Unidentified and Confounded?

Understanding Two-Tower Models for Unbiased Learning to Rank

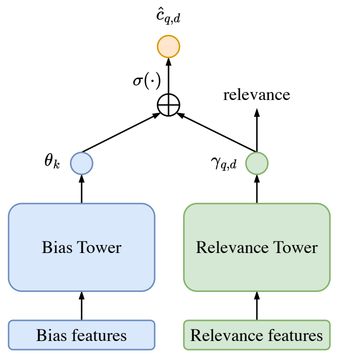
Philipp Hager (UvA), Onno Zeoter (Booking.com), and Maarten de Rijke (UvA)

I. The better the production system, the worse your next two-tower model?

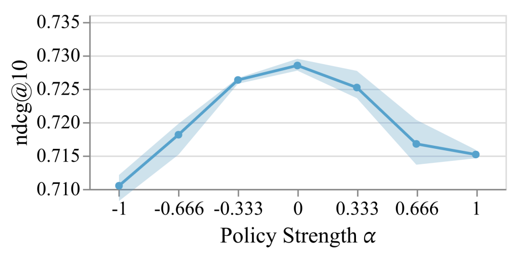
Additive two-tower models are neural architectures to address position bias in click data:

$$P(C = 1 \mid q, d, k) = \sigma(\theta_k + \gamma_{q,d}),$$

and a popular unbiased learning to rank technique in industry settings.



Recent work found that training two-tower models on data collected by strong production systems leads to declining ranking performance and inflated bias estimations [1, 2].

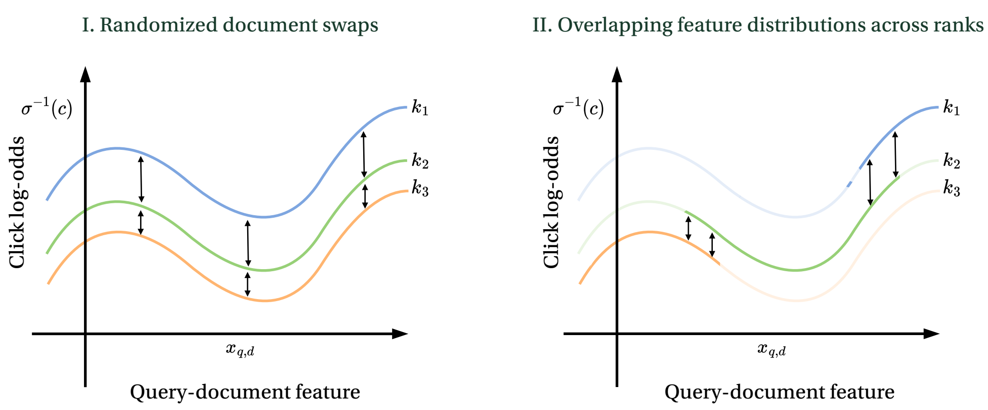


Two-tower models trained on deterministic logging policies of varying strengths (α) on MSLR30K: $\alpha = 1$ represents sorting by expert annotations, $\alpha = 0$ random sorting, and $\alpha = -1$ inversely ranking from least to most relevant.

Are these observations (that we can replicate) due to **logging policy confounding** [1, 2], **model identifiability** issues [3], or **something else?**

II. Identifiability: When can we recover model parameters from observed data?

Our work shows that two-tower models can be identified from:



I. Identification through randomization: Two-tower models are identifiable (up to a constant) when observing **document swaps** across positions*.

II. Identification through overlapping features: When generalizing over shared query-document features:

$$P(C = 1 \mid q, d, k) = \sigma(f(x_{q,d}) + \theta_k),$$

we need **overlapping support in our feature distributions** between positions*:

$$\text{supp}(P(x \mid k)) \cap \text{supp}(P(x \mid k')) \neq \emptyset,$$

and a **continuous relevance tower.**

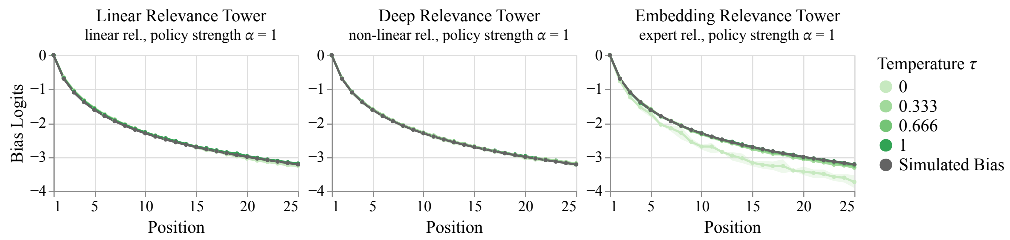
*such that all positions form a connected graph [3].

III. Influence beyond identifiability?

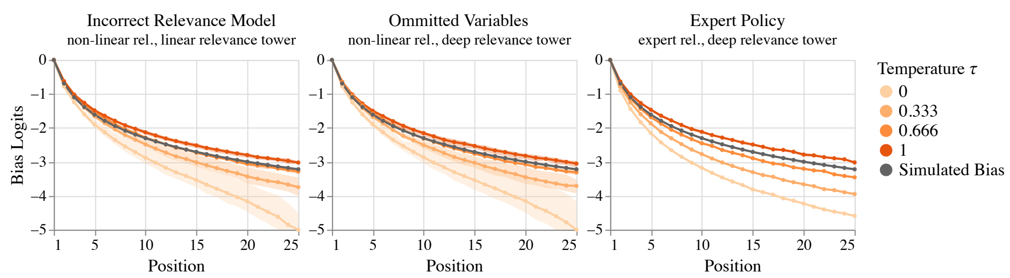
Logging policies impact identifiability by collecting either enough document swaps or overlapping features. But is there an influence beyond identifiability?

$$\frac{\partial \mathcal{L}}{\partial \theta_k} = \sum_q P(q) \sum_d \underbrace{\pi(d, k \mid q)}_{\text{Logging Policy}} \big[\underbrace{P(C=1 \mid q, d, k)}_{\text{Observed CTR}} - \underbrace{\sigma(\theta_k + \gamma_{q,d})}_{\text{Predicted CTR}}\big] = 0.$$

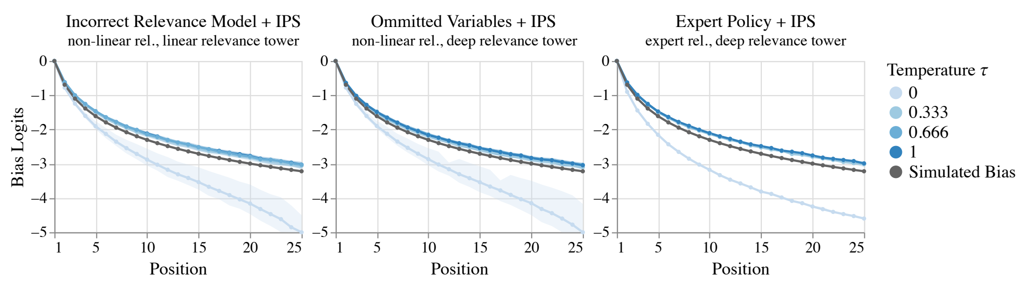
Logging policies have no influence on well-specified and identified two-tower models:



Logging policies can amplify biases in misspecified two-tower models:



We propose an IPS technique that can dampen the effect, but cannot fully remove bias from model misspecification:



IV. Takeaways

- **Identifiability:** Collect document swaps or ensure feature overlap across positions.
- **Misspecification:** Logging policies have no impact on well-specified models, but can amplify bias in misspecified ones.
- **Residuals:** Monitor residuals for correlations between prediction errors and logging policy to detect model misspecification.
- **Simulation:** Never sort by expert annotations as this introduces omitted variable bias. Results from [1] are mostly a simulation artifact.

Our paper contains many more tips for practitioners!

References

[1] Zhang et al. Towards Disentangling Relevance and Bias in Unbiased Learning to Rank. In KDD 2023.
[2] Luo et al. Unbiased Learning-to-Rank Needs Unconfounded Propensity Estimation. In SIGIR 2024.
[3] Chen et al. Identifiability Matters: Revealing the Hidden Recoverable Condition in Unbiased Learning to Rank. In ICML 2024.