Correcting for Recency Bias in Job Recommendation

Ruey-Cheng Chen[1], Qingyao Ai[2], Gaya Jayasinghe[3], W. Bruce Croft[1]

[1]RMIT University [2]University of Utah [3]GO1

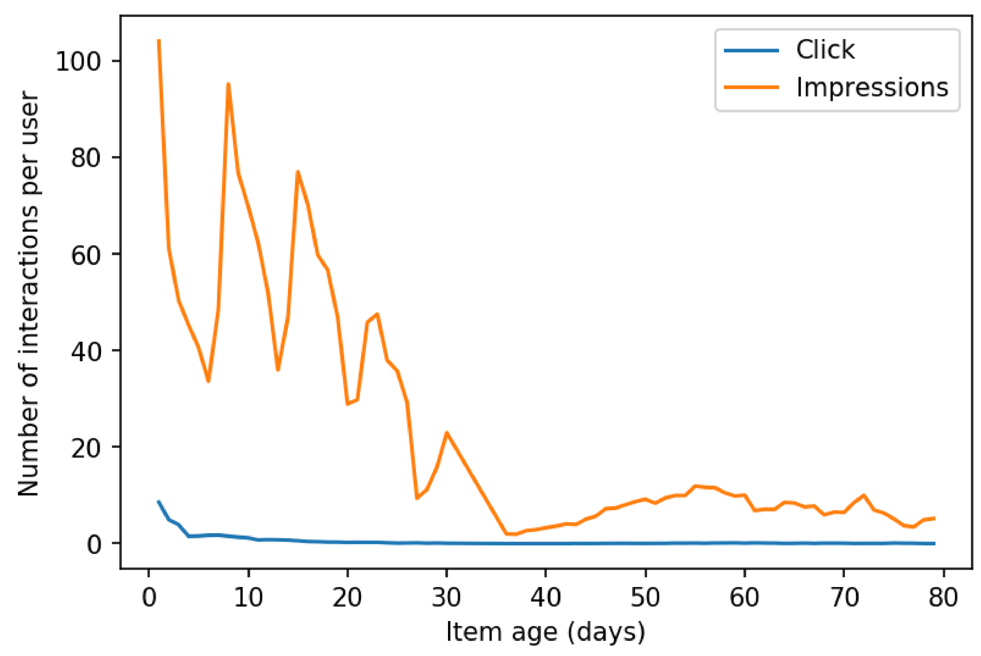
Abstract

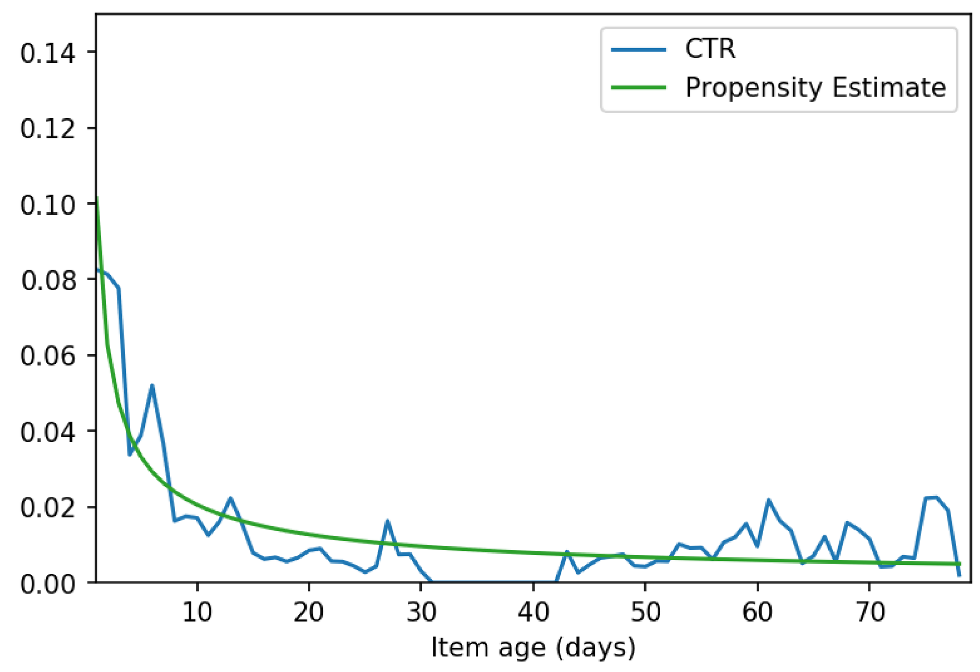
- Users prefer interacting more with fresh content in temporally associated domains, e.g. news search or job seeking.
- Recently published content can be over-represented in the model training phase if data is collected under this influence.
- Consequently, recent content is more likely to be recommended by the model, thereby reinforcing this inherent bias which we called *recency bias*.
- We show that, by actively treating for the recency bias, one can improve the quality of recommendation significantly over a recent neural collaborative filtering model for job recommendation.

What is Recency Bias?

Item exposure is not balanced across different item ages

- "Temporal domain": news/job search & recommendation, etc.
- Strong user interest towards fresh content
- ML algorithms reinforcing recorded selection bias





Different views of item exposure in the RecSys '17 Challenge Data

Relation to other known biases Recency bias might be related to popularity (i.e. fresh content being popular), trust (i.e. users enticed by Related Items module), and position bias (depending on module layout).

Research Question

Can unbiased learning to rank be used to reduce recency bias?

Task: Job Recommendation

Job recommendation (Kenthapadi et al. 2017) is the task of recommending job advertisements (job ads) to potential candidates (users).

- Job ads have a shorter lifespan on the market - new positions are advertised on a daily basis, and can be filled within a few weeks
- Users prefer to apply for job ads as soon as they are discovered.

Data & Experimental Setup

- Experiments were conducted on the training set of RecSys Challenge 2017 dataset (Abel et al, 2017), using only user-item interactions
- Level of interactions: impression (0), click (1), bookmark (2), apply (3), delete (4), and recruiter action (5).
- Time-based splitting: train 14d, dev 7d, test 7d
- Cast as a reranking task by incorporating true impressions

Methodology

1. Model exposure with clickthrough rate to obtain a simplified propensity estimate (assuming $y = o_{uxt}\, r_{ux}$)
2. Formulate a new binary cross-entropy loss (to be used on neural collaborative filtering) using inverse propensity weighting

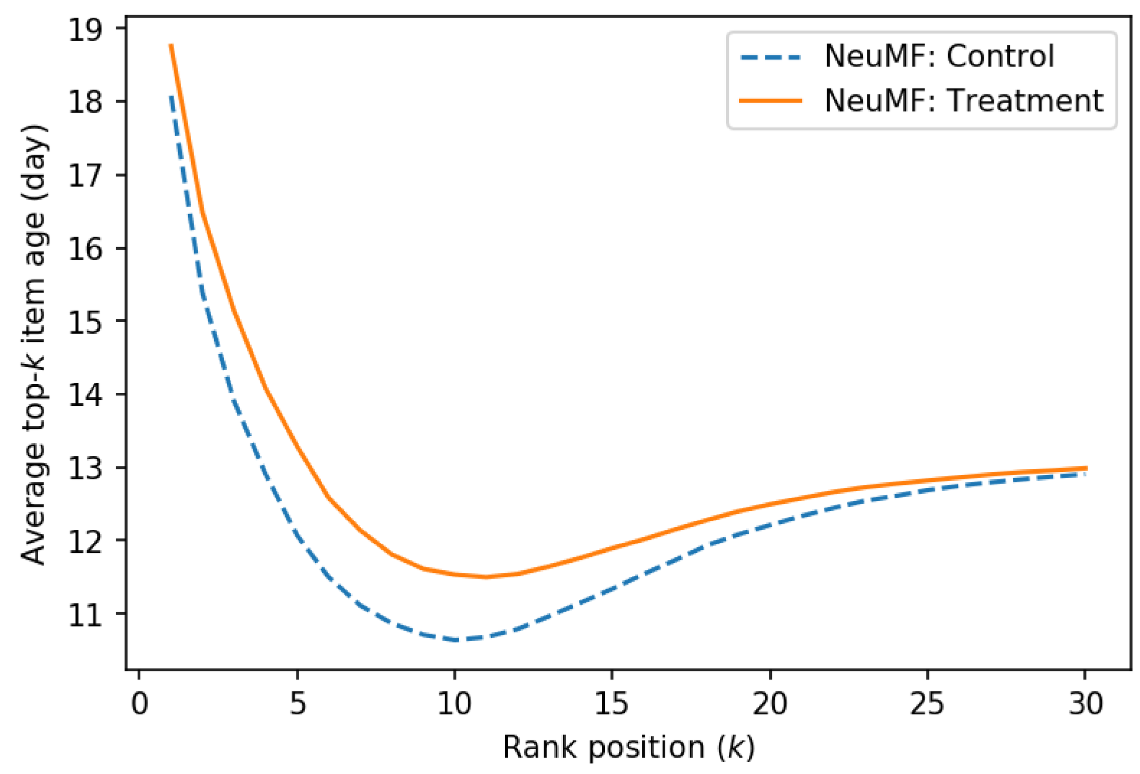
$$-\sum_{(x,y,t)\in\mathcal{D}} \left[\frac{y}{\hat{p}_{xt}} \log \sigma(f(x)) + \frac{1-y}{\hat{p}_{xt}} \log(1-\sigma(f(x)))\right]. \quad (1)$$

Here, $f(x)$ denotes the model score and $\hat{p}_{xt}$ is the recency propensity of item x at time t.

Debiasing Improves Recommendation Effectiveness

We used a tuned NeuMF (He et al., 2017) as the base model.

<u>**Result**</u>: Treatment model covered more slightly dated job ads (by roughly 1 days), which led to improved early NDCG and HitRate



	NDCG@5	HitRate@5
NeuMF: Control	0.5055	0.8903
NeuMF: Treatment	**0.5383‡ (+6.5%)**	**0.9027† (+1.4%)**