

Towards Conversational Recommender Systems



Konstantina Christakopoulou



Filip Radlinski



Katja Hofmann

There are too many options
for where to eat dinner tonight...

If you asked a local:


Do you
like
seafood?

Do you
have a
car?






TOWARDS THIS EXAMPLE SCENARIO




Sure! How hungry are you?

Not very hungry

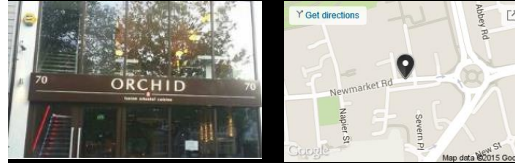


Do you prefer Chinese or Indian today?

Chinese.



Ok. Would you like to go to Orchid?



Yes, thanks!

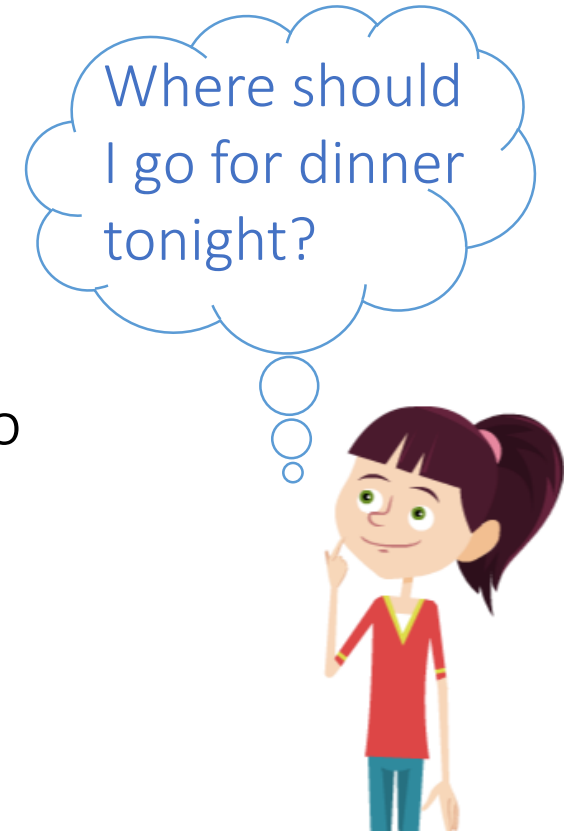


ROADMAP

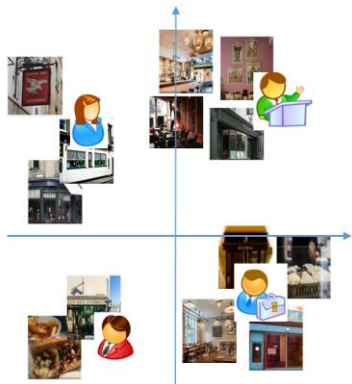
Model exploits implicit structure among items and users to efficiently propagate feedback

Explore/exploit strategy to probe the space of items and allow continuous learning

Feedback elicitation mechanism to select **absolute** and **relative** questions.

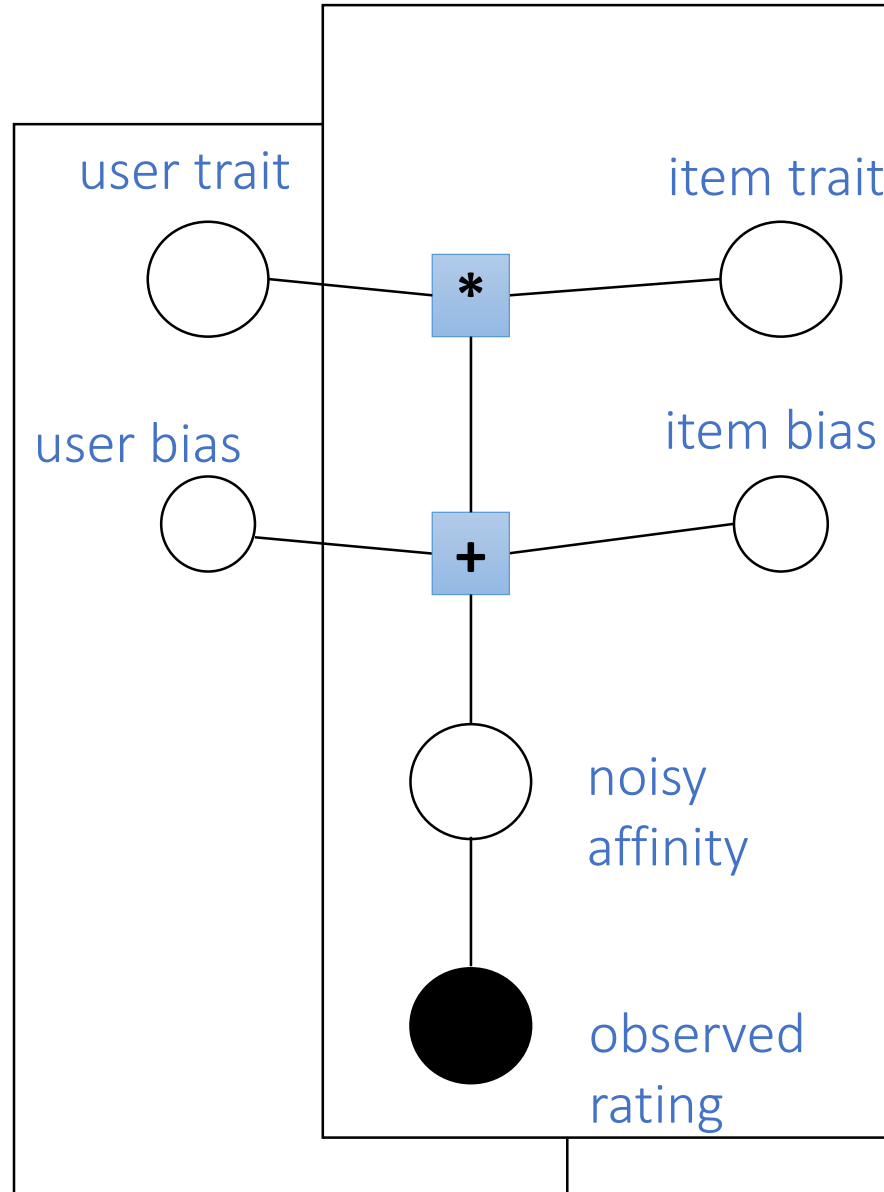


MODEL: PROBABILISTIC MATRIX FACTORIZATION



$$\mathbf{u}_i \sim \mathcal{N}(\mathbf{0}, \sigma_1^2 \mathbf{I})$$

$$\alpha_i \sim \mathcal{N}(0, \sigma_2^2)$$



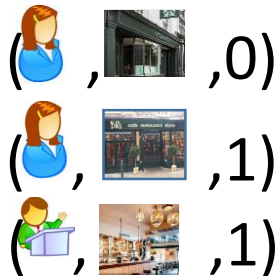
$$\mathbf{v}_j \sim \mathcal{N}(\mathbf{0}, \sigma_1^2 \mathbf{I})$$

$$\beta_j \sim \mathcal{N}(0, \sigma_2^2)$$

$$\mathcal{N}(y_{ij}, \epsilon_{ij})$$

$$y_{ij} = \alpha_i + \beta_j + \mathbf{u}_i^T \mathbf{v}_j$$

$$\hat{r}_{ij} = \mathbf{1}[\hat{y}_{ij} > 0]$$





INITIALIZATION FROM OFFLINE DATA



Learn offline embedding from logged observations



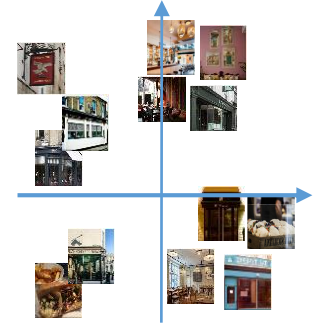
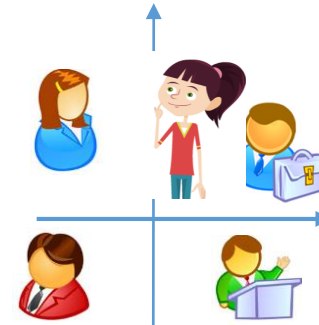
1. Initialize prior of **every item j** from corresponding trait posterior \mathbf{v}_j and bias β_j



2. For **cold-start** user

$$\mathbf{u}^{cold} \sim \mathbf{E}_{i=1, \dots, M} [\mathbf{u}_i]$$

$$\alpha^{cold} \sim \mathbf{E}_{i=1, \dots, M} [\alpha_i]$$





QUESTION SELECTION STRATEGIES

- Thompson Sampling (TS):

BANDIT LEARNING

Pick item with **max. sampled** noisy affinity

- Upper Confidence (UCB):

Pick item with **highest mean plus variance**

- Max. Variance:


ACTIVE LEARNING



Explore-only, **variance reduction**.

Pick item with highest noisy affinity variance 

- Max. Item Trait:

Pick item whose trait vector contains the **most information**, (i.e., highest L2 norm) 

- Greedy: **Exploit-only** strategy 

- Random: **Explore-only** strategy 

- Min. Item Trait: Baseline, least carrying information 

ONLINE UPDATING



Do you like

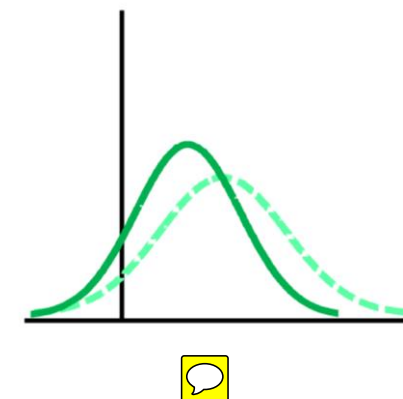
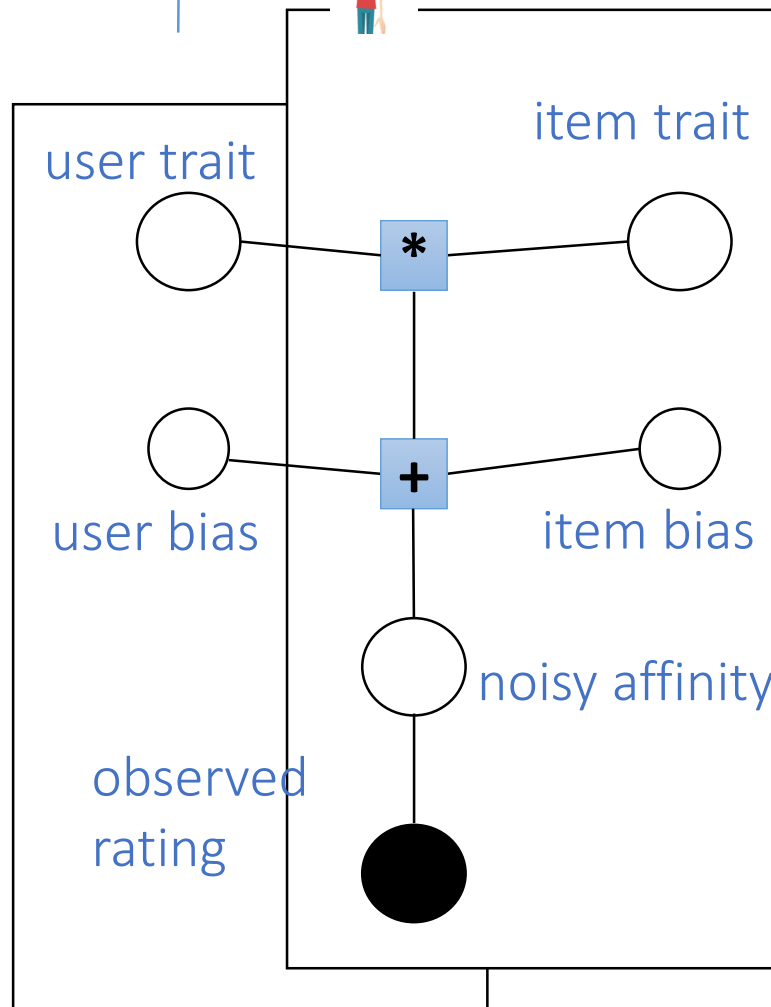
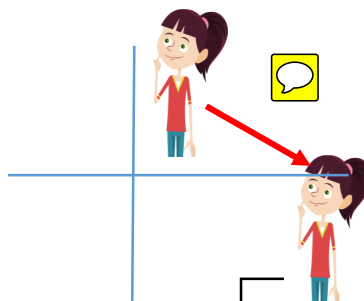


?

No



(, , 0)



ABSOLUTE MODEL, **ABSOLUTE** QUESTIONS

Do you like



?



Generalized Thompson
Sampling



$$j^* = \arg \max_{j \in \mathcal{J}} \hat{y}_{ij}$$



FEEDBACK ELICITATION FOR RELATIVE QUESTIONS

People are often better at giving [pairwise comparisons](#) instead of absolute judgements



VS



Yes/ No/ I
like
neither



ABSOLUTE MODEL, RELATIVE QUESTIONS



Insight : Restaurants compared should be far apart in the latent embedding

Do you prefer






over



?



1. Virtual observation ( ,  , 0)
2. Virtual prior = posterior after incorporating virtual obs. 
3. Pick item B: $j^* = \arg \max_{j \in \mathcal{J}} \hat{y}_{ij}$



Abs

$$j^* = \arg \max_{j \in \mathcal{J}} \hat{y}_{ij}$$

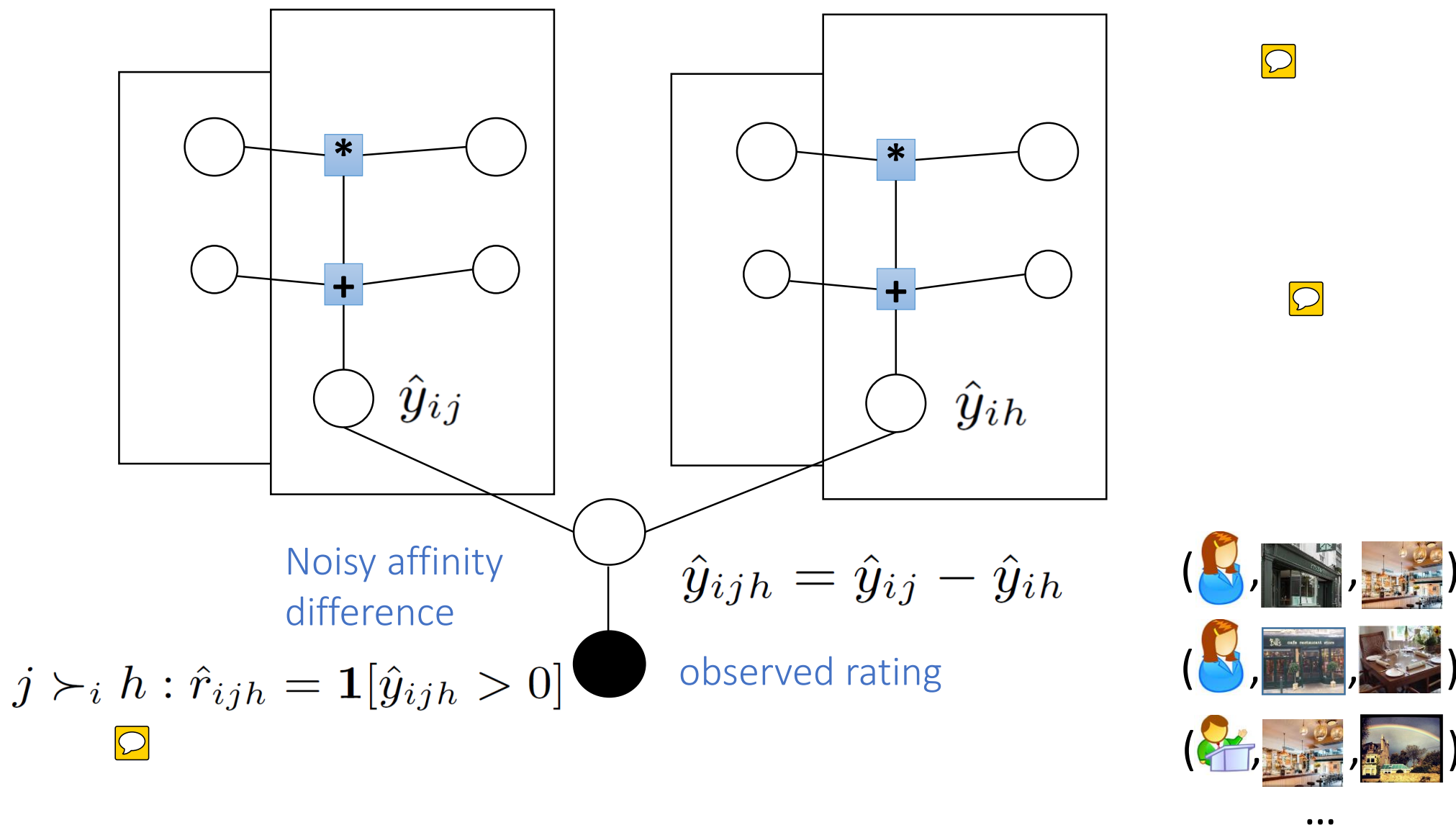


Abs Pos

Abs Pos & Neg



PAIRWISE MODEL



PAIRWISE MODEL, RELATIVE QUESTIONS

Do you
prefer



over



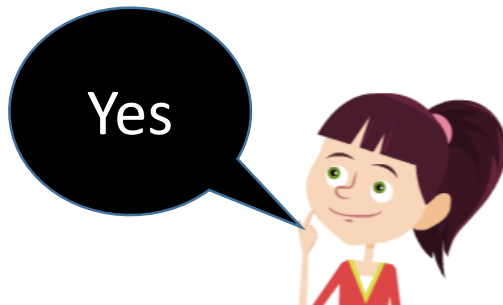
?

Abs $j^* = \arg \max_{j \in \mathcal{J}} \hat{y}_{ij}$



Pick item B = $h^* = \arg \max_{h \in \mathcal{J}} \hat{y}_{ihj^*}$
most preferred compared to A

Yes



EXPERIMENTS

EXPERIMENTAL SETUP

Offline phase: M users interact with N items → get offline embedding

Online phase: model interacts with cold-start users, asking questions on the N items

Domain: restaurant recommendation

$$AP@k = \sum_{\ell=0}^{k-1} \frac{P@{\ell} \cdot r_{i[\ell]}^{\text{true}}}{\min(k, \# \text{ of liked items})}$$

Search data
for offline
embedding

- 26/12/14 – 26/04/15

search logs: 3,549 cookies

with 289 restaurants →

9330 positive observations

- Sample negative

observations

User Study
as basis for
online
evaluation

“Would you
consider restaurant
X for your next
Friday night
dinner”?

28 participants

10 restaurants

Obtaining
Ground
Truth

1

Sample user
= one of 28
participants

2

Observe
user's ratings
on the 10
restaurants

3

Infer user
traits u_i

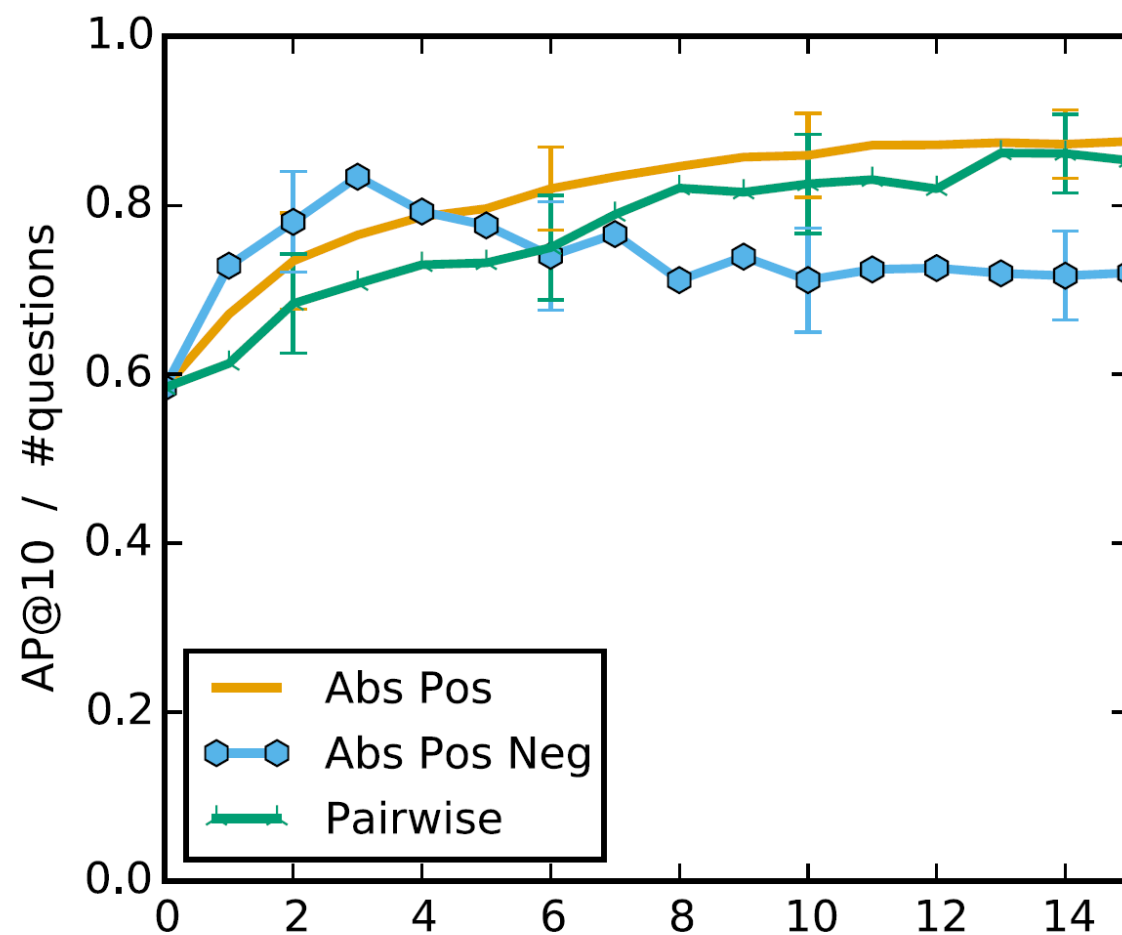
4

Set user prior
= sampled
value from u_i

5

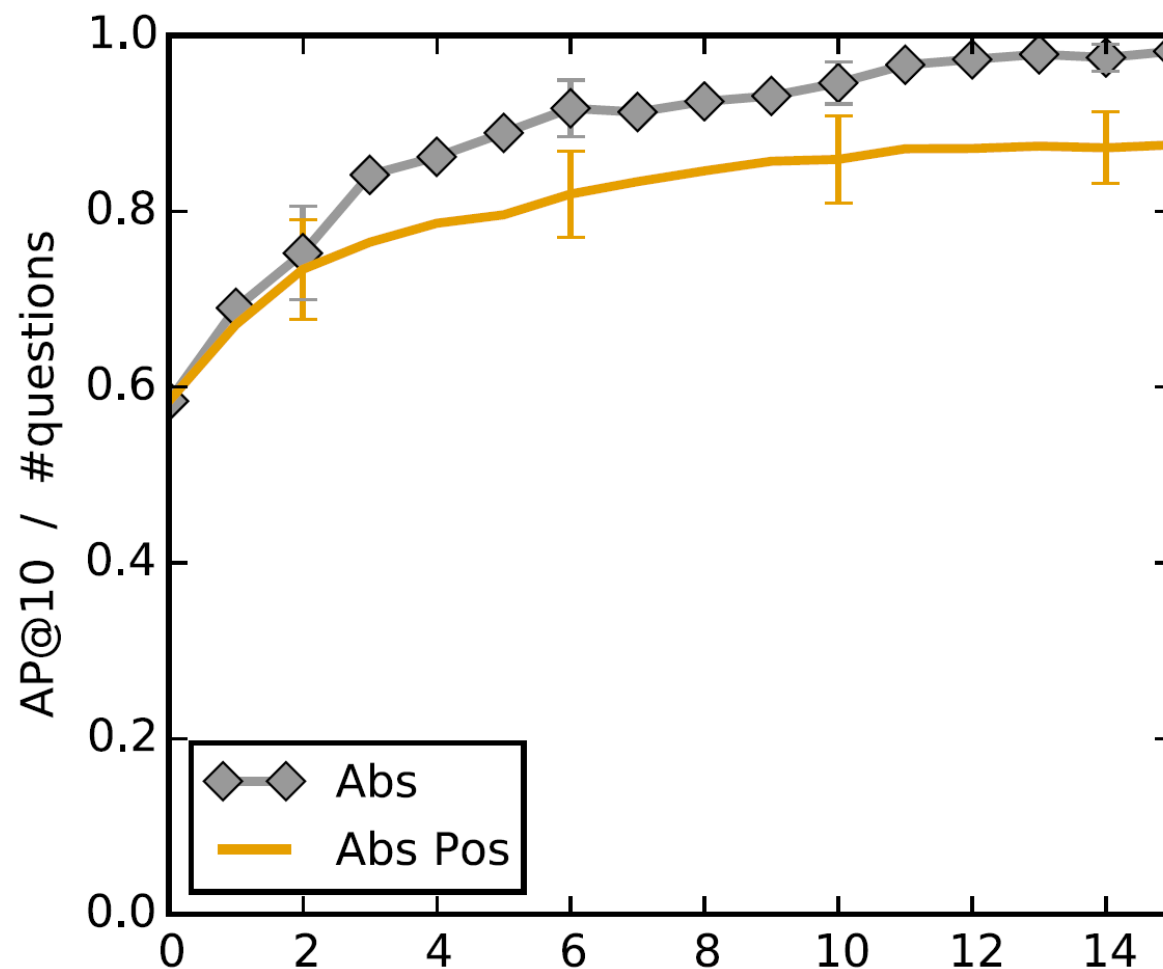
Infer &
sample
ratings r_i

Which method for relative questions is better?

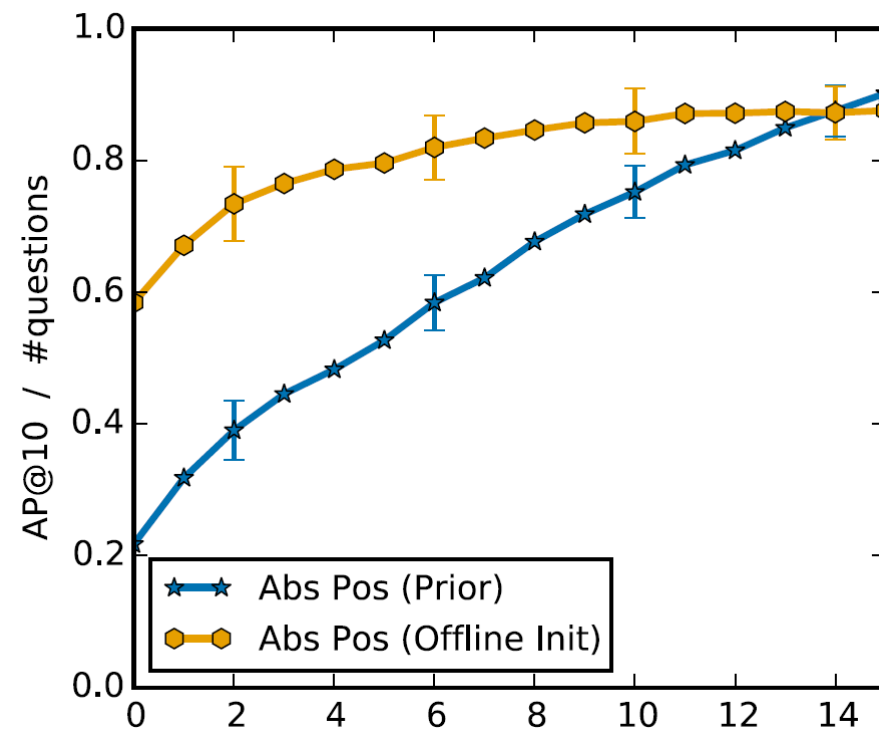
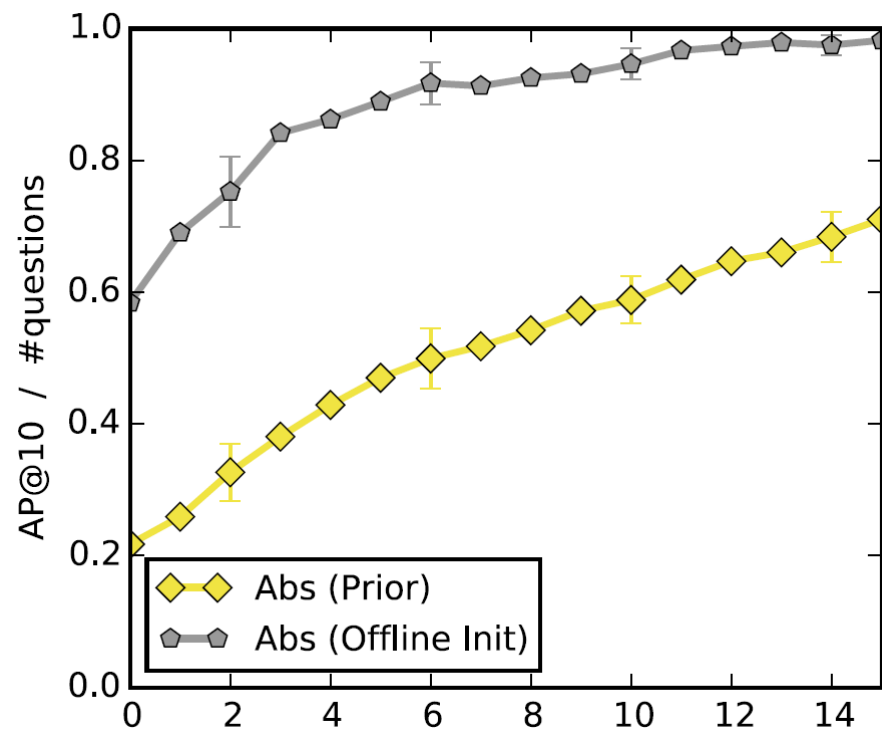


Are absolute or relative questions better?

After only 2 questions,
25% improvement
over a static model.



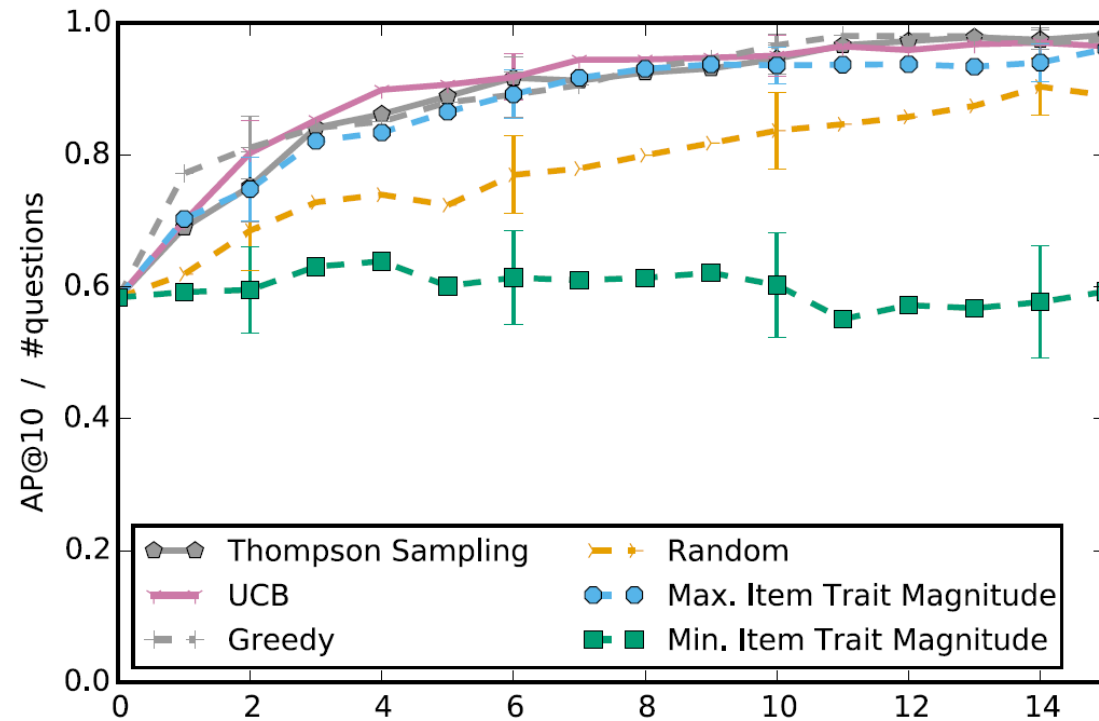
Does offline initialization help?



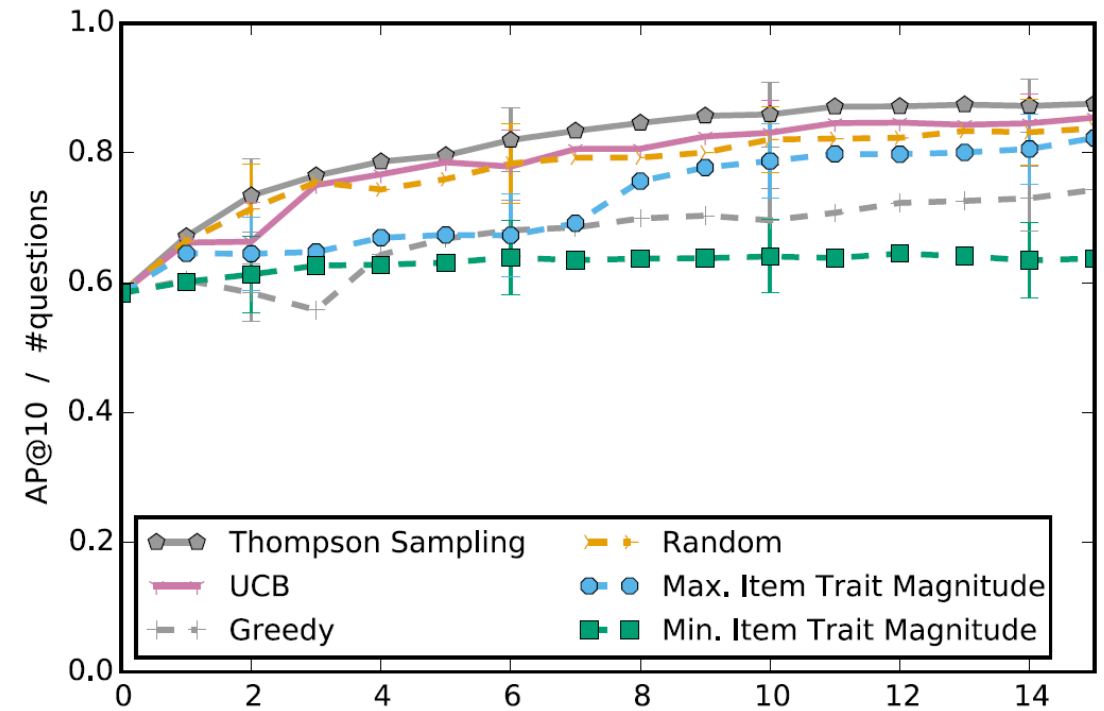
Dramatic benefits from offline embeddings.

Which question selection strategy is best?

ABS



ABS POS



Bandit-inspired strategies perform the most robustly.

CONCLUSIONS

- 1. Envision recommender systems that **converse with new users** to learn their preferences.
- 2. Fully **online learning** approach for recommendation -- both using **absolute** and **relative** feedback
- 3. Proposed various question selection strategies
- 4. Best performance can be achieved with **absolute** questions.
- 5. Effective learning with relative feedback is also possible.
- 6. Offline **learned embedding** greatly boosts initial performance.
- 7. Bandit-inspired question selection strategies are very effective.

Thank you!

