Combining Reward & Rank Signals for Slate Recommendation

Imad Aouali, Sergey Ivanov, Mike Gartrell, David Rohde, Flavian Vasile,Victor Zaytsev, and Diego Legrand

CRITEO
AI Lab

OPEN TOGETHER IMPACTFUL

CRITEO

Combining Reward and Rank Signals for Slate Recommendation

Introduction

- **Slate recommendation:** Task of recommending a collection of K items at once to the user.
- **Two pieces of feedback:**
  - Was the slate clicked? (*reward* signal)
  - If it was clicked, which of its items was clicked? (*rank* signal)
- **Example: 3 products**, *phone, couscous and beer.*

Slate	non-clicks	clicks on 1	clicks on 2
phone, couscous	661	10	29
phone, beer	644	9	47
couscous, beer	626	46	28

Combining Reward and Rank Signals for Slate Recommendation

Reward Model - Intuition

- **Original Data:**

Slate	non-clicks	clicks on 1	clicks on 2
phone, couscous	661	10	29
phone, beer	644	9	47
couscous, beer	626	46	28

- **Reward Model:** Ignores items ranking and only uses the reward signal.

Slate	non-clicks	clicks
phone, couscous	661	39
phone, beer	644	56
couscous, beer	626	74

Combining Reward and Rank Signals for Slate Recommendation

Rank Model - Intuition

- **Original Data:**

Slate	non-clicks	clicks on 1	clicks on 2
phone, couscous	661	10	29
phone, beer	644	9	47
couscous, beer	626	46	28

- **Rank Model:** Ignores the reward signal and only uses items ranking.

Slate	clicks on 1	clicks on 2
phone, couscous	10	29
phone, beer	9	47
couscous, beer	46	28

Combining Reward and Rank Signals for Slate Recommendation

Full Model - Intuition

- **Full Model:** Uses both reward and rank signals. It takes data in its raw form.

Slate	non-clicks	clicks on 1	clicks on 2
phone, couscous	661	10	29
phone, beer	644	9	47
couscous, beer	626	46	28

- It is intuitive that both the reward and the rank signals shall contain important information of user preferences.

Combining Reward and Rank Signals for Slate Recommendation

Bayesian Models: Formulation

- **Models formulation in the case of slates of size 2:**

$$\phi \sim \Gamma(1,1) \quad \theta_i \sim \Gamma(1,1), i \in [N]$$

Model	Description
Full	$nc, c_1, c_2 \mid I, \phi, \theta, a_1, a_2 \sim \text{Multinomial}\left(I, \frac{\phi}{\phi+\theta_{a_1}+\theta_{a_2}}, \frac{\theta_{a_1}}{\phi+\theta_{a_1}+\theta_{a_2}}, \frac{\theta_{a_2}}{\phi+\theta_{a_1}+\theta_{a_2}}\right)$
Reward	$nc, c \mid I, \phi, \theta, a_1, a_2 \sim \text{Multinomial}\left(I, \frac{\phi}{\phi+\theta_{a_1}+\theta_{a_2}}, \frac{\theta_{a_1}+\theta_{a_2}}{\phi+\theta_{a_1}+\theta_{a_2}}\right)$
Rank	$c_1, c_2 \mid I_c, \theta, a_1, a_2 \sim \text{Multinomial}\left(I_c, \frac{\theta_{a_1}}{\theta_{a_1}+\theta_{a_2}}, \frac{\theta_{a_2}}{\theta_{a_1}+\theta_{a_2}}\right)$

- **Remark:** These models can be extended smoothly to arbitrary slate sizes.

Combining Reward and Rank Signals for Slate Recommendation

Experiments

- **Synthetic data:** We generate n samples of users interaction with each slate, using a multinomial distribution with known parameters **Φ** and **θ**.
- **Evaluation:** Models ability to estimate the probability of a click on the first item in the slates

$$L_1(p_{\hat{\theta}}, p_\theta) = \sum_{\text{all slates } a} \left| \frac{\hat{\theta}_{a_1}}{\sum_{j \in [K]} \hat{\theta}_{a_j}} - \frac{\theta_{a_1}}{\sum_{j \in [K]} \theta_{a_j}} \right|$$

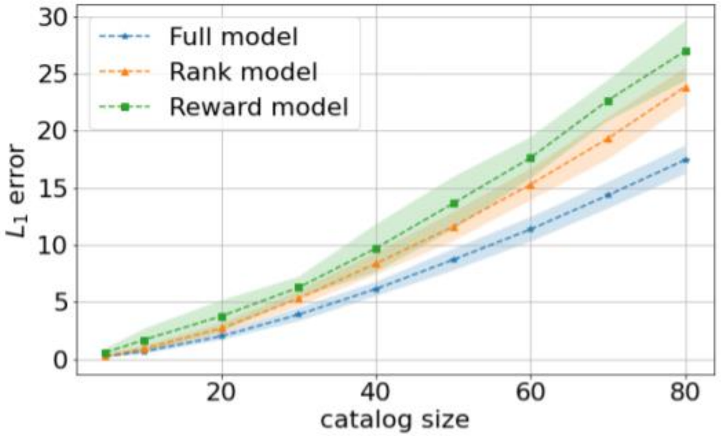
- **Additional Experiment:** Models ability to estimate the probability of a non-click on the slates.

$$L_1(\hat{q}, q) = \sum_{\text{all slates } a} \left| \frac{\hat{\phi}}{\hat{\phi} + \sum_{j \in [K]} \hat{\theta}_{a_j}} - \frac{\phi}{\phi + \sum_{j \in [K]} \theta_{a_j}} \right|$$

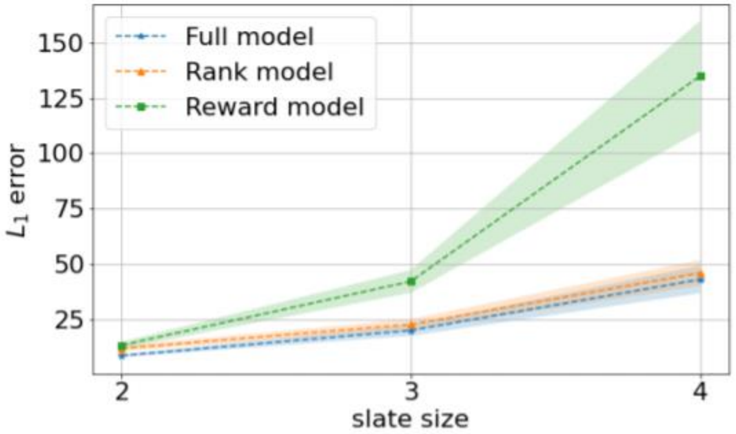
Combining Reward and Rank Signals for Slate Recommendation

Experiments

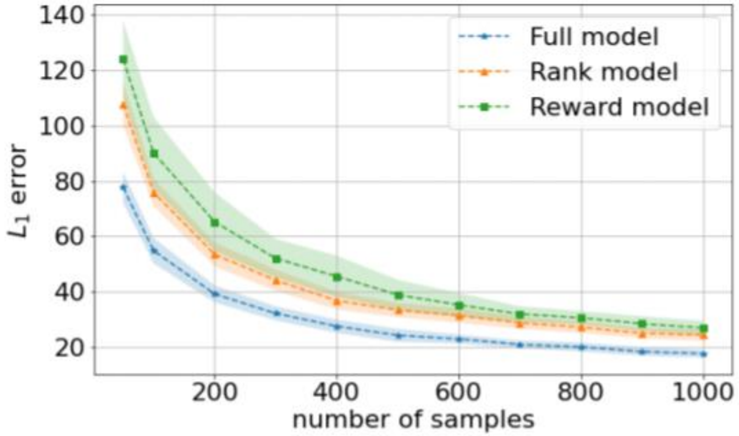
- **Results**



(a) Varying catalog size.



(b) Varying slate size.



(c) Varying N° of samples.

Combining Reward and Rank Signals for Slate Recommendation

Conclusion

- Three Bayesian models for non-personalized slate recommendation.
- Combining both information is beneficial in slate recommendation, especially as the catalog size and slate size grow.
- **Future Work:** Extend the Full model to personalized slate recommendation.