Interaction-Grounded Learning for Recommender Systems

Jessica Maghakian, Kishan Panaganti, Paul Mineiro,
Akanksha Saran, Cheng Tan

September 23, 2022
Workshop on Online Recommender Systems and User Modeling (ORSUM '22)

What should recommender systems optimize?

What should recommender systems optimize?

- Optimizating Click-Through Rate (CTR) promotes clickbait

What should recommender systems optimize?

- Optimizating Click-Through Rate (CTR) promotes clickbait
    - position bias (where on page result is displayed) [2]

What should recommender systems optimize?

- Optimizating Click-Through Rate (CTR) promotes clickbait
  - position bias (where on page result is displayed) [2]
  - caption bias (content of caption) [1]

What should recommender systems optimize?

- Optimizating Click-Through Rate (CTR) promotes clickbait
  - position bias (where on page result is displayed) [2]
  - caption bias (content of caption) [1]
  - curiosity gap (you won't believe this trick, doctors hate her!) [5, 6]

What should recommender systems optimize?

- Optimizating Click-Through Rate (CTR) promotes clickbait
  - position bias (where on page result is displayed) [2]
  - caption bias (content of caption) [1]
  - curiosity gap (you won't believe this trick, doctors hate her!) [5, 6]

- Clicks may even be a signal of user *dissatisfaction*

What should recommender systems optimize?

- Optimizating Click-Through Rate (CTR) promotes clickbait
  - position bias (where on page result is displayed) [2]
  - caption bias (content of caption) [1]
  - curiosity gap (you won't believe this trick, doctors hate her!) [5, 6]

- Clicks may even be a signal of user *dissatisfaction*
  - laboratory study of online news reading [4]

What should recommender systems optimize?

- Optimizating Click-Through Rate (CTR) promotes clickbait
  - position bias (where on page result is displayed) [2]
  - caption bias (content of caption) [1]
  - curiosity gap (you won't believe this trick, doctors hate her!) [5, 6]

- Clicks may even be a signal of user *dissatisfaction*
  - laboratory study of online news reading [4]
  - laboratory study of Spotify listening sessions [7]

What should recommender systems optimize?

- Optimizating Click-Through Rate (CTR) promotes clickbait
  - position bias (where on page result is displayed) [2]
  - caption bias (content of caption) [1]
  - curiosity gap (you won't believe this trick, doctors hate her!) [5, 6]
- Clicks may even be a signal of user *dissatisfaction*
  - laboratory study of online news reading [4]
  - laboratory study of Spotify listening sessions [7]
- Other implicit signals are equally complicated and nuanced

What should recommender systems optimize?

- Optimizating Click-Through Rate (CTR) promotes clickbait
  - position bias (where on page result is displayed) [2]
  - caption bias (content of caption) [1]
  - curiosity gap (you won't believe this trick, doctors hate her!) [5, 6]
- Clicks may even be a signal of user *dissatisfaction*
  - laboratory study of online news reading [4]
  - laboratory study of Spotify listening sessions [7]
- Other implicit signals are equally complicated and nuanced
  - dwell time is impacted by topic, readability and content length [3]

What should recommender systems optimize?

- Optimizating Click-Through Rate (CTR) promotes clickbait
  - position bias (where on page result is displayed) [2]
  - caption bias (content of caption) [1]
  - curiosity gap (you won't believe this trick, doctors hate her!) [5, 6]

- Clicks may even be a signal of user *dissatisfaction*
  - laboratory study of online news reading [4]
  - laboratory study of Spotify listening sessions [7]

- Other implicit signals are equally complicated and nuanced
  - dwell time is impacted by topic, readability and content length [3]
  - viewport time, dwell time and scroll patterns have a complicated temporal relationship [4]

What should recommender systems optimize?

- Optimizating Click-Through Rate (CTR) promotes clickbait
  - position bias (where on page result is displayed) [2]
  - caption bias (content of caption) [1]
  - curiosity gap (you won't believe this trick, doctors hate her!) [5, 6]
- Clicks may even be a signal of user *dissatisfaction*
  - laboratory study of online news reading [4]
  - laboratory study of Spotify listening sessions [7]
- Other implicit signals are equally complicated and nuanced
  - dwell time is impacted by topic, readability and content length [3]
  - viewport time, dwell time and scroll patterns have a complicated temporal relationship [4]

Implicit user feedback is not the true satisfaction signal!

Premise: learner optimizes for latent rewards by interacting with the environment and associating observed feedback with the unobservable true reward

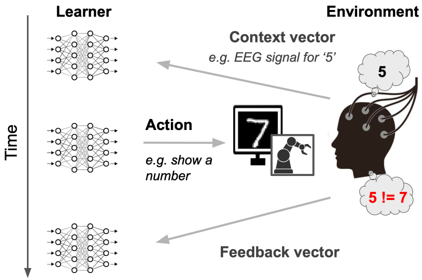


Figure 1: Original motivation for IGL [8]

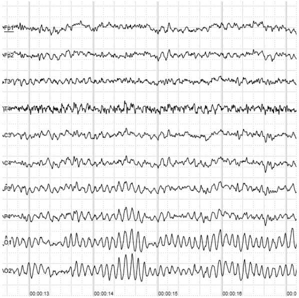
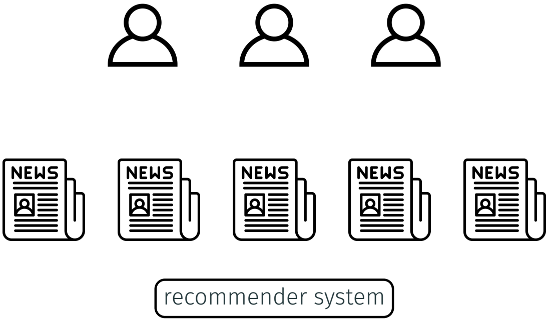


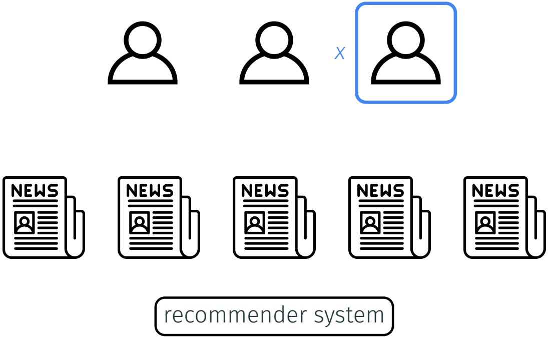
Figure 2: EEG signals[1]

[1]https://team.inria.fr/potioc/old-research-topics/eeg-signal-processing/

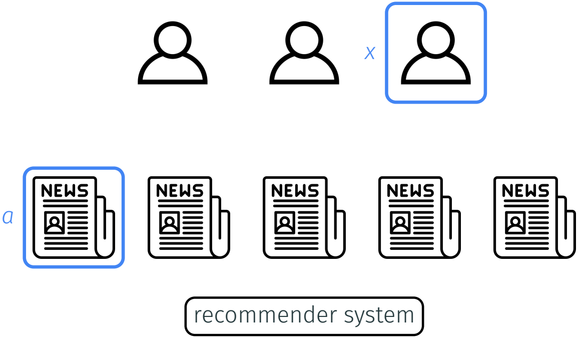
IGL FOR RECOMMENDER SYSTEMS



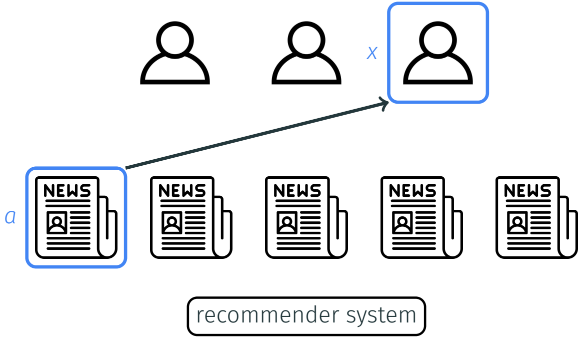
IGL FOR RECOMMENDER SYSTEMS



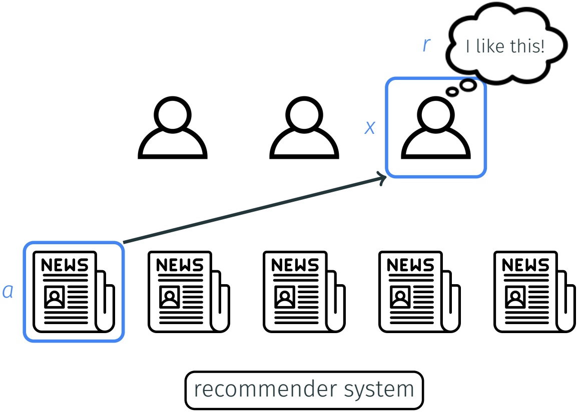
IGL FOR RECOMMENDER SYSTEMS

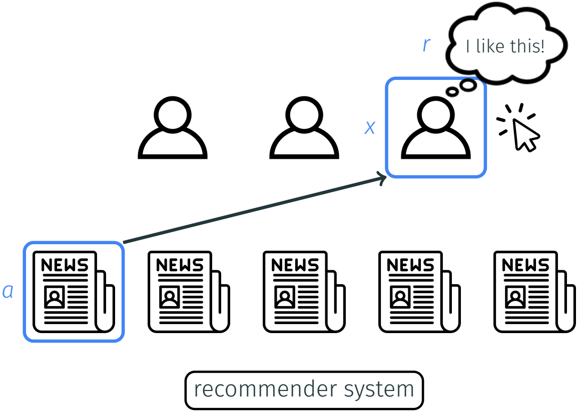


IGL FOR RECOMMENDER SYSTEMS

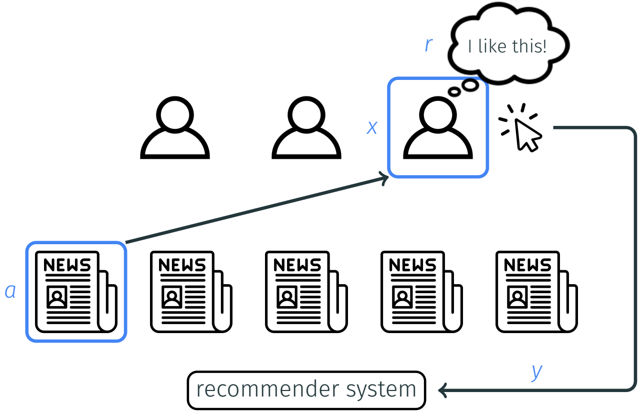


IGL FOR RECOMMENDER SYSTEMS

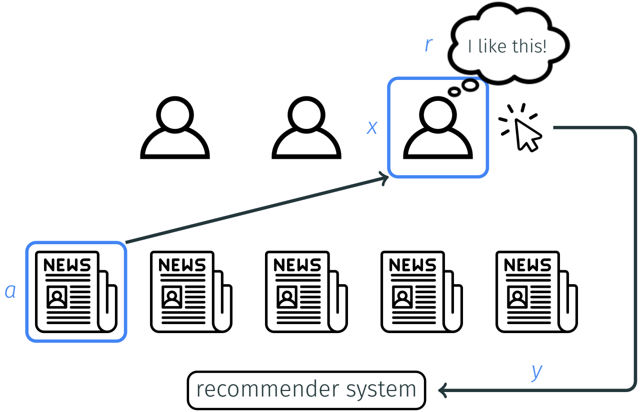



r
I like this!
x
NEWS
NEWS
NEWS
NEWS
NEWS
a
recommender system

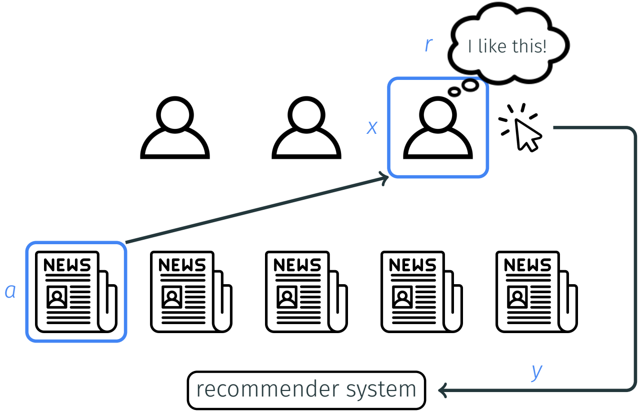
IGL FOR RECOMMENDER SYSTEMS



IGL FOR RECOMMENDER SYSTEMS



It is information theoretically impossible to solve IGL without assumptions about the relation between x, y and a [9].



It is information theoretically impossible to solve IGL without assumptions about the relation between x, y and a [9].

Consistent communication assumption: $y \perp a | x, r$

ARE 2 LATENT STATES SUFFICIENT FOR RECOMMENDER SYSTEMS?

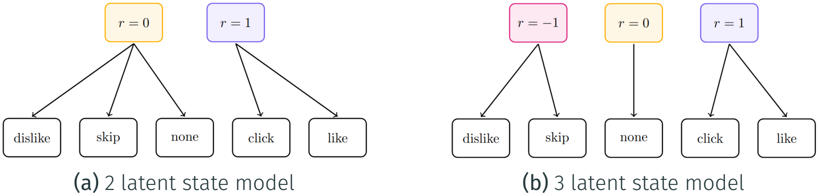


Figure 3: Example of the latent reward structure for a music recommendation system. In Fig. 1a, $r = 0$ corresponds to anything other than the user actively enjoying the content, whereas in Fig. 1b, lack of user enjoyment is split into indifference and active dissatisfaction.

Algorithm idea: anytime the posterior probability of an action is predicted to be more than twice the prior probability, we deduce $r \neq 0$.

Algorithm idea: anytime the posterior probability of an action is predicted to be more than twice the prior probability, we deduce $r \neq 0$.

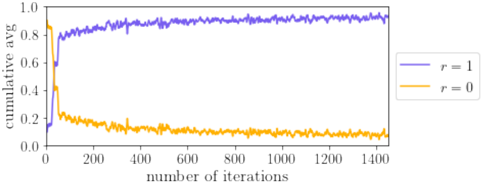


Figure 4: Performance of *IGL-P(2)* when there are two latent states.

IGL-P(2): AN ALGORITHM FOR TWO LATENT STATES

Algorithm idea: anytime the posterior probability of an action is predicted to be more than twice the prior probability, we deduce $r \neq 0$.

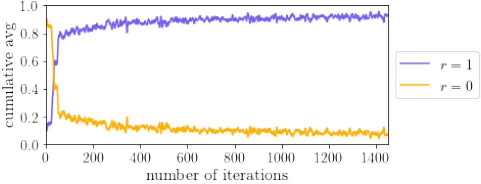


Figure 4: Performance of *IGL-P(2)* when there are two latent states.

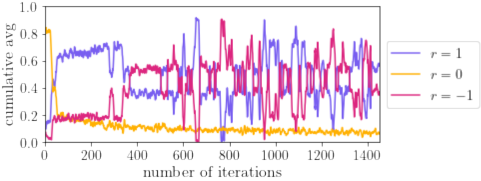


Figure 5: Performance of *IGL-P(2)* when there are three latent states.

IGL-P(3): AN ALGORITHM FOR THREE LATENT STATES

To tackle the three state case, we need to be able to distinguish between $r = 1$ and $r = -1$.

IGL-P(3): AN ALGORITHM FOR THREE LATENT STATES

To tackle the three state case, we need to be able to distinguish between $r = 1$ and $r = -1$.

Algorithm idea: use a negative oracle to detective “definitely negative” events, then use extreme event detection to identify $r = 1$.

To tackle the three state case, we need to be able to distinguish between $r = 1$ and $r = -1$.

Algorithm idea: use a negative oracle to detective "definitely negative" events, then use extreme event detection to identify $r = 1$.

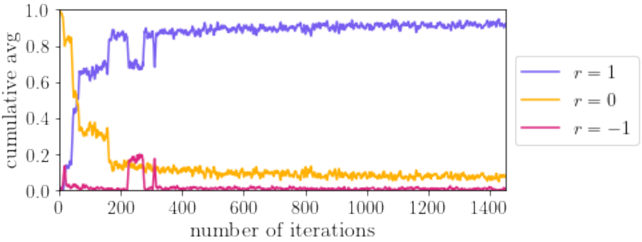


Figure 6: *IGL-P(3)* successfully maximizes user satisfaction while minimizing dissatisfaction.

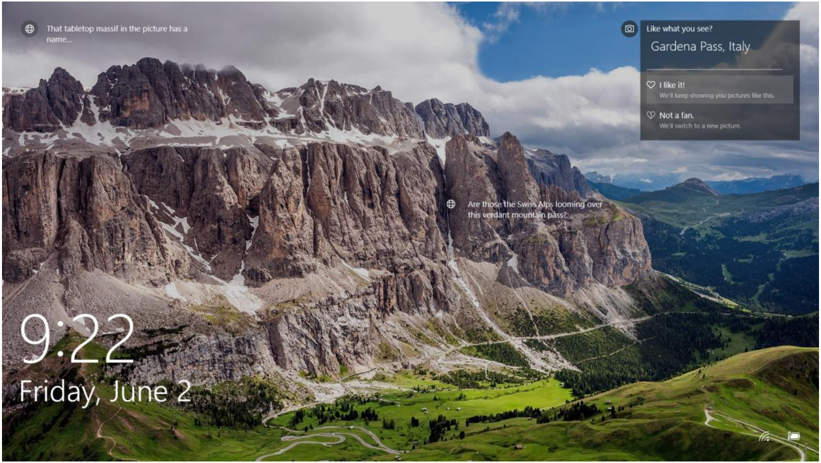


Figure 7: An example of a Windows Spotlight lock screen. Users can interact with the interface using 'click', 'like' or 'dislike' feedback.

Production setting: Windows Spotlight

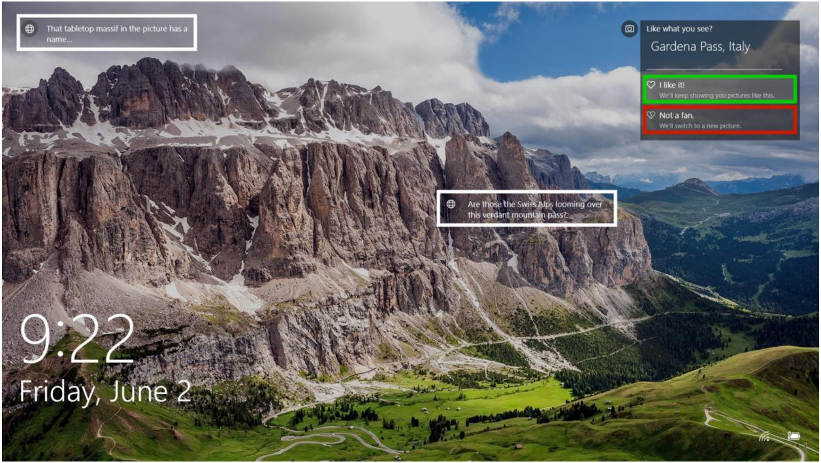


Figure 7: An example of a Windows Spotlight lock screen. Users can interact with the interface using 'click', 'like' or 'dislike' feedback.

Algorithm	*IGL-P(2)*	*IGL-P(3)*
Clicks	[0.926, 1.005, 1.091]	[0.999, 1.067, 1.152]
Likes	[0.914, 0.949, 0.988]	[0.985, 1.029, 1.054]
Dislikes	[1.141, 1.337, 1.557]	[0.751, 1.072, 1.274]

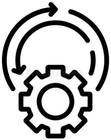
Table 1: Relative metrics lift over a production baseline. The production baseline uses a hand-engineered reward function which is not available to IGL algorithms. Shown are point estimates and associated bootstrap 95% confidence regions. *IGL-P(2)* erroneously increases dislikes to the detriment of other metrics. *IGL-P(3)* directionally improves over the hand-engineered reward function.

Benefits of IGL

Cheaper than
manual feature
engineering

Benefits of IGL

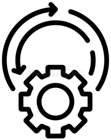
Cheaper than manual feature engineering



Easily adapt to evolving systems and users

Benefits of IGL

Cheaper than manual feature engineering



Easily adapt to evolving systems and users

Improved fairness for diverse users

Key takeaways from today's talk

1. Despite its popularity as a reward objective, implicit user feedback is ***not*** the true satisfaction signal.

Key takeaways from today's talk

1. Despite its popularity as a reward objective, implicit user feedback is ***not*** the true satisfaction signal.

2. The Interaction-Grounded Learning (IGL) framework directly targets the underlying user satisfaction.

Key takeaways from today's talk

1. Despite its popularity as a reward objective, implicit user feedback is ***not*** the true satisfaction signal.

2. The Interaction-Grounded Learning (IGL) framework directly targets the underlying user satisfaction.

3. IGL outperforms state-of-the-art real world systems by learning personalized reward functions for users.

Thank you! Any questions?

References I

K. Hofmann, F. Behr, and F. Radlinski.
On caption bias in interleaving experiments.
In *Proceedings of the 21st ACM international conference on Information and knowledge management*, pages 115–124, 2012.

K. Hofmann, A. Schuth, A. Bellogin, and M. d. Rijke.
Effects of position bias on click-based recommender evaluation.
In *European Conference on Information Retrieval*, pages 624–630. Springer, 2014.

Y. Kim, A. Hassan, R. W. White, and I. Zitouni.
Modeling dwell time to predict click-level satisfaction.
In *Proceedings of the 7th ACM international conference on Web search and data mining*, pages 193–202, 2014.

H. Lu, M. Zhang, and S. Ma.
Between clicks and satisfaction: Study on multi-phase user preferences and satisfaction for online news reading.
In *The 41st International ACM SIGIR Conference on Research & Development in Information Retrieval*, pages 435–444, 2018.

M. Potthast, S. Köpsel, B. Stein, and M. Hagen.
Clickbait detection.
In *European conference on information retrieval*, pages 810–817. Springer, 2016.

References II

K. Scott.
You won't believe what's in this paper! clickbait, relevance and the curiosity gap.
Journal of pragmatics, 175:53–66, 2021.

H. Wen, L. Yang, and D. Estrin.
Leveraging post-click feedback for content recommendations.
In *Proceedings of the 13th ACM Conference on Recommender Systems*, pages 278–286, 2019.

T. Xie, J. Langford, P. Mineiro, and I. Momennejad.
Interaction-grounded learning.
In *International Conference on Machine Learning*, pages 11414–11423. PMLR, 2021.

T. Xie, A. Saran, D. J. Foster, L. Molu, I. Momennejad, N. Jiang, P. Mineiro, and J. Langford.
Interaction-grounded learning with action-inclusive feedback.
arXiv preprint arXiv:2206.08364, 2022.