Optimizing Ranking Models in an Online Setting

based on a publication at ECIR 2019

Harrie Oosterhuis and Maarten de Rijke

November 29, 2019

University of Amsterdam

oosterhuis@uva.nl, derijke@uva.nl

https://staff.fnwi.uva.nl/h.r.oosterhuis

Online Learning to Rank

Online Learning to Rank

Online Learning to Rank: **learn by interacting with users**.

Solves most of the problems of expert-annotations:

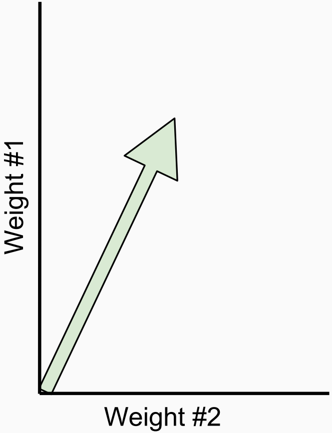
- Interactions are **virtually free** if you have users.
- User **behaviour** gives **implicit feedback**.

These methods have to handle:

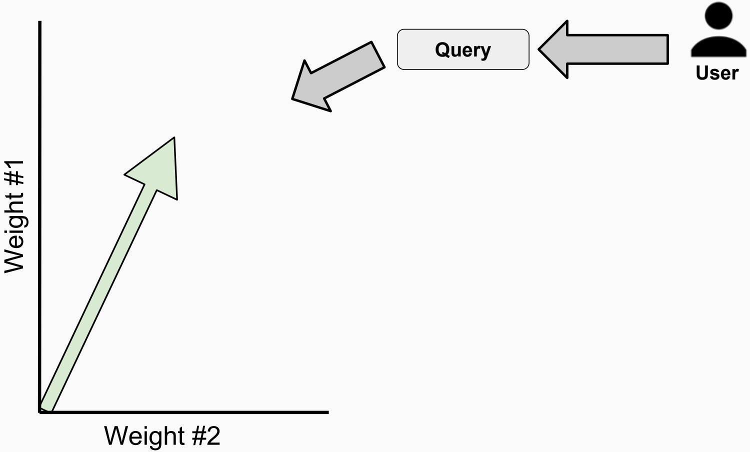
- **Noise**: Users click for **unexpected reasons**.
- **Biases**: Interactions are affected by **position and selection bias**.

Dueling Bandit Gradient Descent

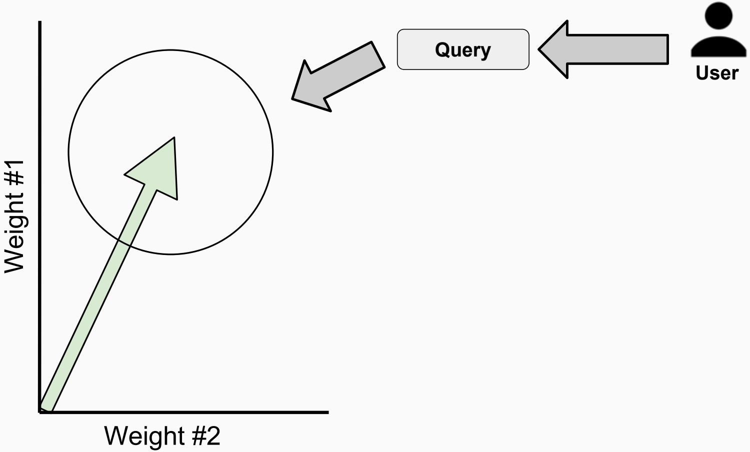
Dueling Bandit Gradient Descent: Visualization



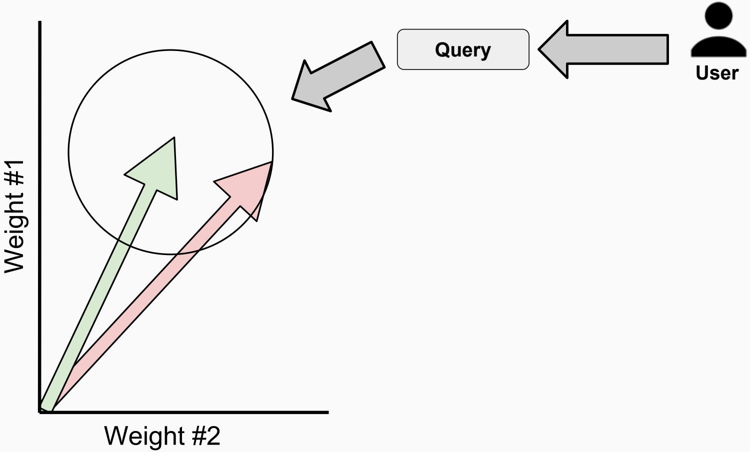
Dueling Bandit Gradient Descent: Visualization



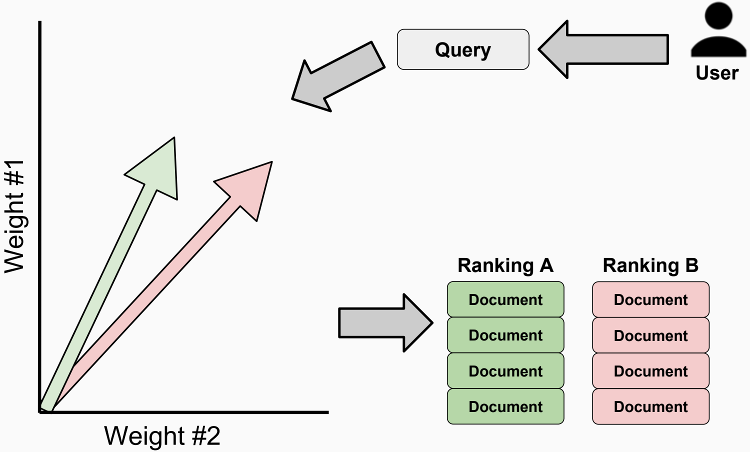
Dueling Bandit Gradient Descent: Visualization



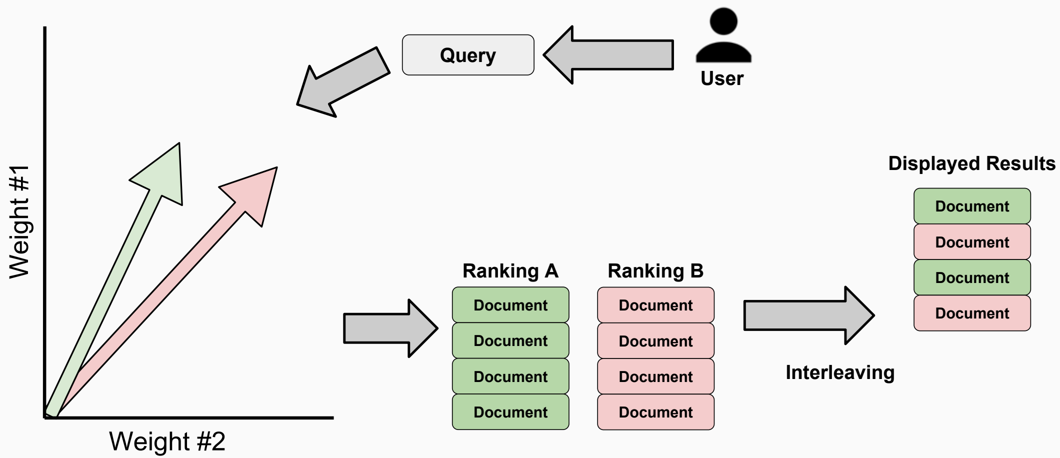
Dueling Bandit Gradient Descent: Visualization



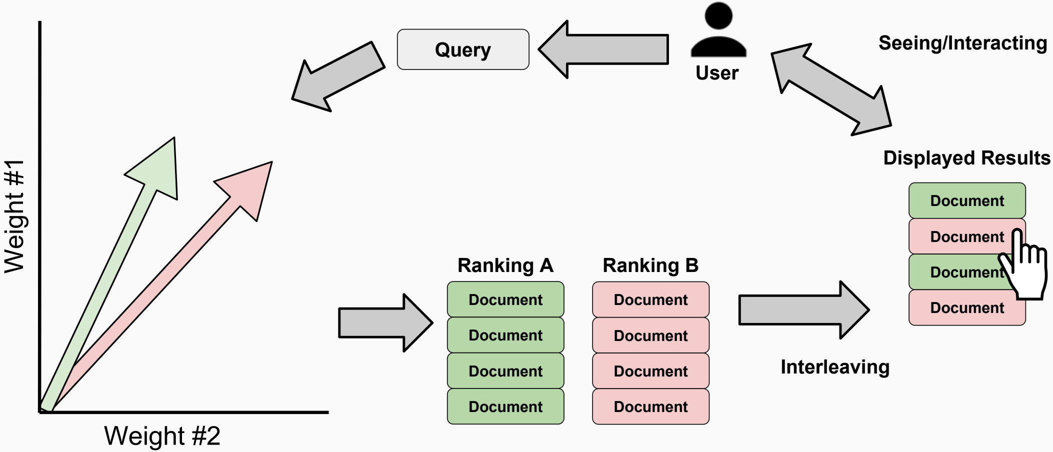
Dueling Bandit Gradient Descent: Visualization



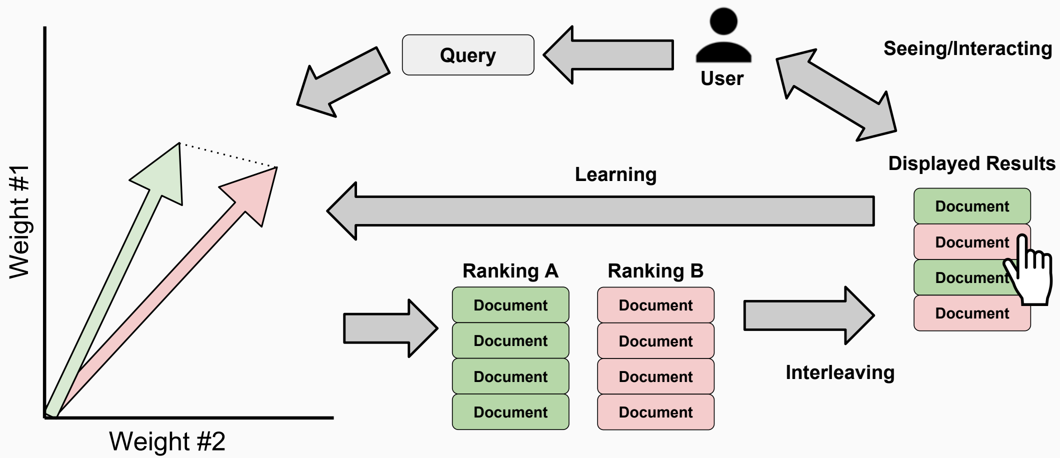
Dueling Bandit Gradient Descent: Visualization



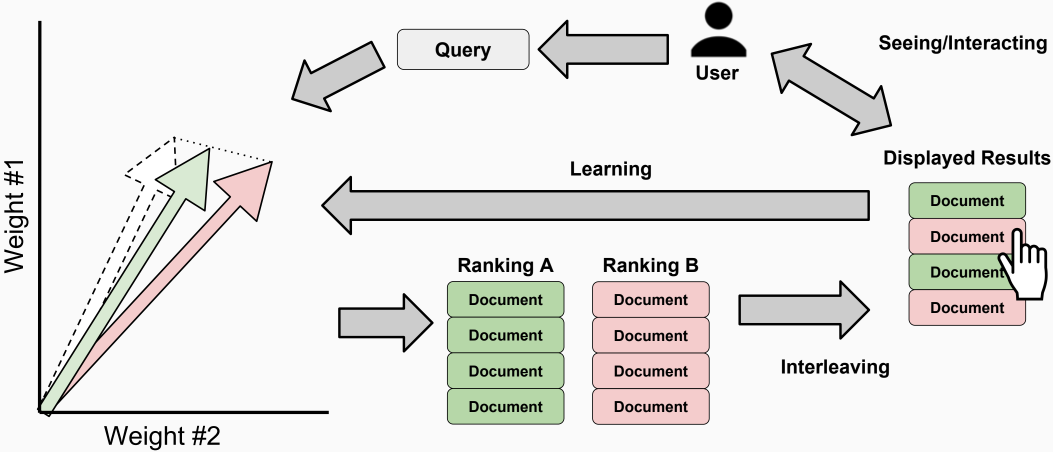
Dueling Bandit Gradient Descent: Visualization



Dueling Bandit Gradient Descent: Visualization



Dueling Bandit Gradient Descent: Visualization



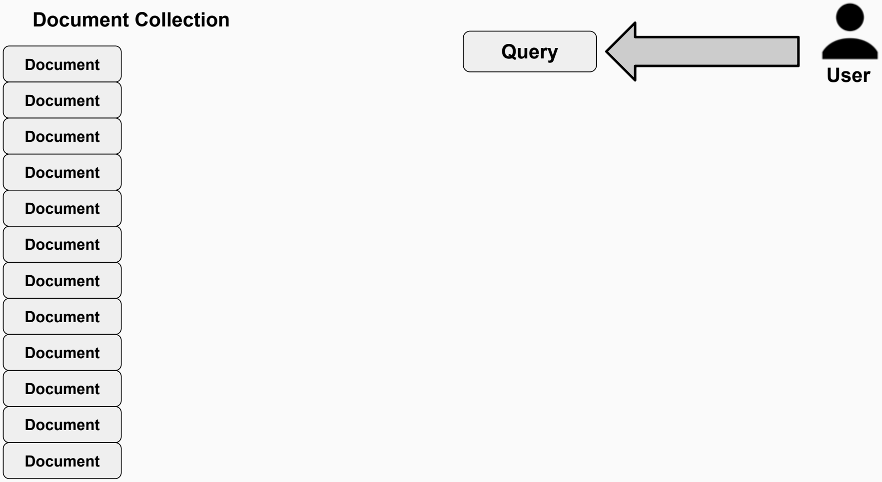
Dueling Bandit Gradient Descent

Basis of the online LTR **field**, virtually everything is an **extension** of DBGD: (Yue and Joachims, 2009; Schuth et al., 2016; Hofmann et al., 2013; Zhao and King, 2016; Wang et al., 2018).

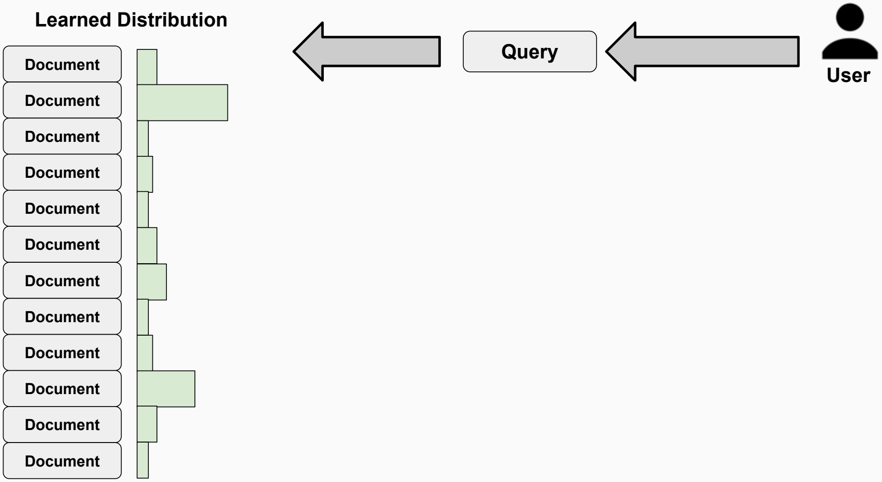
Pairwise Differentiable Gradient Descent

Pairwise Differentiable Gradient Descent: Visualization

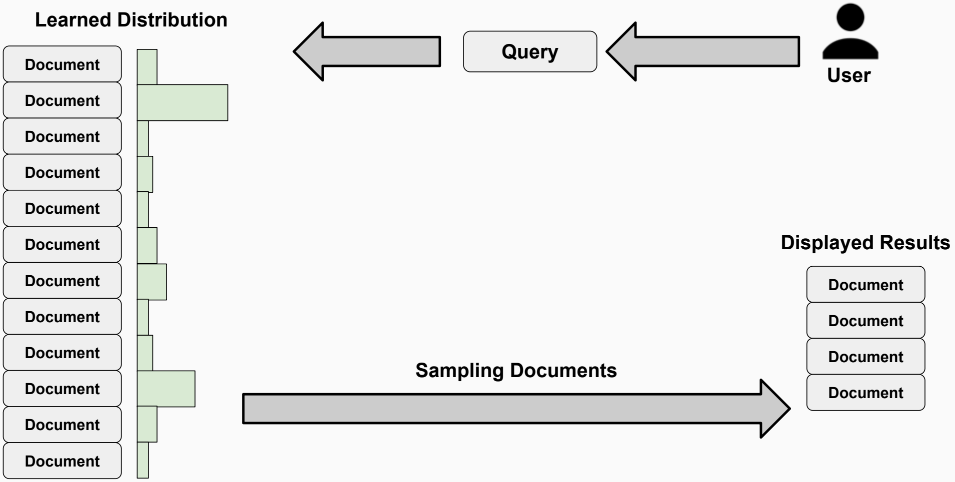
Pairwise Differentiable Gradient Descent: Visualization



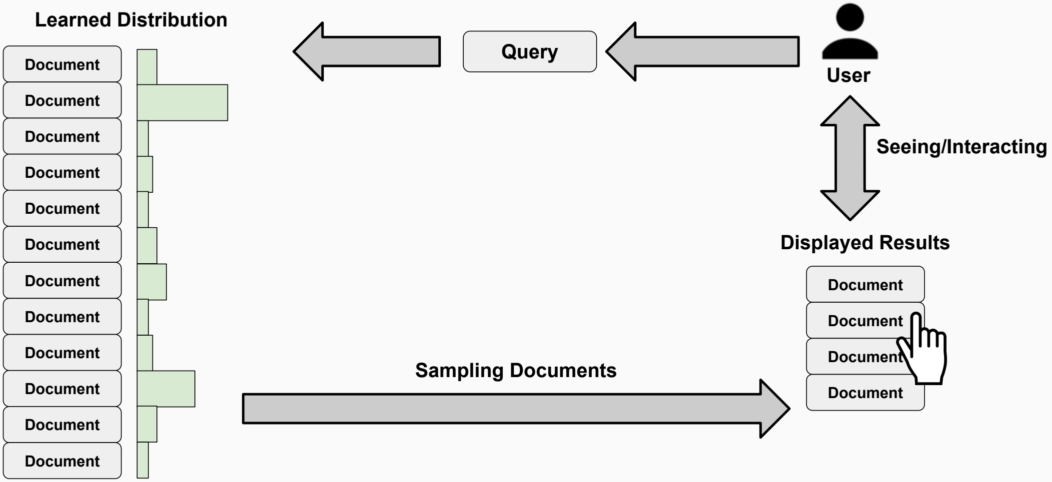
Pairwise Differentiable Gradient Descent: Visualization



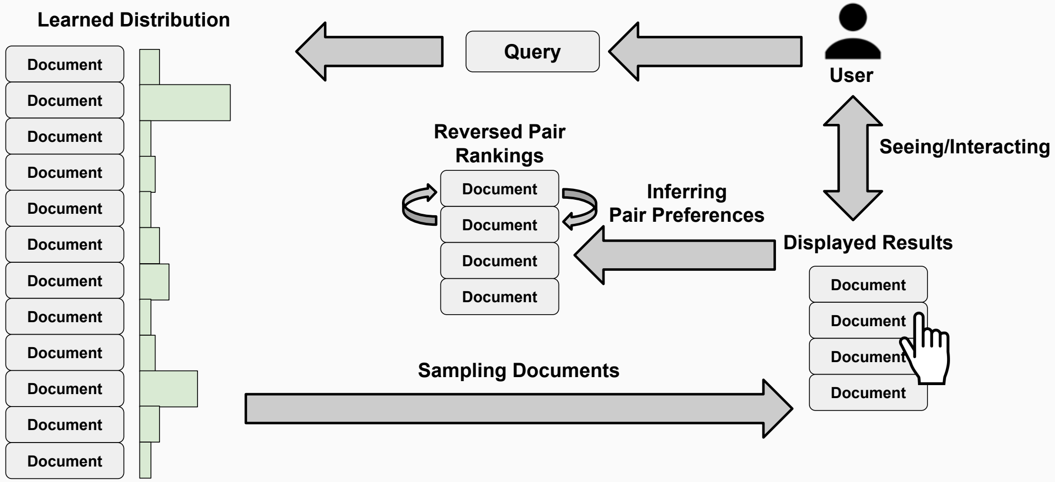
Pairwise Differentiable Gradient Descent: Visualization



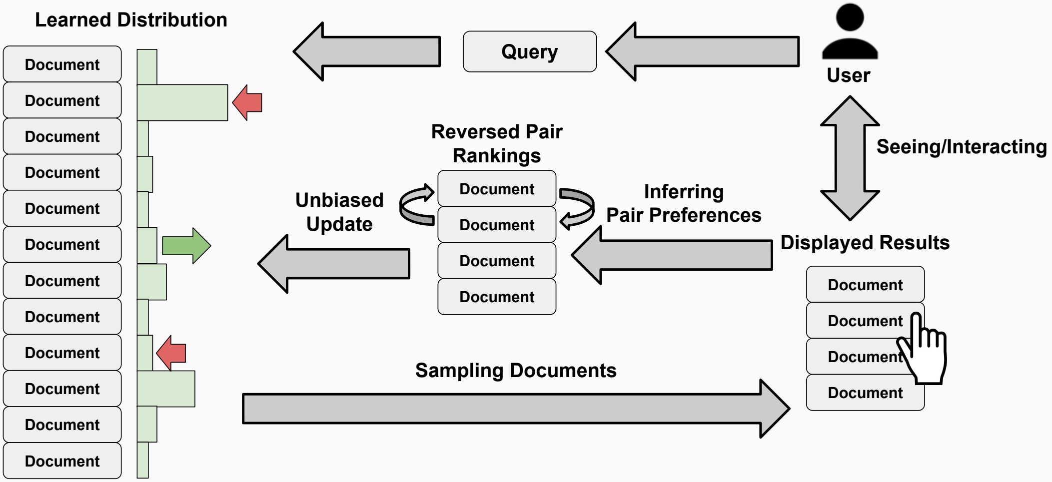
Pairwise Differentiable Gradient Descent: Visualization



Pairwise Differentiable Gradient Descent: Visualization



Pairwise Differentiable Gradient Descent: Visualization



Limitations of Existing Results

Limitations of Previous Existing Results

Comparison based on **previous work** (Oosterhuis and de Rijke, 2018):

- Pairwise Differentiable Gradient Descent **outperforms** DBGD.
- Dueling Bandit Gradient Descent has **proven sublinear regret bounds**.

Problems with the current state of affairs:

- Past comparisons are based on **low-noise, cascading click models**, Pairwise Differentiable Gradient Descent **assumes cascading behaviour!**
- There is a **conflict** between the **proven regret bounds** of Dueling Bandit Gradient Descent and its **observed performance**.

Our contribution

First, we **critically look** at the regret bounds of Dueling Bandit Gradient Descent:

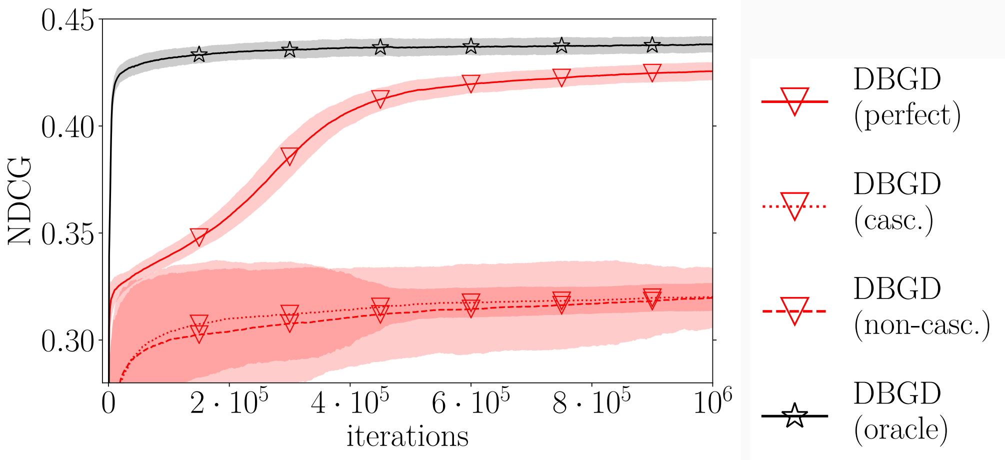
- We **prove** that its **assumptions cannot be true** for standard ranking models.

Then we reproduce the comparison under different conditions using simulation:

- **Both cascading** and **non-cascading** click behaviour.
- Simulated conditions ranging from **ideal to extremely difficult**:
    - **ideal:** no noise, no position bias,
    - **near-random:** mostly noise, very high position bias.

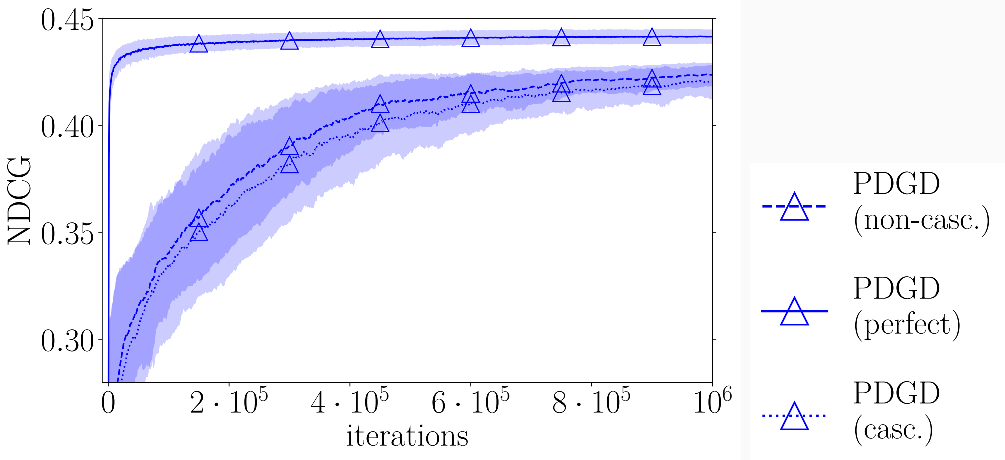
Experimental Results

Comparison: Dueling Bandit Gradient Descent



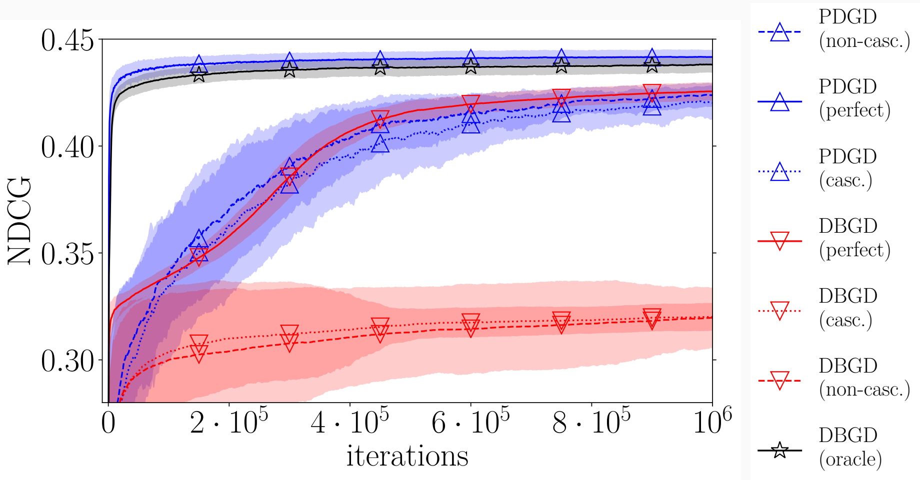
Results from simulations on the MSLR-WEB10k dataset.

Comparison: Pairwise Differentiable Gradient Descent



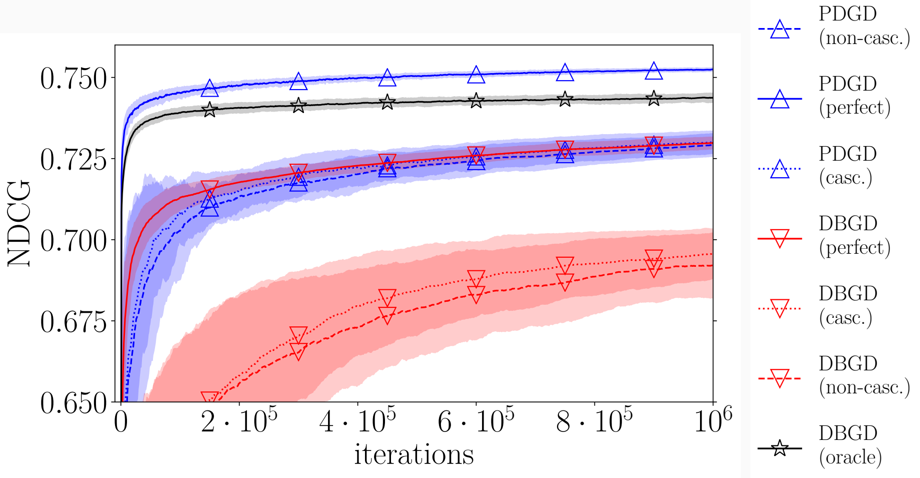
Results from simulations on the MSLR-WEB10k dataset.

Comparison: Complete



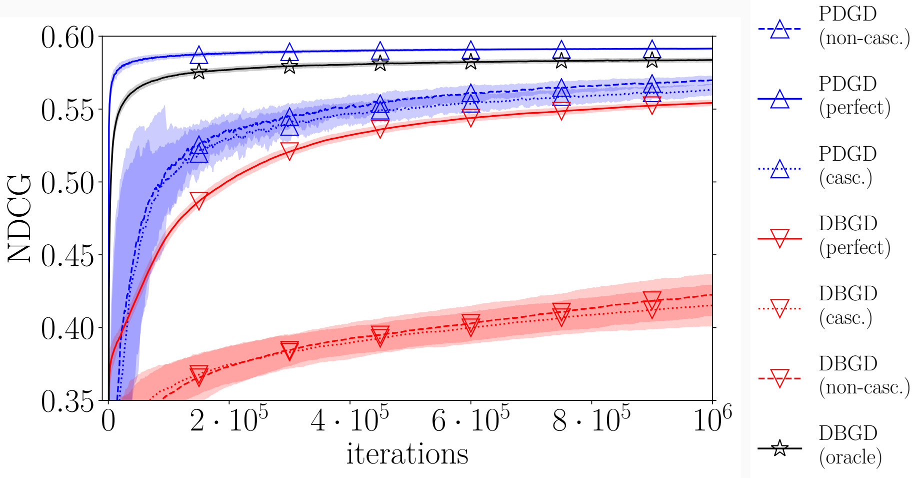
Results from simulations on the MSLR-WEB10k dataset.

Comparison: Complete



Results from simulations on the Yahoo Webscope dataset.

Comparison: Complete



Results from simulations on the Istella dataset.

Conclusion

Conclusion

In this reproducibility paper we have:

- shown that an **existing proof** for regret bounds is **unsound**.
- **Reproduced a comparison** between Pairwise Differentiable Gradient Descent and Dueling Bandit Gradient Descent, and **generalized their conclusions** from ideal circumstances to extremely difficult circumstances.
- Shown that **under all experimental conditions** we could simulate, Pairwise Differentiable Gradient Descent **outperforms** previous methods by **large margins**.

Please continue our work: `https://github.com/HarrieO/OnlineLearningToRank`

K. Hofmann, A. Schuth, S. Whiteson, and M. de Rijke. Reusing historical interaction data for faster online learning to rank for ir. In *Proceedings of the sixth ACM international conference on Web search and data mining*, pages 183–192. ACM, 2013.

H. Oosterhuis and M. de Rijke. Differentiable unbiased online learning to rank. In *Proceedings of the 27th ACM International Conference on Information and Knowledge Management*, pages 1293–1302. ACM, 2018.

A. Schuth, H. Oosterhuis, S. Whiteson, and M. de Rijke. Multileave gradient descent for fast online learning to rank. In *Proceedings of the Ninth ACM International Conference on Web Search and Data Mining*, pages 457–466. ACM, 2016.

H. Wang, R. Langley, S. Kim, E. McCord-Snook, and H. Wang. Efficient exploration of gradient space for online learning to rank. *arXiv preprint arXiv:1805.07317*, 2018.

Y. Yue and T. Joachims. Interactively optimizing information retrieval systems as a dueling bandits problem. In *Proceedings of the 26th Annual International Conference on Machine Learning*, pages 1201–1208. ACM, 2009.

T. Zhao and I. King. Constructing reliable gradient exploration for online learning to rank. In *Proceedings of the 25th ACM International on Conference on Information and Knowledge Management*, pages 1643–1652. ACM, 2016.

Acknowledgments

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