

Social Collaborative Filtering for Cold-start Recommendations

Suvash Sedhain^{1,2}, Scott Sanner^{2,1}, Darius Braziunas³, Lexing Xie^{1,2}, Jordan Christensen³
 ANU¹, NICTA², Kobo Inc³

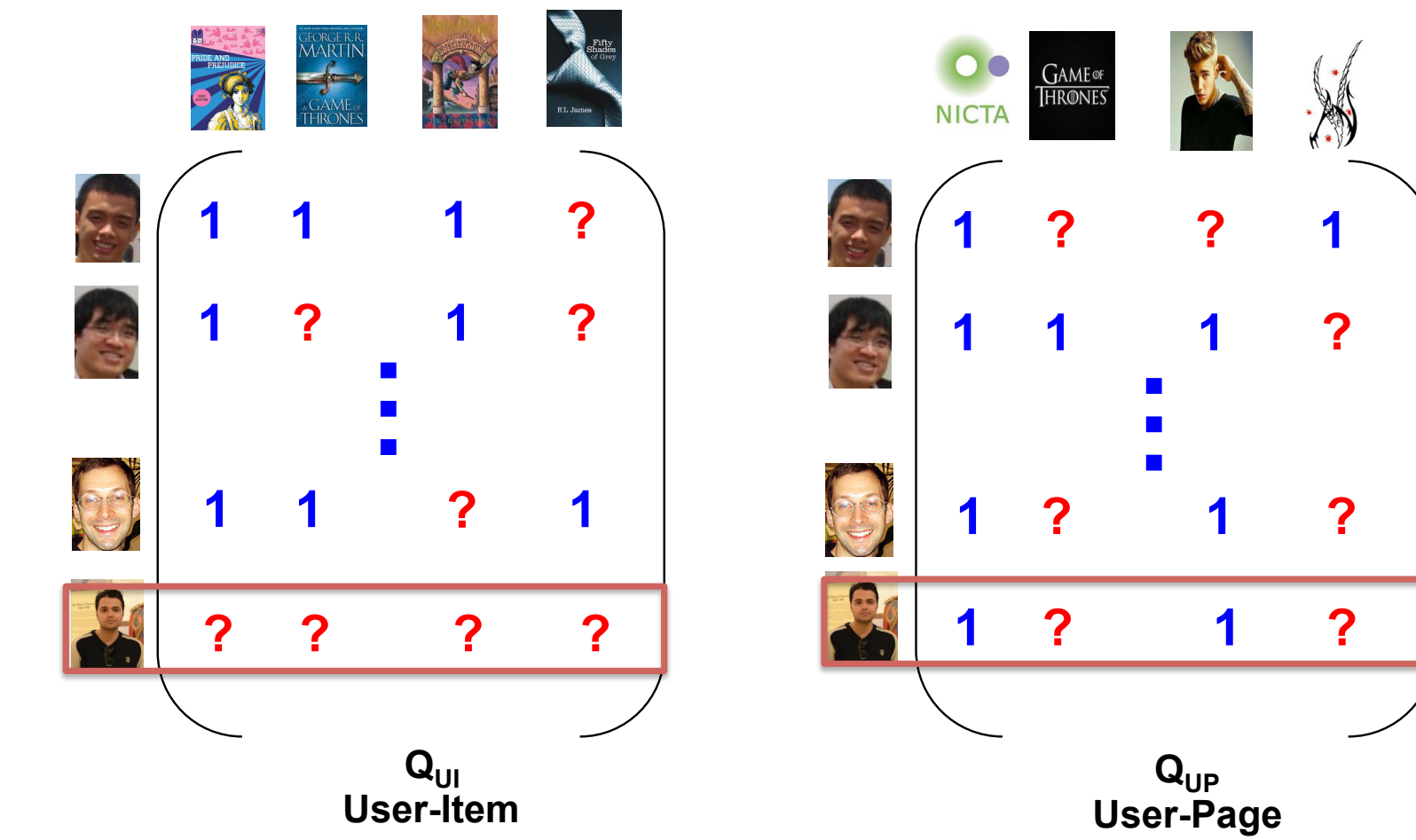
Highlights

- Question:** How to incorporate users' social information in an implicit (positive only) feedback recommendation setting? ARE YOU WORRIED ABOUT COLD START USERS?
- Results:**
 - We cast neighborhood-based recommendation in a matrix view and showed how this view allows us to directly incorporate users' social side information for cold-start recommendations.
 - We showed that Facebook page likes are an extremely valuable source of this user side information.

NO WORRIES, SOCIAL SIDE INFORMATION FROM FACEBOOK CAN HELP.



Social Cold-start Recommendation



$$U_{\text{Target}} \begin{bmatrix} r \\ Q_{Up} \end{bmatrix} \star MM_2 \begin{bmatrix} r \\ U_{\text{Train}} \end{bmatrix} \star MM_1 \begin{bmatrix} I \\ c \end{bmatrix} = U_{\text{Target}} \begin{bmatrix} I \\ R_{Ui} \end{bmatrix}$$

$MM_* = \{IP, \text{LogIP}, \text{BinIP}, \text{Cos}\}$

$$IP(r, c) = \langle r, c \rangle$$

$$\text{LogIP}(r, c) = \log \langle r, c \rangle$$

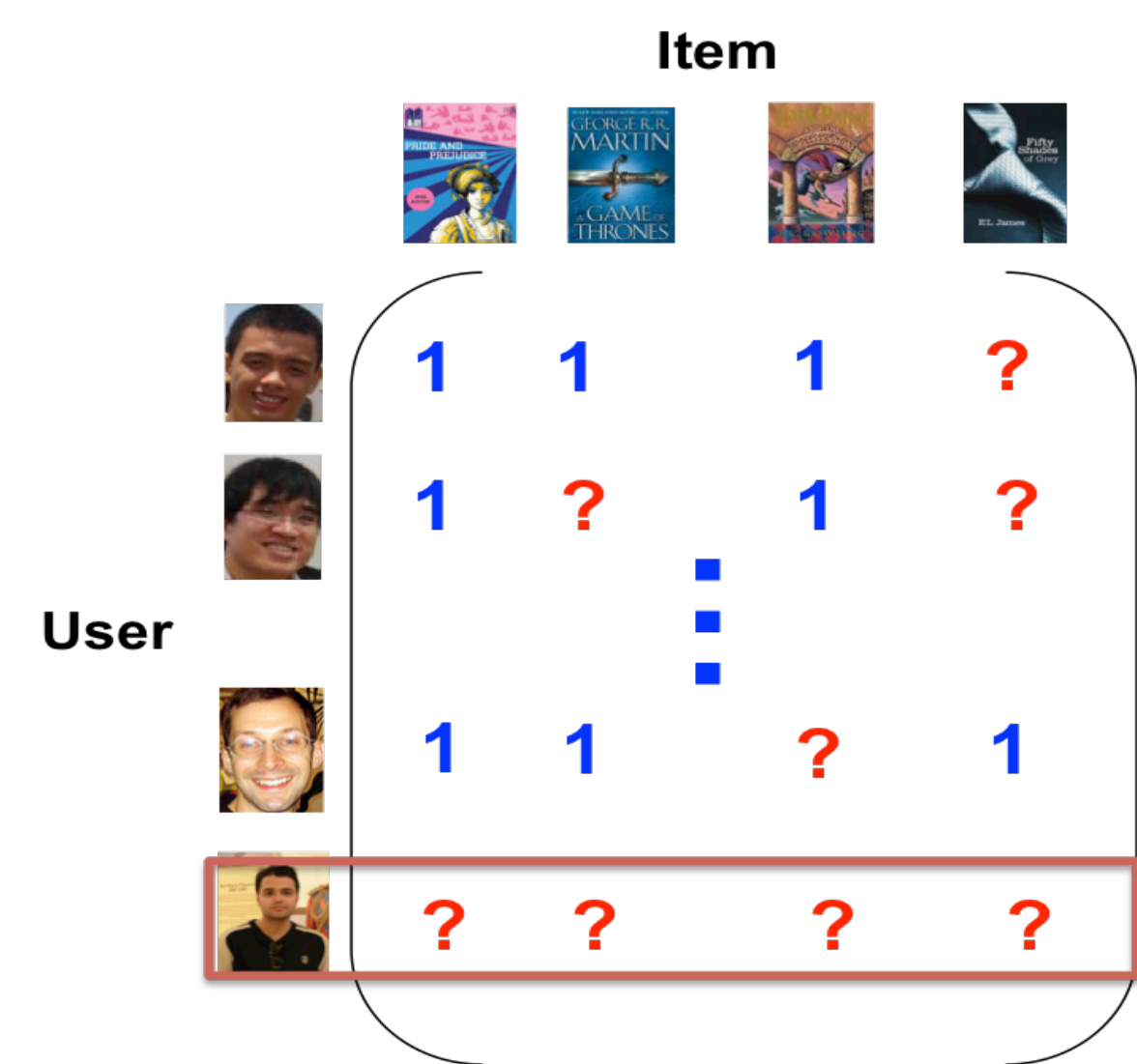
$$\text{BinIP}(r, c) = \begin{cases} 1 & \text{if } \langle r, c \rangle > \tau \\ 0 & \text{otherwise} \end{cases}$$

$$\text{Cos}(r, c) = \frac{\langle r, c \rangle}{\|r\| \|c\|}$$

WELL, I THINK I CAN DEAL WITH COLD-START USERS IF THEY ARE SOCIAL



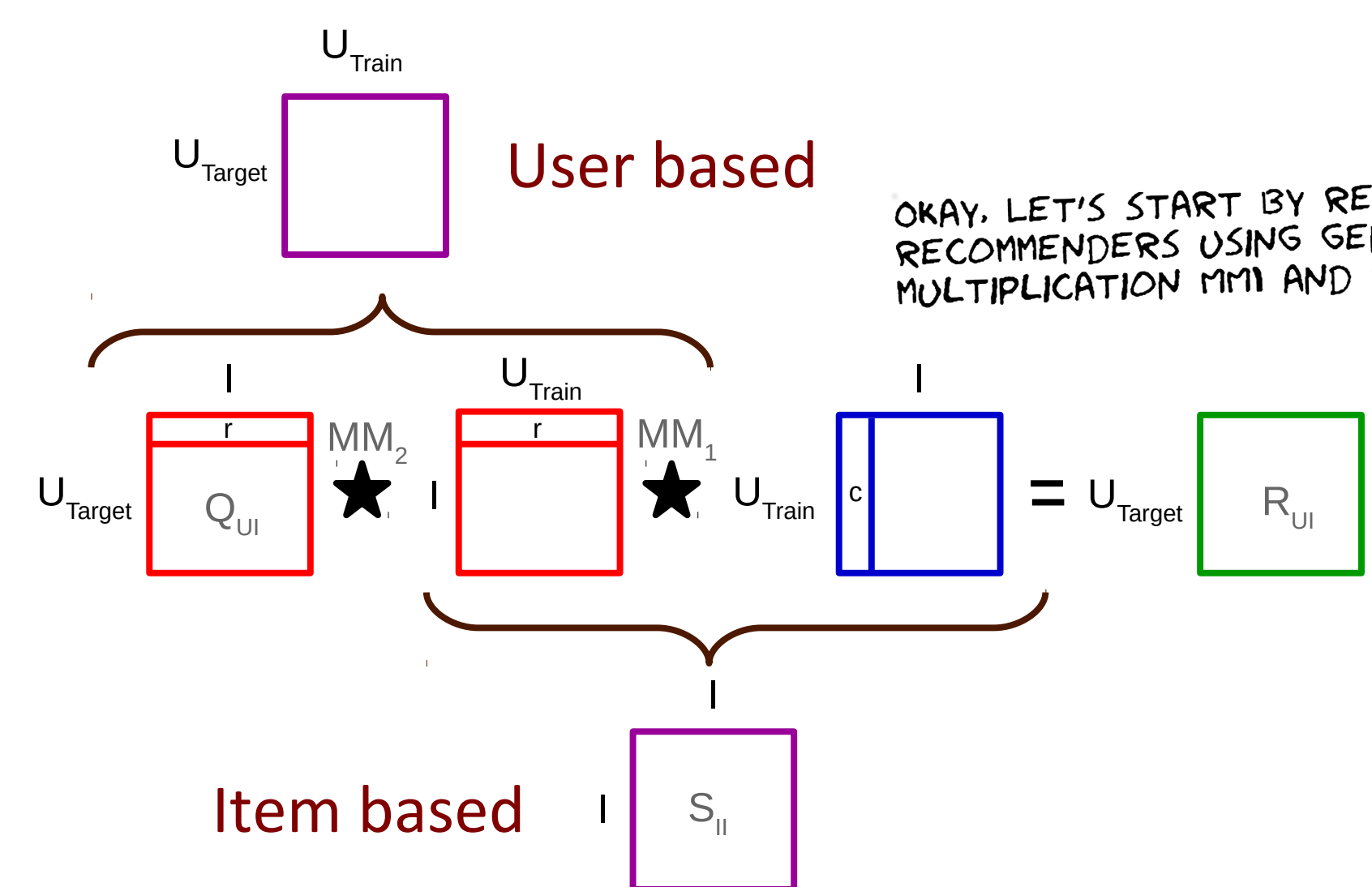
Cold-start Recommendation



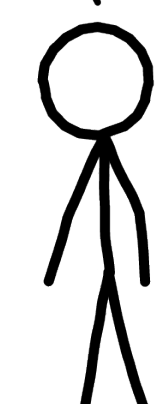
I KNOW NOTHING ABOUT THIS GUY!! HOW SHALL I RECOMMEND STUFF TO HIM? I HATE COLD-START USERS!!



Neighborhood-based Recommendation



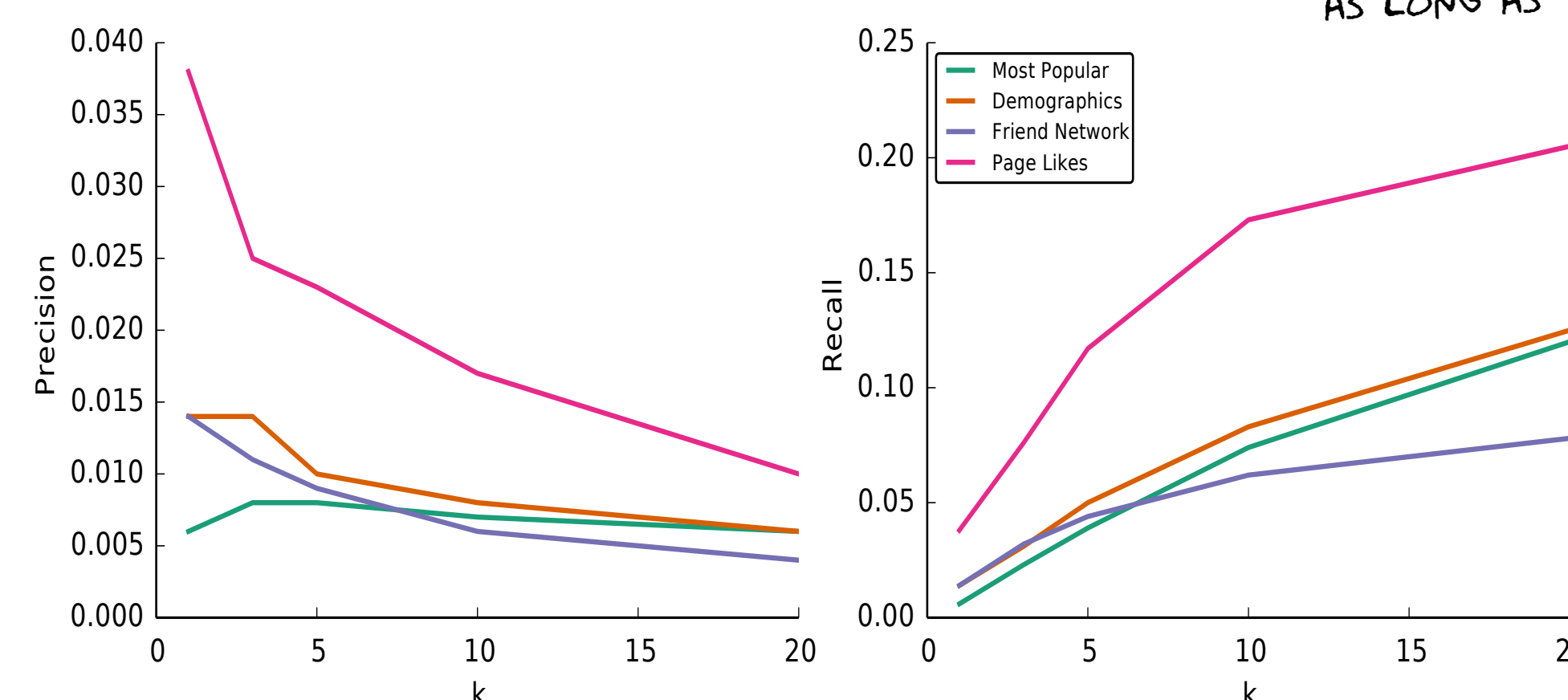
OKAY, LET'S START BY REWRITING NEIGHBORHOOD RECOMMENDERS USING GENERALIZED MATRIX MULTIPLICATION MM1 AND MM2



Evaluation and Analysis

Data Description	
# users	32,027
# items	88,810
# pages	6,218,698
# pages liked by >5 and <5000 users	606,780

Social Cold start recommendation vs. baselines



NOW, I DON'T MIND NEW USER'S AS LONG AS THEY ARE SOCIAL. HIGH FINE!!

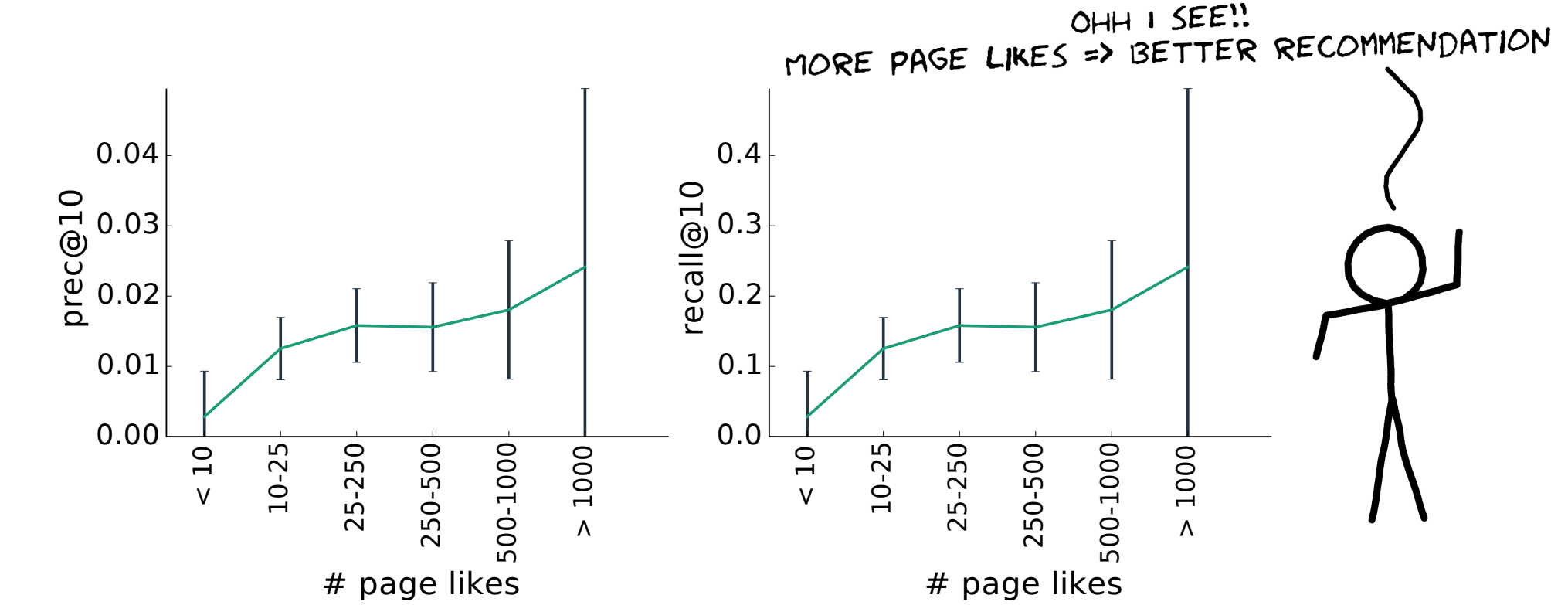


How does the performance vary with the choice of MM₁–MM₂?

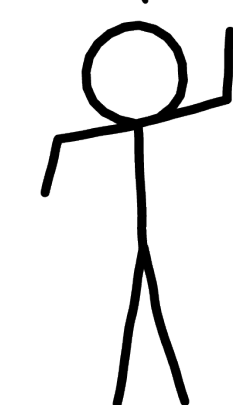
k	IP-IP		LogIP-IP		BinIP-IP		Cos-IP	
	Precision	Recall	Precision	Recall	Precision	Recall	Precision	Recall
@1	0.027+/-0.013	0.027+/-0.013	0.030+/-0.012	0.030+/-0.012	0.028+/-0.012	0.028+/-0.012	0.030+/-0.012	0.030+/-0.012
@3	0.018+/-0.007	0.053+/-0.020	0.020+/-0.008	0.061+/-0.023	0.022+/-0.009	0.066+/-0.026	0.022+/-0.008	0.066+/-0.024
@5	0.017+/-0.006	0.086+/-0.032	0.019+/-0.008	0.097+/-0.038	0.020+/-0.008	0.102+/-0.040	0.020+/-0.008	0.102+/-0.038
@10	0.014+/-0.005	0.136+/-0.047	0.015+/-0.005	0.148+/-0.050	0.015+/-0.005	0.153+/-0.051	0.015+/-0.005	0.154+/-0.052
@20	0.009+/-0.003	0.181+/-0.060	0.009+/-0.003	0.185+/-0.061	0.009+/-0.003	0.186+/-0.060	0.010+/-0.003	0.191+/-0.060
mAP	0.057+/-0.019		0.062+/-0.021		0.063+/-0.022		0.065+/-0.021	

k	IP-Cos		LogIP-Cos		BinIP-Cos		Cos-Cos	
	Precision	Recall	Precision	Recall	Precision	Recall	Precision	Recall
@1	0.031+/-0.014	0.031+/-0.014	0.031+/-0.013	0.031+/-0.013	0.030+/-0.014	0.030+/-0.014	0.038+/-0.015	0.038+/-0.015
@3	0.021+/-0.008	0.064+/-0.023	0.022+/-0.008	0.065+/-0.023	0.021+/-0.008	0.063+/-0.024	0.025+/-0.010	0.076+/-0.029
@5	0.020+/-0.008	0.100+/-0.039	0.020+/-0.008	0.101+/-0.038	0.020+/-0.008	0.100+/-0.037	0.023+/-0.009	0.117+/-0.044
@10	0.015+/-0.005	0.151+/-0.052	0.015+/-0.005	0.153+/-0.052	0.016+/-0.005	0.159+/-0.055	0.017+/-0.006	0.173+/-0.059
@20	0.010+/-0.003	0.193+/-0.061	0.010+/-0.003	0.193+/-0.061	0.010+/-0.003	0.202+/-0.063	0.010+/-0.003	0.205+/-0.063
mAP	0.065+/-0.022		0.065+/-0.022		0.065+/-0.021		0.075+/-0.025	

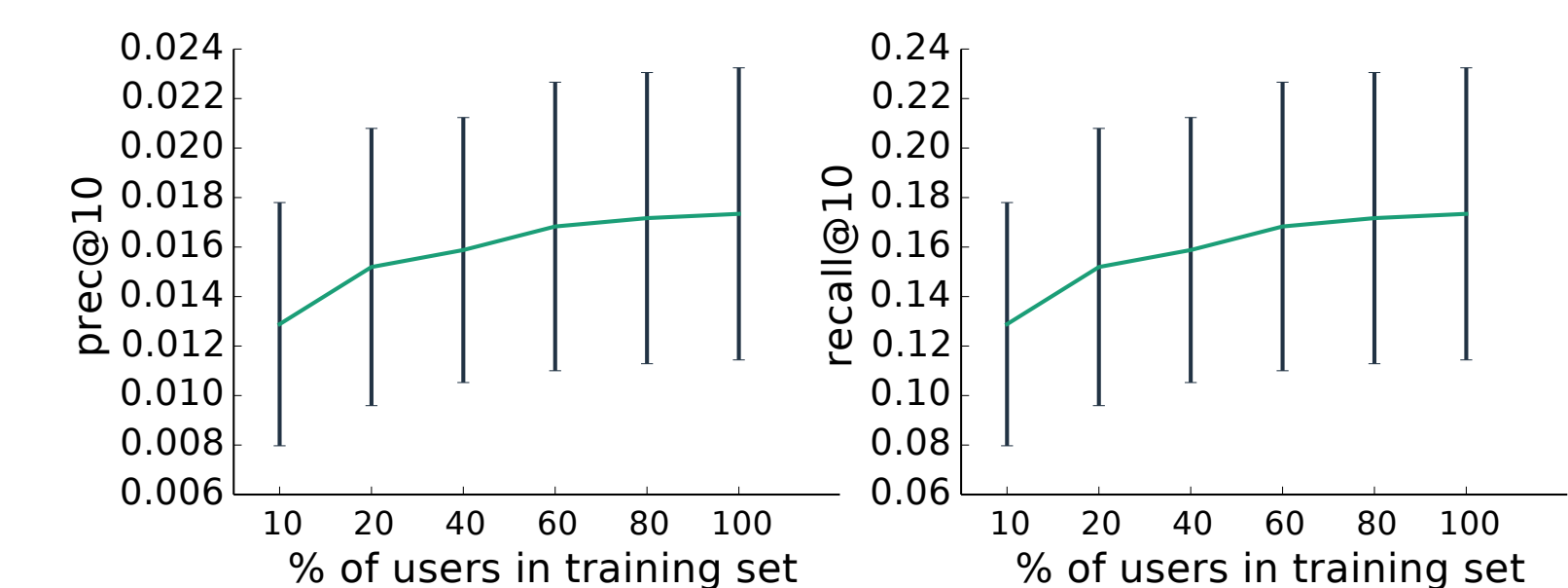
How does the performance vary with the number of page likes for target user?



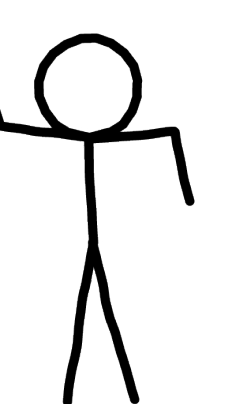
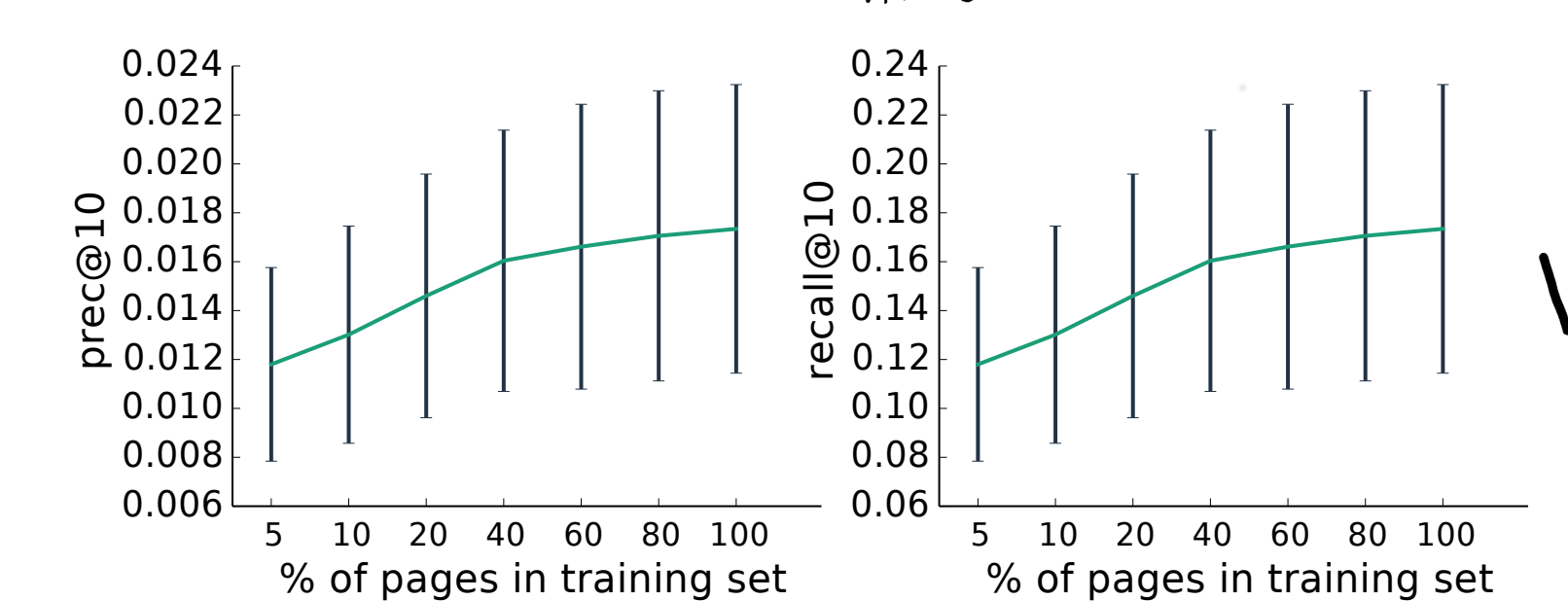
OH! I SEE!! MORE PAGE LIKES => BETTER RECOMMENDATION



How does the performance vary with the number of users/pages in training set?



DIMINISHING MARGINAL RETURN AFTER 40%-60% OF THE USERS/PAGES ARE INCLUDED IN TRAINING.



Conclusions & Future Work

- We showed that social cold-start recommendation using Facebook page likes provides up to 600% improvement over various baselines.
- However, our proposed framework does not jointly leverage both user and item side information, which we plan to explore in future work.