Popularity Bias in Dynamic Recommendation

Ziwei Zhu, Yun He, Xing Zhao, and James Caverlee
Department of Computer Science and Engineering, Texas A&M University, USA
{zhuziwei, yunhe, xingzhao, caverlee}@tamu.edu

KDD2021

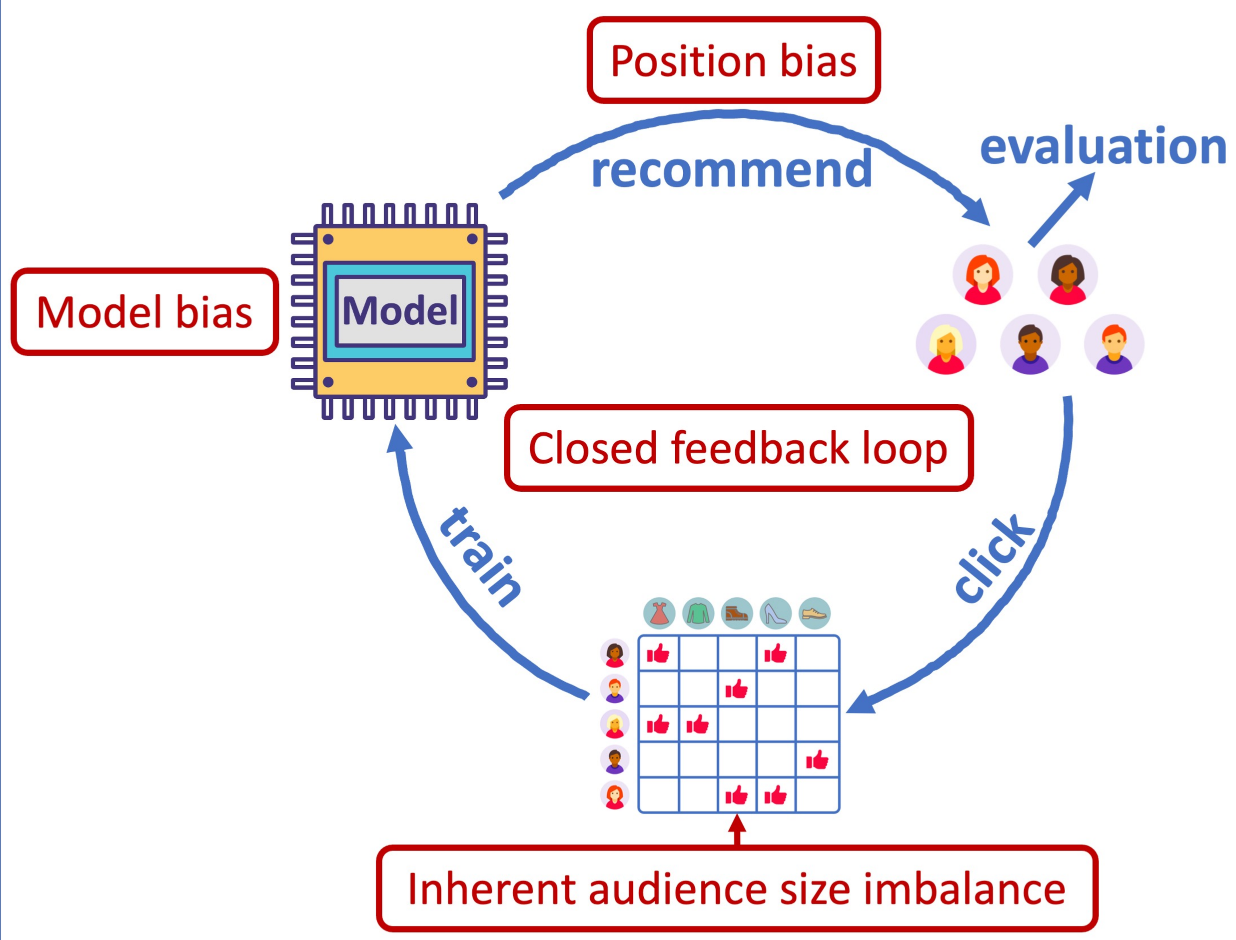
Introduction

Prior works study the long-standing problem of **popularity bias** in RecSys in a **static** setting, where bias is analyzed by conducting a single round of recommendation. There is a significant research gap in our understanding of the **dynamics** of the popularity bias in a real-world **dynamic recommendation** process.

Contributions:

- Conduct a comprehensive **empirical study** by simulation experiments to investigate how the popularity bias **evolves** in dynamic recommendation and how **four key bias factors** impact the bias;
- Proposed a simple but powerful **dynamic debiasing framework** to adapt exiting static debiasing methods to the dynamic scenario;
- Extensive experiments to show the **effectiveness** of the proposed dynamic debiasing method.

Dynamic Recommendation



- **Inherent Audience Size Imbalance**: A few items have very large audience sizes, while the majority have small ones.
- **Model Bias:** The model itself amplify any imbalances in the data it ingests for training.
- **Position Bias:** Once the model makes recommendations, the top-ranked items are more likely to be examined by users.
- **Closed Feedback Loop:** The feedback collected from recommendations by the current model will impact the training of future versions of the model.

Algorithm 1: Dynamic Recommendation Process

1. **Bootstrap:** Randomly show K items to each user and collect initial clicks $\mathcal{D}$ and train the first model ψ by $\mathcal{D}$;
2. **for** $t = 1 : T$ **do**
3. Recommend K items to the current user u_t by ψ;
4. Collect new clicks and add them to $\mathcal{D}$;
5. **if** $t\%L == 0$ **then**
6. Retrain ψ by $\mathcal{D}$;

Popularity Bias

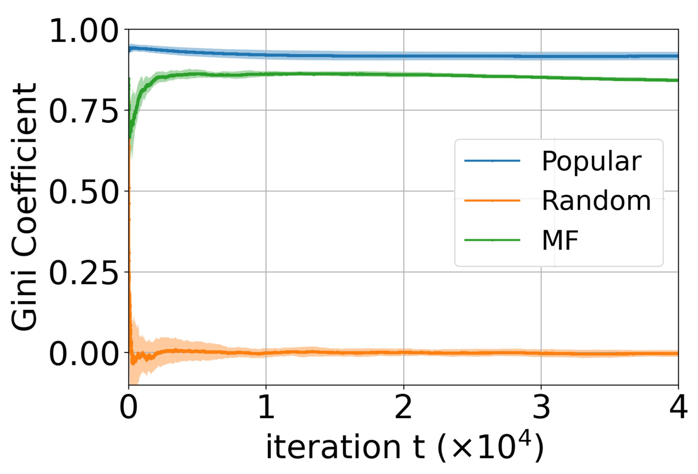
Compared with less popular items, whether popular items are more likely to be correctly recommended to matched users who like them?

⟶ Calculate **Gini Coefficient** of matched exposure (i.e., true positive rate) over items sorted by popularity. (**higher, more severe bias**)

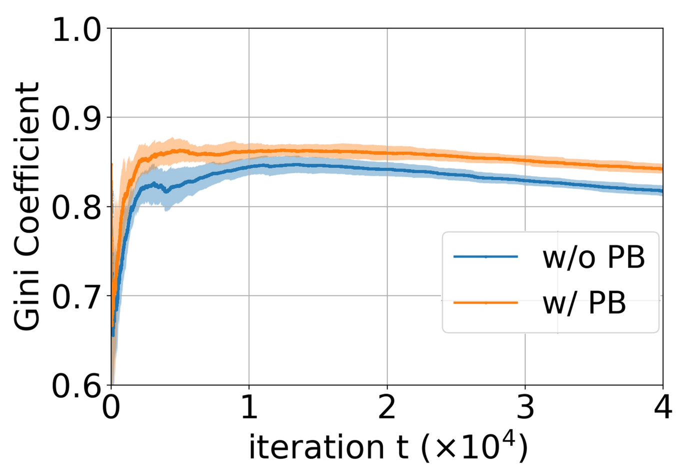
$$Gini_t = \frac{\sum_{i \in \mathcal{I}}(2i - M - 1)TPR_i}{M \sum_{i \in \mathcal{I}} TPR_i}$$

Empirical Study of Popularity Bias in Dynamic Recommendation

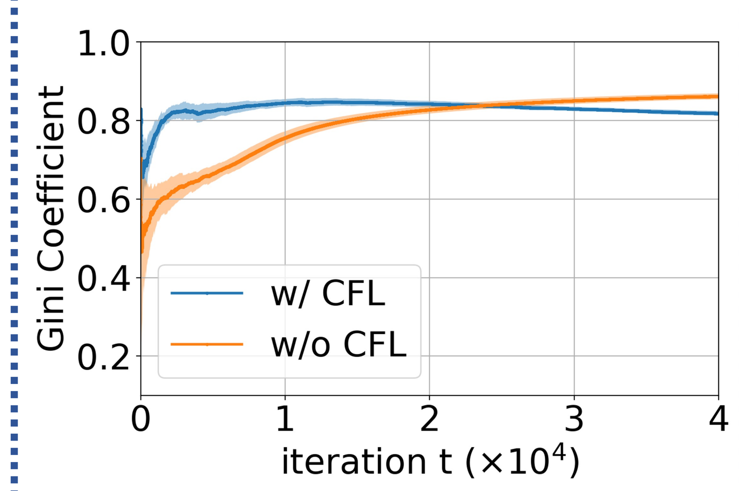
- Evolution of popularity bias:



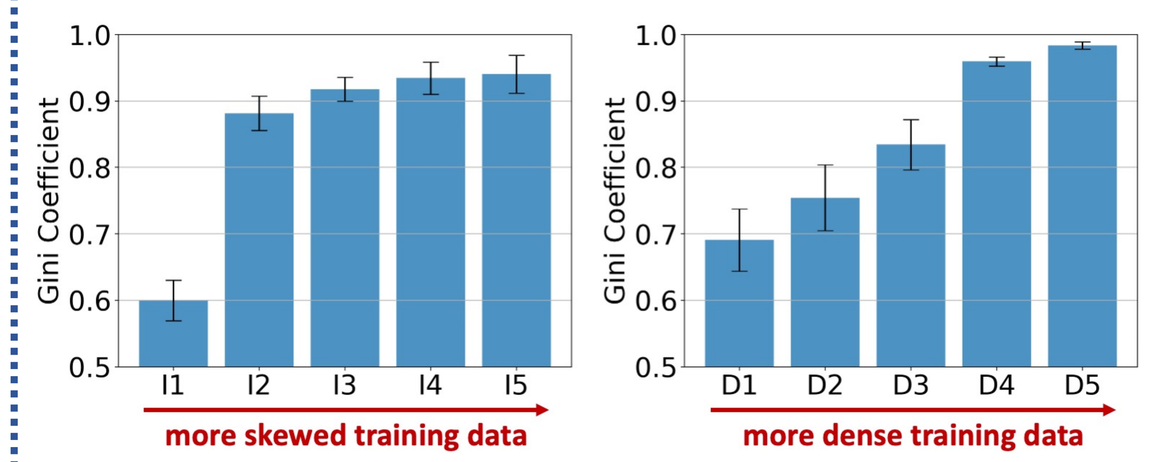
- Impact of position bias:



- Impact of closed feedback loop:



- Impact of model bias:



- **Inherent audience size imbalance** and **model bias** are the main sources of popularity bias;
- **Position bias** and **closed feedback loop** can intensify the bias when inherent audience size imbalance and model bias exist;
- Higher training data **density** and greater **imbalance** can increase the effect of model bias.

Debias in a dynamic way

- Adopt an **existing static debiasing method**, apply it to dynamic recommendation process by gradually **increasing debiasing strength**.

  Example: an existing debiasing method **Scale**

  Debiasing strength hyper-parameter

$$\widehat{r}_{u,i}^{(scaled)} = \widehat{r}_{u,i}^{(model)} / (C_i)^{\boxed{\alpha}}$$

  popularity of item i

- Adopt an **existing static debiasing method**, apply it to dynamic recommendation process by gradually **increasing debiasing strength**.

$$P(r_{u,i} = 1 | c_{k_1} = 0, \dots, c_{k_F} = 0) = 1 - \frac{1 - \boxed{\theta_{u,i}}}{\prod_{f=1}^{F}(1 - \boxed{\delta_{k_f}}\theta_{u,i})}$$

prediction from a recommendation model

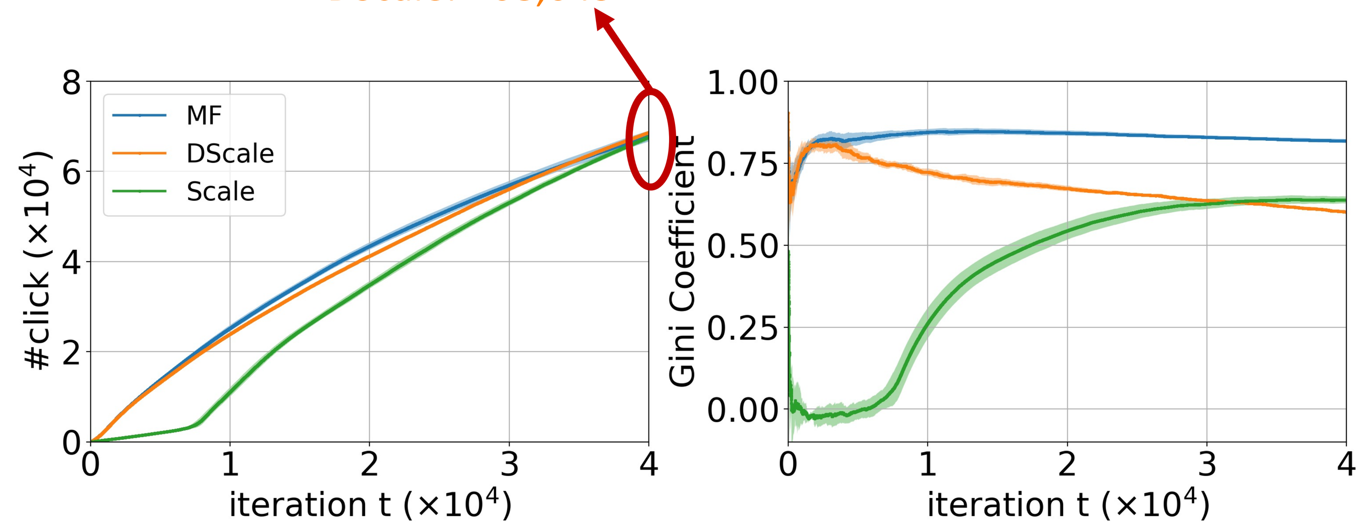
The probability user u likes item i given i has been recommended to u for F times and did not receive any clicks.

Examine probability at position k_f

Experimental Results

- With increasing debiasing strength, we can **continuously decrease** the bias
- Fix the debiasing strength as static debiasing method, the bias starts low but **grows to high level**.

MF:	67,816
Scale:	66,630
DScale:	68,645



- Integrate DScale and false positive correction, the popularity **bias is further decreased**;
- **More clicks** are collected by debiasing by the proposed method (higher recommendation utility is achieved).

MF:	67,816
DScale:	68,645
FPC-DScale:	73,145