Radbound University

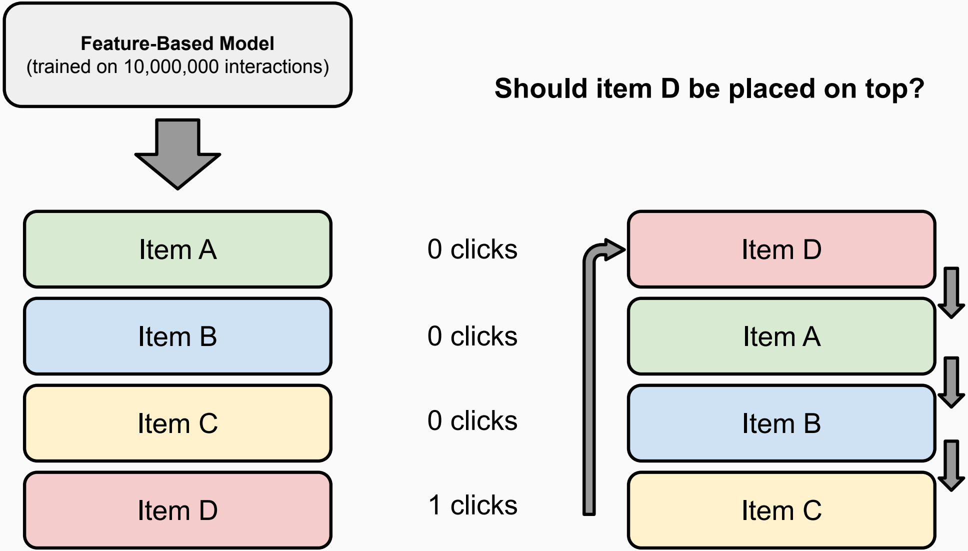
Robust Generalization and Safe Query-Specialization in Counterfactual Learning to Rank

Harrie Oosterhuis[1], Maarten de Rijke[2,3]

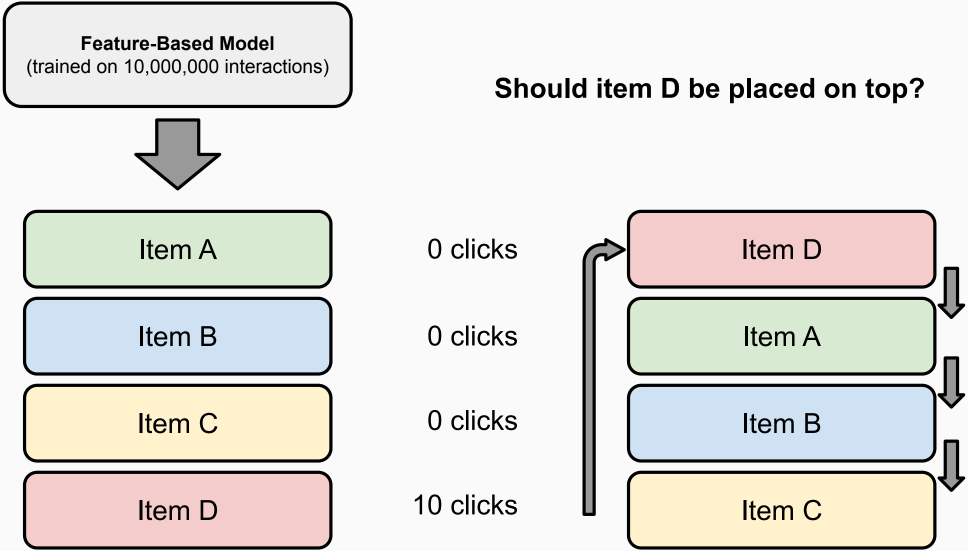
April 20, 2021

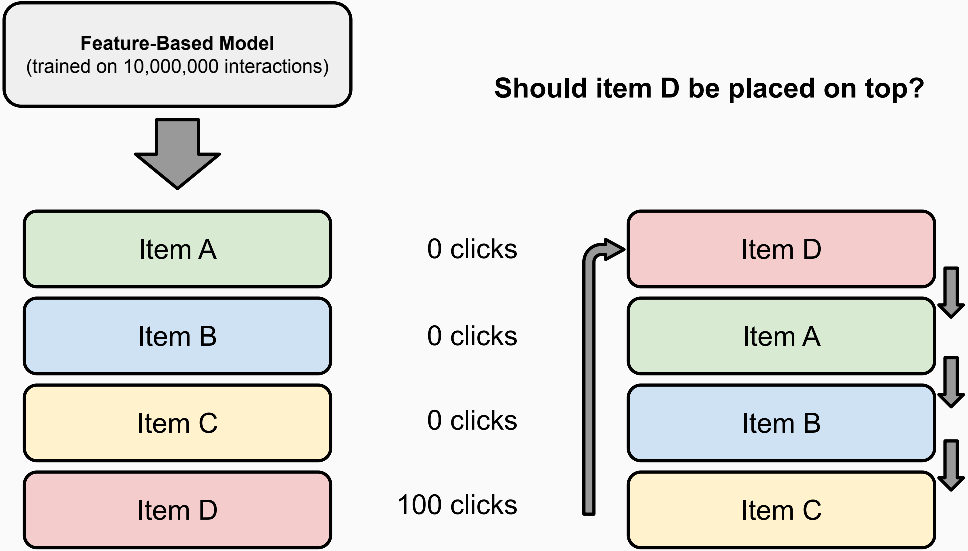
Radboud University[1], University of Amsterdam[2], Ahold Delhaize[3]
harrie.oosterhuis@ru.nl, derijke@uva.nl
https://twitter.com/HarrieOos

A Dilemma

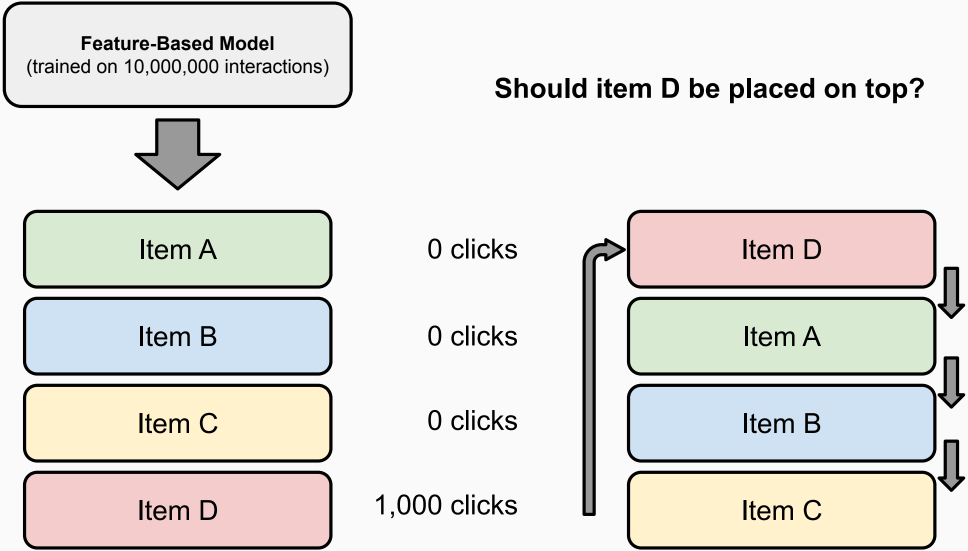


Should item D be placed on top?

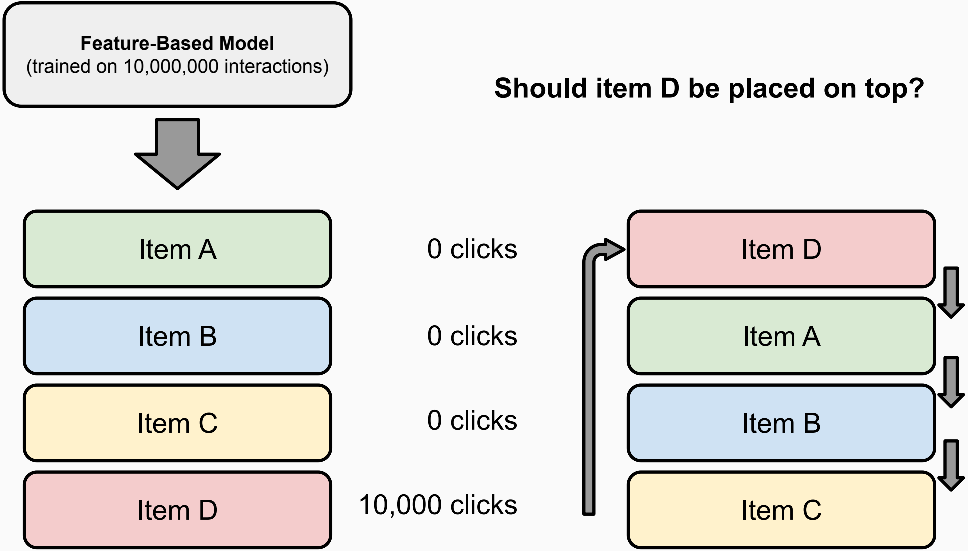

Feature-Based Model
(trained on 10,000,000 interactions)
Should item D be placed on top?
Item A
Item B
Item C
Item D
0 clicks
0 clicks
0 clicks
10 clicks
Item D
Item A
Item B
Item C


Feature-Based Model
(trained on 10,000,000 interactions)
Should item D be placed on top?
Item A
Item B
Item C
Item D
0 clicks
0 clicks
0 clicks
100 clicks
Item D
Item A
Item B
Item C

Dilemma


Feature-Based Model
(trained on 10,000,000 interactions)
Should item D be placed on top?
Item A
Item B
Item C
Item D
0 clicks
0 clicks
0 clicks
1,000 clicks
Item D
Item A
Item B
Item C

Dilemma


Feature-Based Model
(trained on 10,000,000 interactions)
Should item D be placed on top?
Item A
Item B
Item C
Item D
0 clicks
0 clicks
0 clicks
10,000 clicks
Item D
Item A
Item B
Item C

Introduction: Counterfactual Learning to Rank

Counterfactual Learning to Rank:

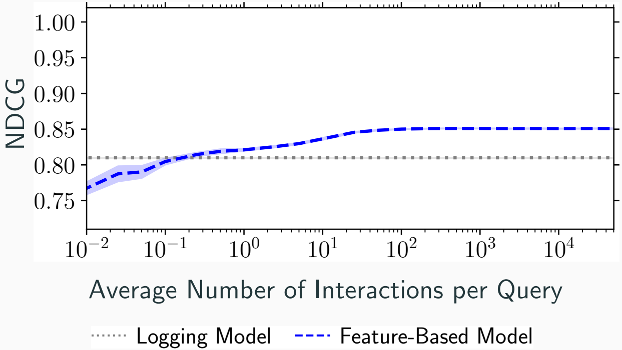
- **Learning from clicks** while correcting for interaction biases caused during gathering of the data (Joachims et al., 2017).

We correct for Position Bias (Craswell et al., 2008) using the policy-aware approach (Oosterhuis and de Rijke, 2020):

$$\hat{R}_{IPS}(d) = \frac{1}{N}\sum_{i=1}^{N}\frac{c_i(d)}{\rho_d}.$$

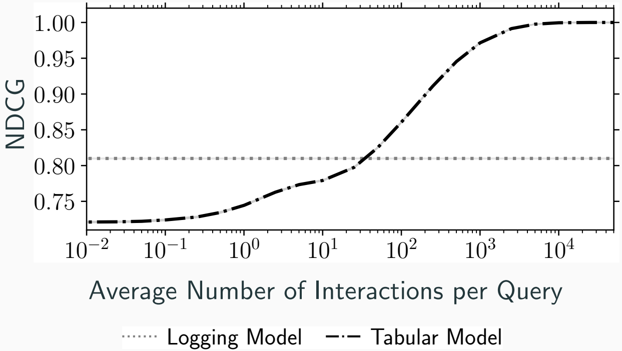
The estimator $\hat{R}_{IPS}$ can be used to unbiasedly estimate a **ranking loss**, e.g. DCG-loss.

For instance, we can optimize a **linear model**:



Performance on the Yahoo! Webscope LTR dataset with simulated biased and noisy clicks.

A **tabular model** simply stores the $\hat{R}_{IPS}$ values and ranks accordingly:



Performance on the Yahoo! Webscope LTR dataset with simulated biased and noisy clicks.

Model Trade-Off

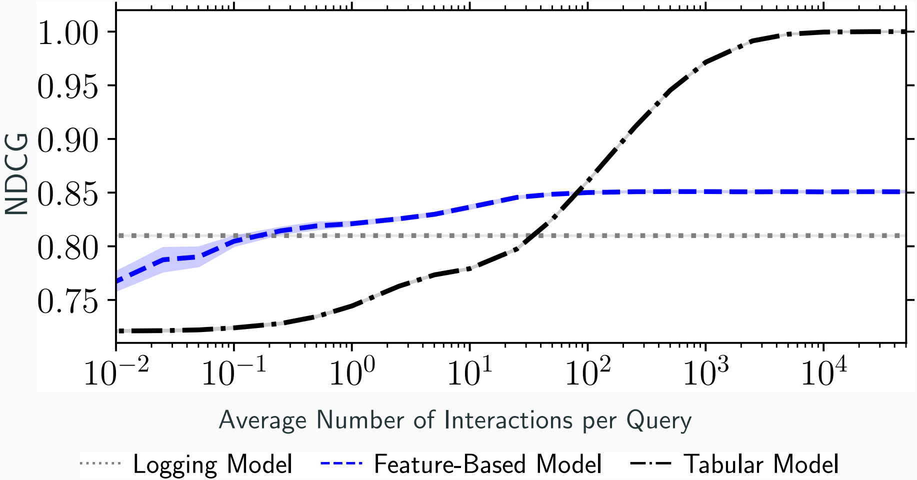
Feature-Based Models:

- **Generalized performance:** Robust over all queries.
- Can be applied to **previously unseen queries**.
- Performance is often **limited** by the quality of the available **features**.

Tabular Models:

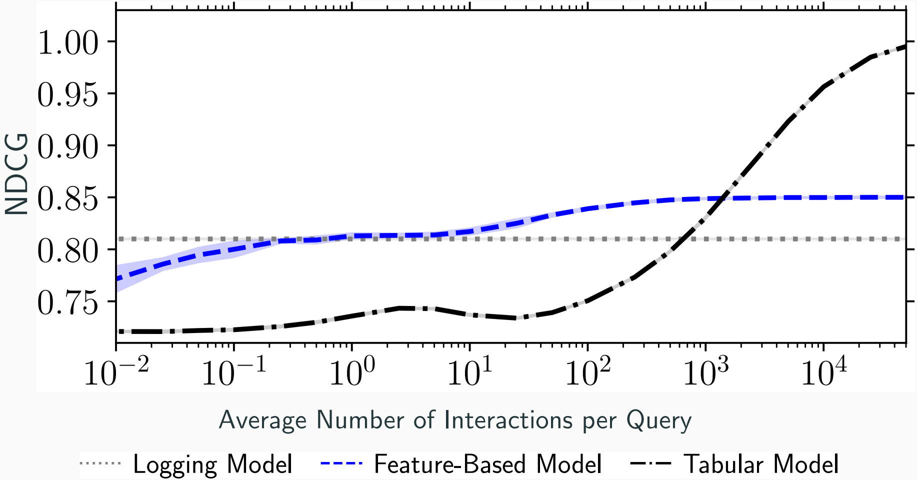
- **Specialized performance**: Independent behavior per query.
- **Cannot** be applied to **previously unseen queries**.
- Performance is **not limited by features**, can learn any possible ranking behavior.

Model Trade-Off Visualized



Performance on the Yahoo! Webscope LTR dataset with simulated biased and noisy clicks.

Model Trade-Off Visualized



Performance on the Yahoo! Webscope LTR dataset with simulated biased and very noisy clicks.

Goal

We want to have both:

- the **safe robust behavior** of feature-based models,
- the **high-performance at convergence** of tabular models,

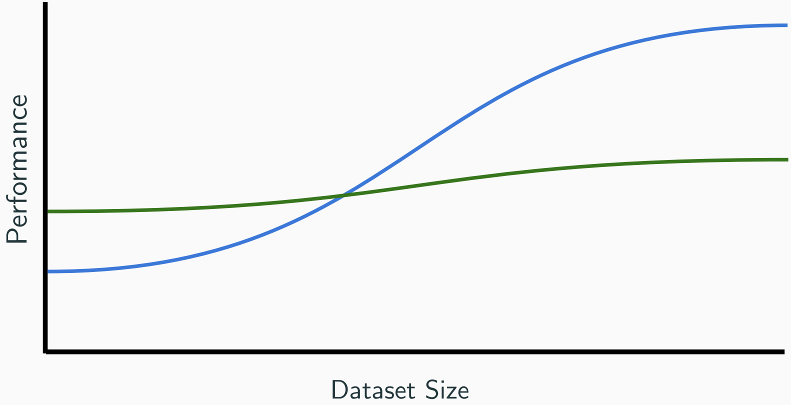
and avoid

- the **detrimental initial performance** of tabular models.

Performance Bounds

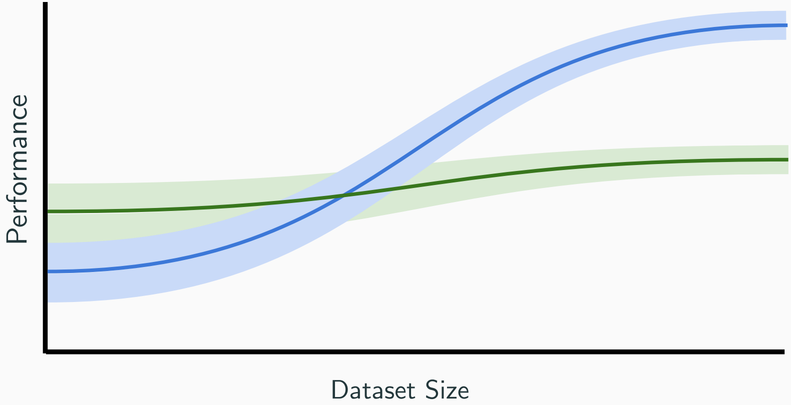
SEA Bound on Performance

Jagerman et al. (2020) introduce the Safe Exploration Algorithm:

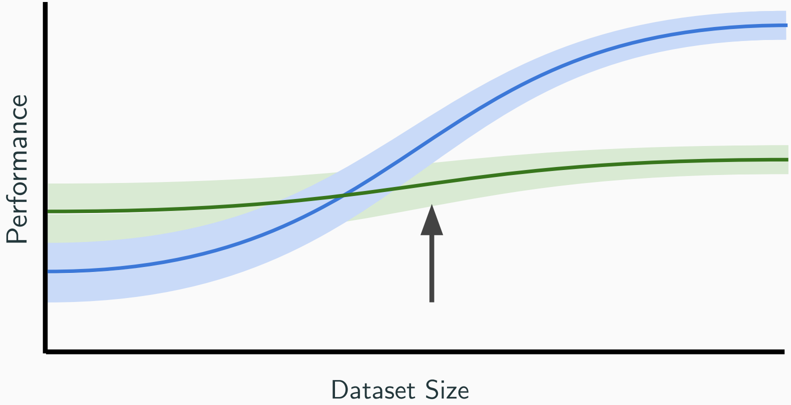


SEA Bound on Performance

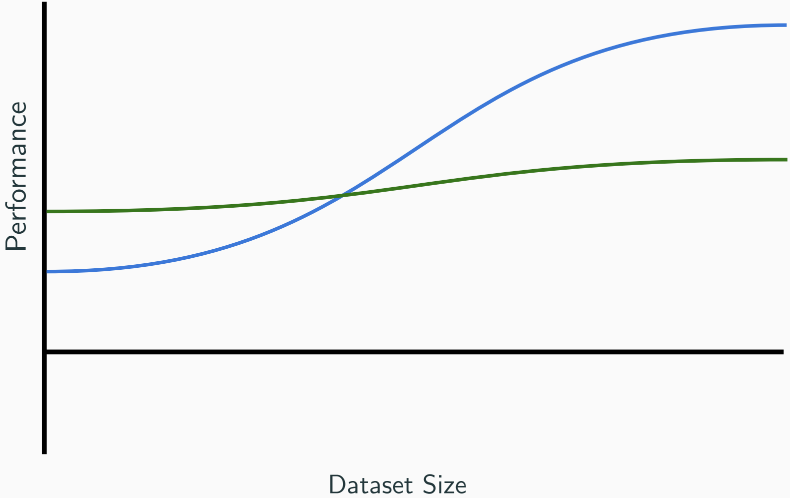
Jagerman et al. (2020) introduce the Safe Exploration Algorithm:



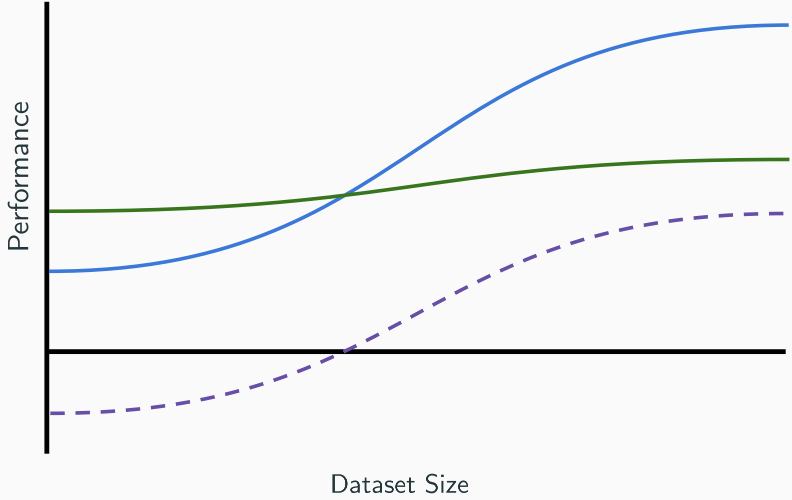
Jagerman et al. (2020) introduce the Safe Exploration Algorithm:



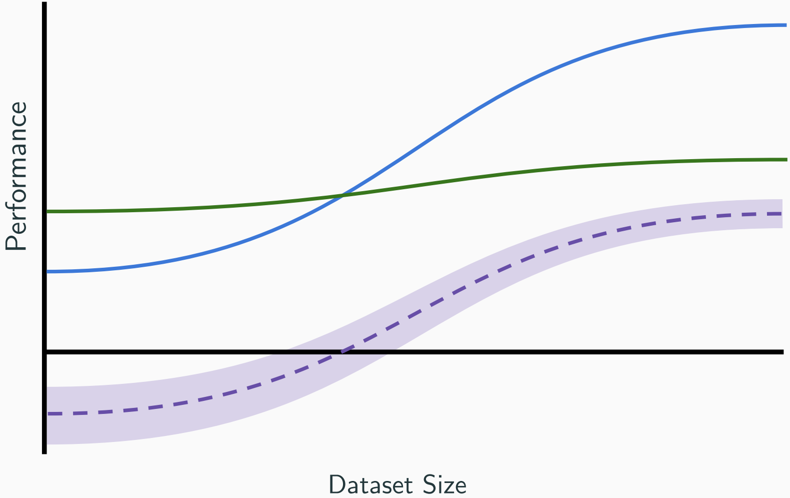
We introduce a new approach that bounds relative performance:



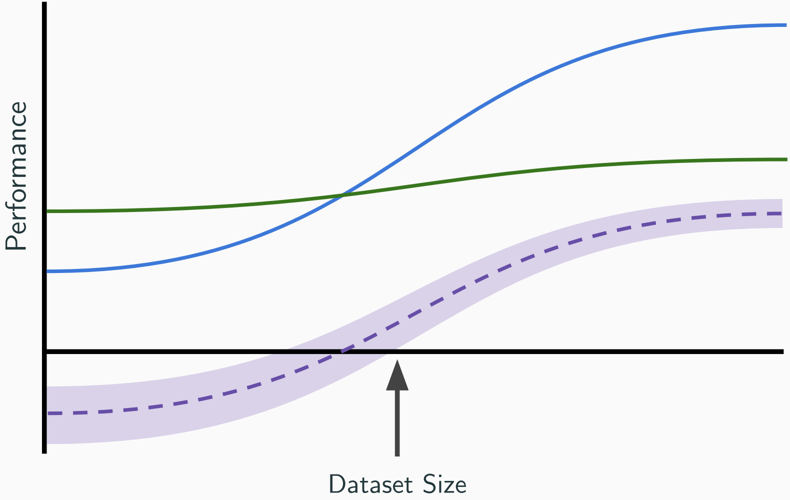
We introduce a new approach that bounds relative performance:



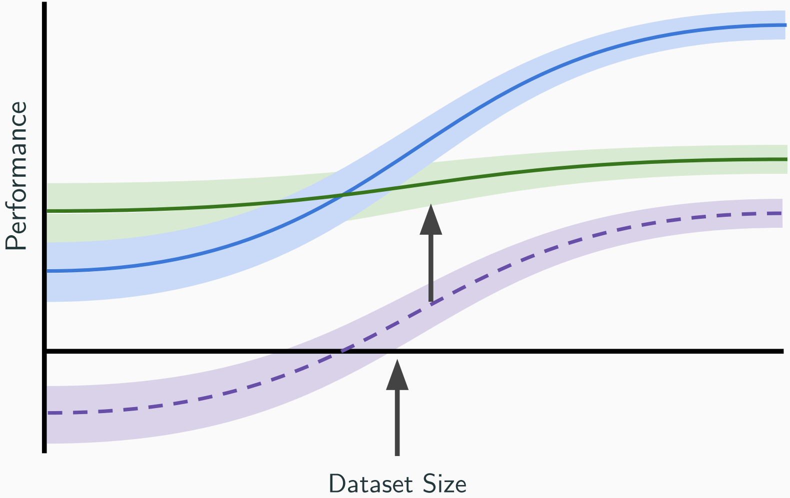
We introduce a new approach that bounds relative performance:



We introduce a new approach that bounds relative performance:



We introduce a new approach that bounds relative performance:



The Generalization and Specialization Framework

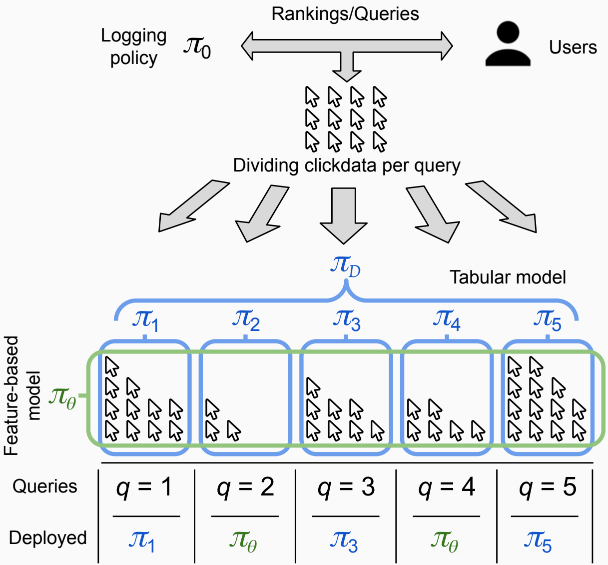
Optimization Phase

- train a **feature-based model** on data over **all queries**,
- compute the values for the **tabular model**.

Serving Procedure

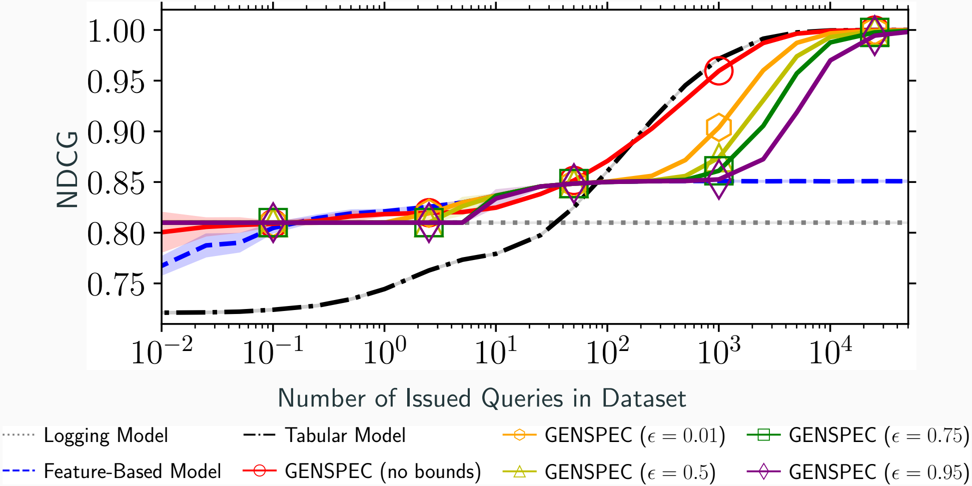
- choose between the **logging policy** and **feature-based model** according to bound computed over **all data**,
- then choose between **tabular model** and **previous choice** according to bound computed only on data for the **specific query**.

GENSPEC overview

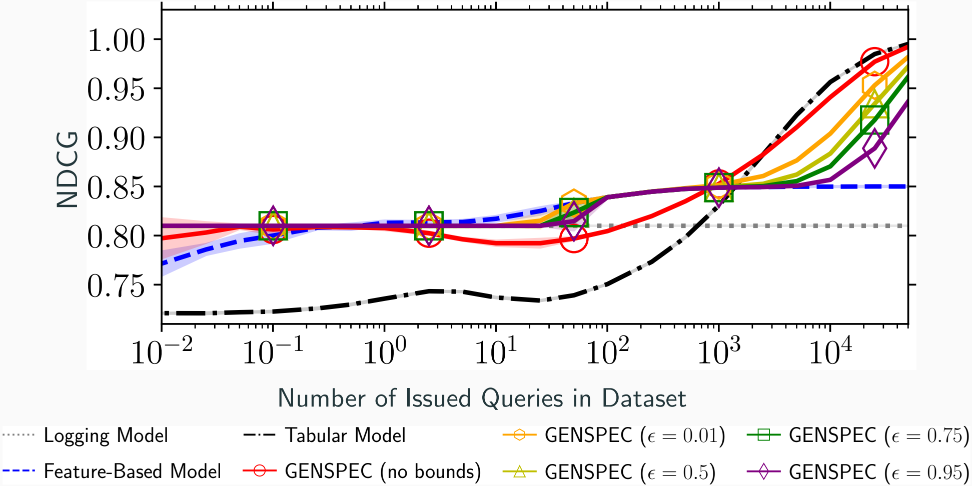


Experimental Results

Performance on the Yahoo! Webscope Dataset



Experiment with simulated biased and noisy clicks.



Experiment with simulated biased and very noisy clicks.

Conclusion

Conclusion

Different models have different advantages and risks:

- **feature-based:** robust generalized performance,
- **tabular:** high performance at convergence, initial detrimental performance.

We introduced **the Generalization and Specialization (GENSPEC) framework:**

- optimizes two models for **generalization** and **specialization**,
- uses performance bounds to **safely choose** to deploy **per query**.

We can have both **robust generalization** and **safe query-specialization** in counterfactual learning to rank.

Continue our work: `https://github.com/HarrieO/2021WWW-GENSPEC`

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. ACM, 2008.

R. Jagerman, I. Markov, and M. de Rijke. Safe exploration for optimizing contextual bandits. *ACM Transactions on Information Systems*, 38(3):Article 24, 2020.

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, 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*. ACM, 2020.

Acknowledgments

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