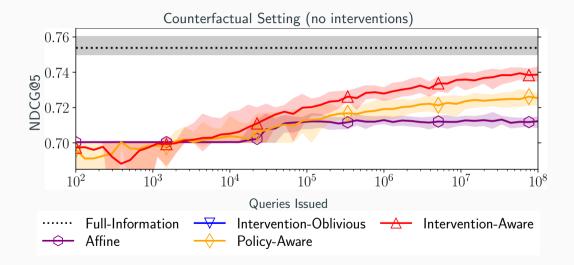
Affine top-5 click model based parameters inferred by Agarwal et al. (2019a).

Both counterfactual and online experiments,

online interventions are spread evenly on a logarithmic scale.

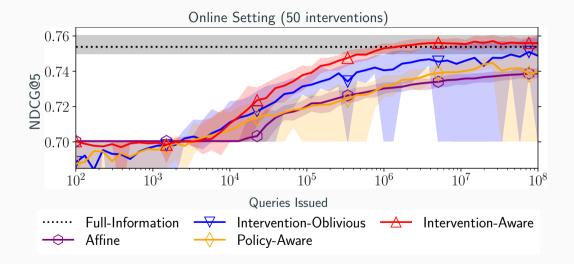
Counterfactual Methods: Counterfactual Comparison





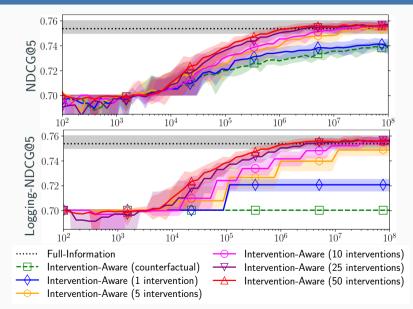
Counterfactual Methods: Online Comparison





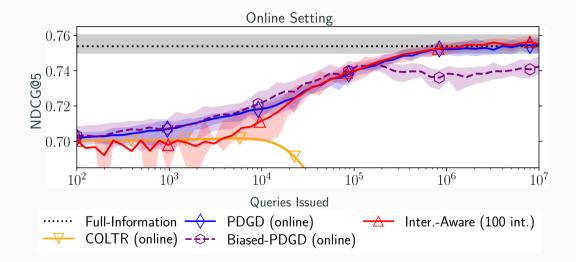
Effect of Online Interventions





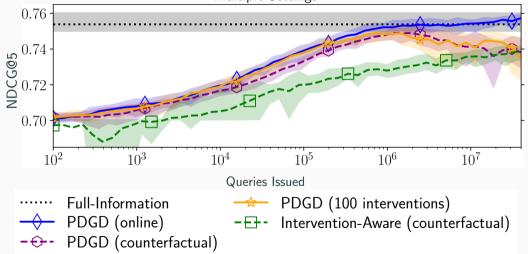
Online Methods: Online Comparison







Multiple Settings



Conclusion: Part III



Unifying Online and Counterfactual Learning to Rank Harrie Oosterhuis and Maarten de Rijke - WSDM 2021

Main Takeaways:

- Intervention-Aware Estimator:
 - Novel counterfactual/online estimator.
 - Most reliable choice for counterfactual learning.
 - Online performance comparable to state-of-the-art.



Unifying Online and Counterfactual Learning to Rank Harrie Oosterhuis and Maarten de Rijke - WSDM 2021

Main Takeaways:

- Intervention-Aware Estimator:
 - Novel counterfactual/online estimator.
 - Most reliable choice for counterfactual learning.
 - Online performance **comparable to state-of-the-art**.
- PDGD is not reliable when not applied fully online.



Unifying Online and Counterfactual Learning to Rank Harrie Oosterhuis and Maarten de Rijke - WSDM 2021

Main Takeaways:

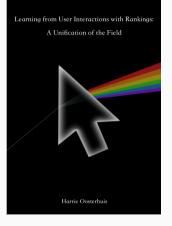
- Intervention-Aware Estimator:
 - Novel counterfactual/online estimator.
 - Most reliable choice for counterfactual learning.
 - Online performance **comparable to state-of-the-art**.
- PDGD is not reliable when not applied fully online.
- A single method that is the best choice for both online and counterfactual learning to rank.

Overview of this Talk



Papers in this talk:

- Policy-Aware Unbiased Learning to Rank for Top-k Rankings
 Harrie Oosterhuis and Maarten de Rijke - SIGIR 2020
- When Inverse Propensity Scoring does not Work: Affine Corrections for Unbiased Learning to Rank Ali Vardasbi, Harrie Oosterhuis and Maarten de Rijke -CIKM 2020
- Unifying Online and Counterfactual Learning to Rank Harrie Oosterhuis and Maarten de Rijke - WSDM 2021





- A. Agarwal, X. Wang, C. Li, M. Bendersky, and M. Najork. Addressing trust bias for unbiased learning-to-rank. In *The World Wide Web Conference*, pages 4–14. ACM, 2019a.
- A. Agarwal, I. Zaitsev, X. Wang, C. Li, M. Najork, and T. Joachims. Estimating position bias without intrusive interventions. In *Proceedings of the Twelfth ACM International Conference on Web Search and Data Mining*, pages 474–482. ACM, 2019b.
- Q. Ai, K. Bi, C. Luo, J. Guo, and W. B. Croft. Unbiased learning to rank with unbiased propensity estimation. In *Proceedings of the 41st International ACM SIGIR conference on Research and Development in Information Retrieval*, pages 385–394. ACM, 2018.
- O. Chapelle and Y. Chang. Yahoo! Learning to Rank Challenge Overview. *Journal of Machine Learning Research*, 14:1–24, 2011.
- N. Craswell, O. Zoeter, M. Taylor, and B. Ramsey. An experimental comparison of click position-bias models. In *Proceedings of the 2008 international conference on web search and data mining*, pages 87–94, 2008.



- T. Joachims, L. Granka, B. Pan, H. Hembrooke, and G. Gay. Accurately interpreting clickthrough data as implicit feedback. In *SIGIR Forum*, pages 154–161. ACM, 2005.
- T. Joachims, A. Swaminathan, and T. Schnabel. Unbiased learning-to-rank with biased feedback. In *Proceedings of the Tenth ACM International Conference on Web Search and Data Mining*, pages 781–789. ACM, 2017.
- H. Oosterhuis and M. de Rijke. Policy-aware unbiased learning to rank for top-k rankings. In *Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 489–498. ACM, 2020.
- Z. Ovaisi, R. Ahsan, Y. Zhang, K. Vasilaky, and E. Zheleva. Correcting for selection bias in learning-to-rank systems. In *Proceedings of The Web Conference 2020*, pages 1863–1873, 2020.
- A. Vardasbi, H. Oosterhuis, and M. de Rijke. When inverse propensity scoring does not work: Affine corrections for unbiased learning to rank. In *Proceedings of the 28th ACM International Conference on Information and Knowledge Management*, 2020.



- X. Wang, M. Bendersky, D. Metzler, and M. Najork. Learning to rank with selection bias in personal search. In *Proceedings of the 39th International ACM SIGIR conference on Research and Development in Information Retrieval*, pages 115–124. ACM, 2016.
- X. Wang, N. Golbandi, M. Bendersky, D. Metzler, and M. Najork. Position bias estimation for unbiased learning to rank in personal search. In *Proceedings of the Eleventh ACM International Conference on Web Search and Data Mining*, pages 610–618. ACM, 2018a.
- X. Wang, C. Li, N. Golbandi, M. Bendersky, and M. Najork. The lambdaloss framework for ranking metric optimization. In *Proceedings of the 27th ACM International Conference on Information and Knowledge Management*, pages 1313–1322. ACM, 2018b.





All content represents the opinion of the author(s), which is not necessarily shared or endorsed by their employers and/or sponsors.