ELIXIR: Learning from User Feedback on Explanations to Improve Recommender Models

Azin Ghazimatin, Soumajit Pramanik, Rishiraj Saha Roy, Gerhard Weikum

Max Planck Institute for Informatics, Saarland Informatics Campus
Saarbrücken, Germany

The Web Conference, 2021

Motivation

User u

Recommendation (*rec*)

Fight Club

Explanation (*exp*):

7 years in Tibet

The Prestige

Pulp Fiction

Motivation

User u

Recommendation (*rec*)

Fight Club

Explanation (*exp*):

7 years in Tibet

The Prestige

Pulp Fiction

I want to stop seeing this item and the like.

I want to see more items like this.

Motivation

So far: item-level rating

User u

Recommendation (*rec*)

Fight Club

Explanation (*exp*):

7 years in Tibet

The Prestige

Pulp Fiction

I want to stop seeing this item and the like.

I want to see more items like this.

Motivation

So far: item-level rating

User u

Recommendation (*rec*)

Fight Club

Explanation (*exp*):

7 years in Tibet

The Prestige

Pulp Fiction

What is it that you (dis)like about the *Fight Club*?

Brad Pitt

Surprise Ending

Violence

...

I want to stop seeing this item and the like.

I want to see more items like this.

Motivation

The idea: Leverage feedback on pairs of *explanations* and *recommendations*

User u

Recommendation (*rec*)

Fight Club

Explanation (*exp*):

7 years in Tibet

The Prestige

Pulp Fiction

Similar aspect of (*rec, exp*):

Cast: Brad Pitt

Storyline: Surprise Ending

Content: Violence

Feedback on (*rec, exp*):

I want to stop seeing this item and the like.

I want to see more items like this.

ELIXIR

Efficient **L**earning from **I**tem-level e**X**planations **I**n **R**ecommenders

(Improving future recommendations through *feedback on pairs of recommendation and explanation items*)

User u

Recommendation (*rec*) at time T

Fight Club

Explanation (*exp*):

7 years in Tibet

The Prestige

Pulp Fiction

Similar aspect of (*rec, exp*):

Cast: Brad Pitt

Storyline: Surprise Ending

Content: Violence

Feedback on (*rec, exp*):

ELIXIR

Collect and densify Feedback

Incorporate Feedback

Recommendation (*rec*) at time $T+1$

Ocean's Eleven

ELIXIR: Feedback Matrix

User u

Recommendation (*rec*) at time T

Fight Club

Explanation (*exp*):

7 years in Tibet

The Prestige

Pulp Fiction

Do you like the similarity between (*rec, exp*)?

$\Rightarrow$

Matrix F_u

ELIXIR: Feedback Densification

★ Label propagation [Zhu and Ghahramani, 2002] on pairs of items to mitigate sparsity $\Rightarrow F_u^d$

Zhu and Ghahramani, *Learning from labeled and unlabeled data with label propagation*, CMU Technical Report (2002)

ELIXIR: Feedback Densification

- ★ Label propagation [Zhu and Ghahramani, 2002] on pairs of items to mitigate sparsity $\Rightarrow F_u^d$
- ★ Present pairs $(\overrightarrow{v_i}, \overrightarrow{v_j})$ with pseudo-items (geometric mean of two vectors)

$$\overrightarrow{v_{ij}} = (\overrightarrow{v_i} \otimes \overrightarrow{v_j})^{\frac{1}{2}}$$

Zhu and Ghahramani, *Learning from labeled and unlabeled data with label propagation*, CMU Technical Report (2002)

ELIXIR: Feedback Densification

- ★ Label propagation [Zhu and Ghahramani, 2002] on pairs of items to mitigate sparsity $\Rightarrow F_u^d$
- ★ Present pairs $(\overrightarrow{v_i}, \overrightarrow{v_j})$ with pseudo-items (geometric mean of two vectors)

$$\overrightarrow{v_{ij}} = (\overrightarrow{v_i} \otimes \overrightarrow{v_j})^{\frac{1}{2}}$$

- ★ The affinity matrix W encodes pair-pair similarity. This makes W huge! ($W \in \mathbb{R}^{|I|^2 \times |I|^2}$, I is the set of items)

Zhu and Ghahramani, *Learning from labeled and unlabeled data with label propagation*, CMU Technical Report (2002)

ELIXIR: Feedback Densification

- ★ Label propagation [Zhu and Ghahramani, 2002] on pairs of items to mitigate sparsity $\Rightarrow F_u^d$
- ★ Present pairs $(\overrightarrow{v_i}, \overrightarrow{v_j})$ with pseudo-items (geometric mean of two vectors)

$$\overrightarrow{v_{ij}} = (\overrightarrow{v_i} \otimes \overrightarrow{v_j})^{\frac{1}{2}}$$

- ★ The affinity matrix W encodes pair-pair similarity. This makes W huge! ($W \in \mathbb{R}^{|I|^2 \times |I|^2}$, I is the set of items)
- ★ Propagate labels only on kNN of the labeled pairs
  - naive approach: $O(|I|^2)$
  - our approach:
    - find kNN of v_{ij} among the items in I using locality sensitive hashing (LSH) , we call this set kNN_{ij}^I.
    - Search for kNN of v_{ij} in $kNN_{ij}^I \times kNN_{ij}^I$.
    - This reduces the time complexity to $O(|I|)$ for small k.

Zhu and Ghahramani, *Learning from labeled and unlabeled data with label propagation*, CMU Technical Report (2002)

ELIXIR: Feedback Incorporation

The Idea: Learn user-specific parameters $\overrightarrow{w_u}$ of a mapping function g to produce user-specific item representations.

$$\min_{\overrightarrow{w_u}} \frac{1}{m} \sum_{v_i, v_j} F_u^d(v_i, v_j) \cdot (sim(\overrightarrow{v_i}, \overrightarrow{v_j}) - sim(g(\overrightarrow{v_i}, \overrightarrow{w_u}), g(\overrightarrow{v_j}, \overrightarrow{w_u}))) + \gamma ||\overrightarrow{w_u}||^2$$

No. feedback pairs

similarity function

Similarity of items in **negative** (**positive**) pairs **decreases** (**increases**).

ELIXIR: Feedback Incorporation

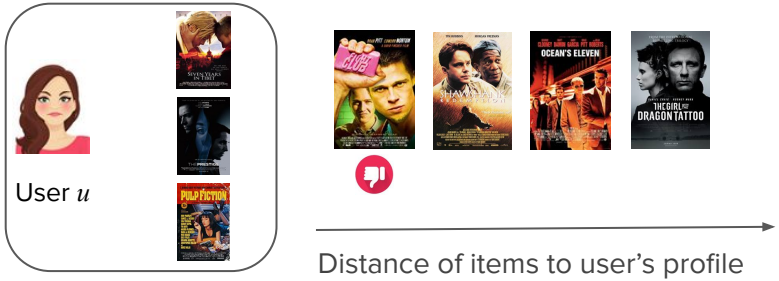
The Idea: Learn user-specific parameters $\overrightarrow{w_u}$ of a mapping function g to produce user-specific item representations.

$$\min_{\overrightarrow{w_u}} \frac{1}{m} \sum_{v_i, v_j} F_u^d(v_i, v_j) \cdot (sim(\overrightarrow{v_i}, \overrightarrow{v_j}) - sim(g(\overrightarrow{v_i}, \overrightarrow{w_u}), g(\overrightarrow{v_j}, \overrightarrow{w_u}))) + \gamma ||\overrightarrow{w_u}||^2$$

No. feedback pairs

similarity function

Similarity of items in **negative** (**positive**) pairs **decreases** (**increases**).



ELIXIR: Feedback Incorporation

The Idea: Learn user-specific parameters $\overrightarrow{w_u}$ of a mapping function g to produce user-specific item representations.

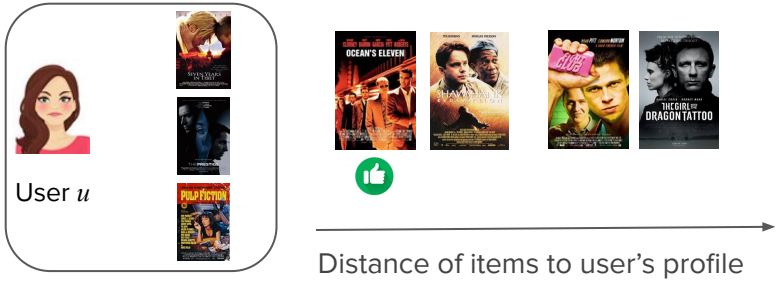
$$\min_{\overrightarrow{w_u}} \frac{1}{m} \sum_{v_i, v_j} F_u^d(v_i, v_j) \cdot (sim(\overrightarrow{v_i}, \overrightarrow{v_j}) - sim(g(\overrightarrow{v_i}, \overrightarrow{w_u}), g(\overrightarrow{v_j}, \overrightarrow{w_u}))) + \gamma ||\overrightarrow{w_u}||^2$$

No. feedback pairs

similarity function

Similarity of items in **negative** (**positive**) pairs **decreases** (**increases**).

After feedback incorporation

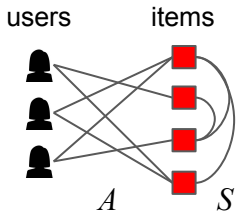


Instantiation of ELIXIR

★ We instantiated our framework with RecWalk^{PR} [Nikolakopoulos and Karypis, WSDM'19]

$$\overrightarrow{PPR}(u,.) = \alpha \cdot \overrightarrow{e_u} + (1-\alpha) \cdot \overrightarrow{PPR}(u,.) \cdot [\beta A + (1-\beta) S]$$

★ Explanations generated by PRINCE [Ghazimatin et al., WSDM'20]

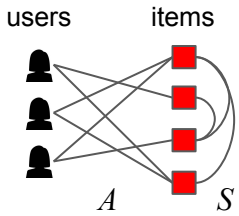


Nikolakopoulos and Karypis, *RecWalk: Nearly uncoupled random walks for top-n recommendation*, WSDM'19
Ghazimatin et al., *PRINCE: Provider-side Interpretability with Counterfactual Explanations in Recommender Systems*, WSDM'20

Instantiation of ELIXIR

- ★ We instantiated our framework with RecWalk^{PR} [Nikolakopoulos and Karypis, WSDM'19]

$$\overrightarrow{PPR}(u,.) = \alpha \cdot \overrightarrow{e_u} + (1-\alpha) \cdot \overrightarrow{PPR}(u,.) \cdot [\beta A + (1-\beta) S]$$



- ★ Explanations generated by PRINCE [Ghazimatin et al., WSDM'20]
- ★ Learning $\overrightarrow{w_u}$ ($g(\overrightarrow{v_i}, \overrightarrow{w_u}) = \overrightarrow{v_i} + \overrightarrow{w_u}$):

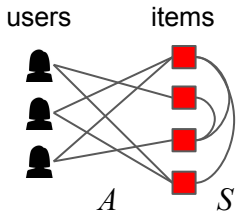
$$\min_{\overrightarrow{w_u}} \frac{1}{m} \sum_{v_i, v_j} F_u^d(v_i, v_j) \cdot (\underline{cos}(\overrightarrow{v_i}, \overrightarrow{v_j}) - \underline{cos}(\underline{\overrightarrow{v_i} + \overrightarrow{w_u}}, \underline{\overrightarrow{v_j} + \overrightarrow{w_u}})) + \gamma ||\overrightarrow{w_u}||^2$$

Nikolakopoulos and Karypis, *RecWalk: Nearly uncoupled random walks for top-n recommendation*, WSDM'19
Ghazimatin et al., *PRINCE: Provider-side Interpretability with Counterfactual Explanations in Recommender Systems*, WSDM'20

Instantiation of ELIXIR

★ We instantiated our framework with RecWalk^{PR} [Nikolakopoulos and Karypis, WSDM'19]

$$\overrightarrow{PPR}(u,.) = \alpha \cdot \overrightarrow{e_u} + (1-\alpha) \cdot \overrightarrow{PPR}(u,.) \cdot [\beta A + (1-\beta) S]$$



★ Explanations generated by PRINCE [Ghazimatin et al., WSDM'20]

★ Learning $\overrightarrow{w_u}$ ($g(\overrightarrow{v_i}, \overrightarrow{w_u}) = \overrightarrow{v_i} + \overrightarrow{w_u}$):

$$\min_{\overrightarrow{w_u}} \frac{1}{m} \sum_{v_i, v_j} F_u^d(v_i, v_j) \cdot (\underline{cos}(\overrightarrow{v_i}, \overrightarrow{v_j}) - \underline{cos}(\underline{\overrightarrow{v_i} + \overrightarrow{w_u}}, \underline{\overrightarrow{v_j} + \overrightarrow{w_u}})) + \gamma ||\overrightarrow{w_u}||^2$$

★ Updating the model:

$$S_u(v_i, v_j) = cos(\overrightarrow{v_i} + \overrightarrow{w_u}, \overrightarrow{v_j} + \overrightarrow{w_u})$$

$$\overrightarrow{PPR}(u,.) = \alpha \cdot \overrightarrow{e_u} + (1-\alpha) \cdot \overrightarrow{PPR}(u,.) \cdot [\beta A + (1-\beta) S_u]$$

Nikolakopoulos and Karypis, *RecWalk: Nearly uncoupled random walks for top-n recommendation*, WSDM'19
Ghazimatin et al., *PRINCE: Provider-side Interpretability with Counterfactual Explanations in Recommender Systems*, WSDM'20

User Study for Data Collection

- ★ User studies in two domains: movies and books
- ★ Phase 1: Building users' profiles
  - 50 liked movies selected from Movielens website
  - 50 liked books selected from Goodreads website
- ★ Phase 2: Collecting feedback on items and pairs
  - 30 recommendations (generated at time T)
    - generated by RecWalk
    - S: cosine similarity of item representations learned by applying NMF on item-feature matrix
  - 150 pairs corresponding to recommendations and their top 5 explanations
  - 150 pairs corresponding to recommendations and their bottom 5 explanations
- ★ Phase 3: Collecting feedback on final recommendations (evaluation at time $T+1$)

Scenario	#Item feedback	#Pair feedback	Sessions	Hours
Phase1	50	-	1	2
Phase2	30	300	5	10
Phase3	72	-	1	2
Total	152	300	7	14
Total (All users)	≈4000	7500	175	350

User Study for Data Collection

- ★ User studies in two domains: movies and books
- ★ Phase 1: Building users' profiles
  - 50 liked movies selected from Movielens website
  - 50 liked books selected from Goodreads website
- ★ Phase 2: Collecting feedback on items and pairs
  - 30 recommendations (generated at time T)
    - generated by RecWalk
    - S: cosine similarity of item representations learned by applying NMF on item-feature matrix
  - 150 pairs corresponding to recommendations and their top 5 explanations
  - 150 pairs corresponding to recommendations and their bottom 5 explanations
- ★ Phase 3: Collecting feedback on final recommendations (evaluation at time $T+1$)

Scenario	#Item feedback	#Pair feedback	Sessions	Hours
Phase1	50	-	1	2
Phase2	30	300	5	10
Phase3	72	-	1	2
Total	152	300	7	14
Total (All users)	≈4000	7500	175	350

User Study for Data Collection

- ★ User studies in two domains: movies and books
- ★ Phase 1: Building users' profiles
  - 50 liked movies selected from Movielens website
  - 50 liked books selected from Goodreads website
- ★ Phase 2: Collecting feedback on items and pairs
  - 30 recommendations (generated at time T)
    - generated by RecWalk
    - S: cosine similarity of item representations learned by applying NMF on item-feature matrix
  - 150 pairs corresponding to recommendations and their top 5 explanations
  - 150 pairs corresponding to recommendations and their bottom 5 explanations
- ★ Phase 3: Collecting feedback on final recommendations (evaluation at time $T+1$)

Scenario	#Item feedback	#Pair feedback	Sessions	Hours
Phase1	50	-	1	2
Phase2	30	300	5	10
Phase3	72	-	1	2
Total	152	300	7	14
Total (All users)	≈4000	7500	175	350

User Study for Data Collection

- ★ User studies in two domains: movies and books
- ★ Phase 1: Building users' profiles
  - 50 liked movies selected from Movielens website
  - 50 liked books selected from Goodreads website
- ★ Phase 2: Collecting feedback on items and pairs
  - 30 recommendations (generated at time T)
    - generated by RecWalk
    - S: cosine similarity of item representations learned by applying NMF on item-feature matrix
  - 150 pairs corresponding to recommendations and their top 5 explanations
  - 150 pairs corresponding to recommendations and their bottom 5 explanations
- ★ Phase 3: Collecting feedback on final recommendations (evaluation at time $T+1$)

Scenario	#Item feedback	#Pair feedback	Sessions	Hours
Phase1	50	-	1	2
Phase2	30	300	5	10
Phase3	72	-	1	2
Total	152	300	7	14
Total (All users)	≈4000	7500	175	350

Evaluation Configurations

- ★ Item-level feedback (baseline)
- ★ Pair-level feedback with explanations
- ★ Pair-level feedback with random items
  - explanation items are replaced by the least relevant items from user's history
- ★ Item+pair-level feedback with explanations
- ★ Item+pair-level feedback with random items
  - explanation items are replaced by the least relevant items from user's history

Results

- ★ Metrics: P@k, MAP@k, nDCG@k at time $T+1$
- ★ Key findings:
  - ○ Pair-level feedback improves recommendation.

Setup [Movies]	**P@3**	**P@5**	**P@10**
Item-level	0.253	0.368	0.484
Pair-level(Exp-1)	0.506*	0.592*	0.58*
Pair-level(Exp-3)	0.506*	0.568*	0.596*
Pair-level(Exp-5)	0.48*	0.512*	0.568*
Pair-level(Rand-1)	0.453*	0.504*	0.56*
Item+Pair-level (Exp-5)	0.533*	0.544*	0.596*
Item+Pair-level (Rand-5)	0.453*	0.488*	0.572*
Setup [Books]	**P@3**	**P@5**	**P@10**
Item-level	0.253	0.336	0.436
Pair-level(Exp-1)	0.506*	0.528*	0.54*
Pair-level(Exp-3)	0.506*	0.536*	0.58*
Pair-level(Exp-5)	0.586*	0.6*	0.656*
Pair-level(Rand-1)	0.48*	0.504	0.504
Item+Pair-level (Exp-5)	0.493*	0.456*	0.56*
Item+Pair-level (Rand-5)	0.386*	0.416*	0.516*

Results

- ★ Metrics: P@k, MAP@k, nDCG@k at time $T+1$
- ★ Key findings:
  - Pair-level feedback improves recommendation.
  - Pair-level feedback is more discriminative than item-level.

Setup [Movies]	**P@3**	**P@5**	**P@10**
Item-level	0.253	0.368	0.484
Pair-level(Exp-1)	0.506*	0.592*	0.58*
Pair-level(Exp-3)	0.506*	0.568*	0.596*
Pair-level(Exp-5)	0.48*	0.512*	0.568*
Pair-level(Rand-1)	0.453*	0.504*	0.56*
Item+Pair-level (Exp-5)	0.533*	0.544*	0.596*
Item+Pair-level (Rand-5)	0.453*	0.488*	0.572*
Setup [Books]	**P@3**	**P@5**	**P@10**
Item-level	0.253	0.336	0.436
Pair-level(Exp-1)	0.506*	0.528*	0.54*
Pair-level(Exp-3)	0.506*	0.536*	0.58*
Pair-level(Exp-5)	0.586*	0.6*	0.656*
Pair-level(Rand-1)	0.48*	0.504	0.504
Item+Pair-level (Exp-5)	0.493*	0.456*	0.56*
Item+Pair-level (Rand-5)	0.386*	0.416*	0.516*

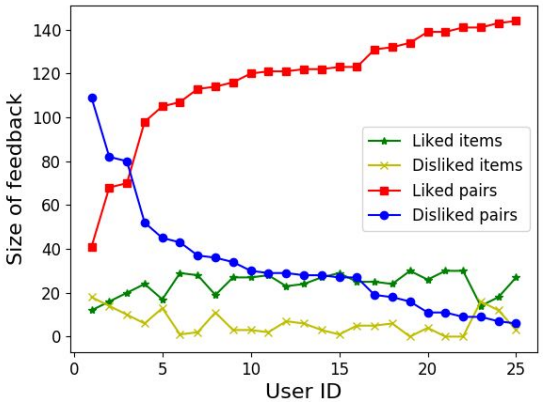
Results

- ★ Metrics: P@k, MAP@k, nDCG@k at time $T+1$
- ★ Key findings:
  - Pair-level feedback improves recommendation.
  - Pair-level feedback is more discriminative than item-level.
  - Using explanations for pair-level feedback is essential.

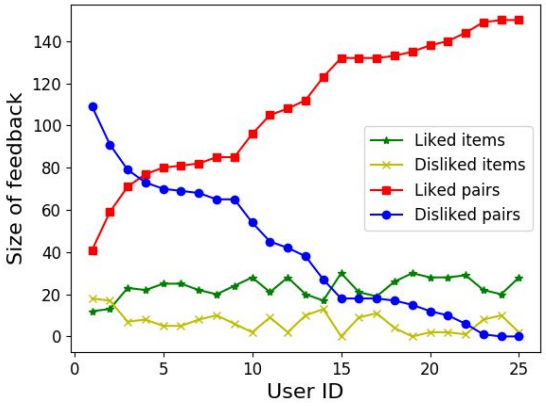
Setup [Movies]	**P@3**	**P@5**	**P@10**
Item-level	0.253	0.368	0.484
Pair-level(Exp-1)	0.506*	0.592*	0.58*
Pair-level(Exp-3)	0.506*	0.568*	0.596*
Pair-level(Exp-5)	0.48*	0.512*	0.568*
Pair-level(Rand-1)	0.453*	0.504*	0.56*
Item+Pair-level (Exp-5)	0.533*	0.544*	0.596*
Item+Pair-level (Rand-5)	0.453*	0.488*	0.572*
Setup [Books]	**P@3**	**P@5**	**P@10**
Item-level	0.253	0.336	0.436
Pair-level(Exp-1)	0.506*	0.528*	0.54*
Pair-level(Exp-3)	0.506*	0.536*	0.58*
Pair-level(Exp-5)	0.586*	0.6*	0.656*
Pair-level(Rand-1)	0.48*	0.504	0.504
Item+Pair-level (Exp-5)	0.493*	0.456*	0.56*
Item+Pair-level (Rand-5)	0.386*	0.416*	0.516*

Some insights

★ Users give much more positive feedback than negative.



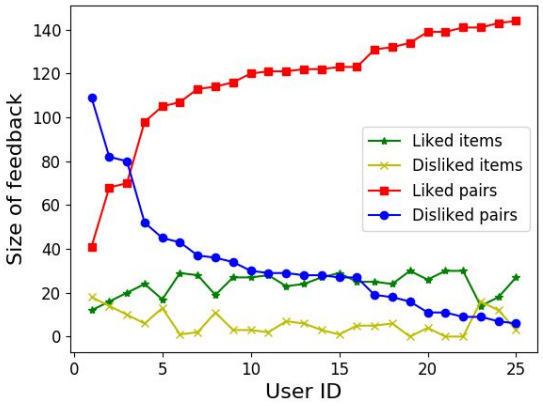
-Movies-



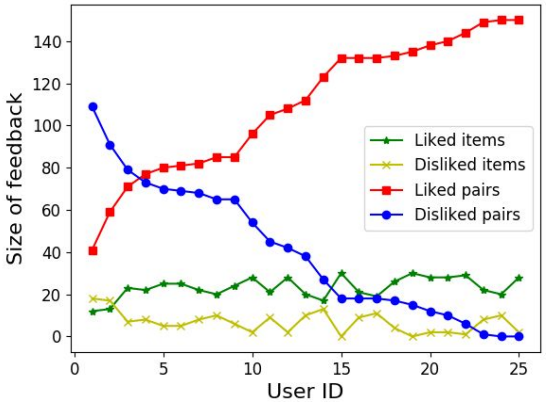
-Books-

Some insights

- ★ Users give much more positive feedback than negative.
- ★ User behavior carries over from items to pairs (pearson correlation ≥ 0.5).



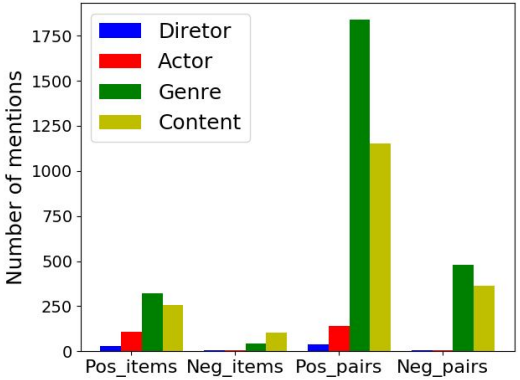
-Movies-



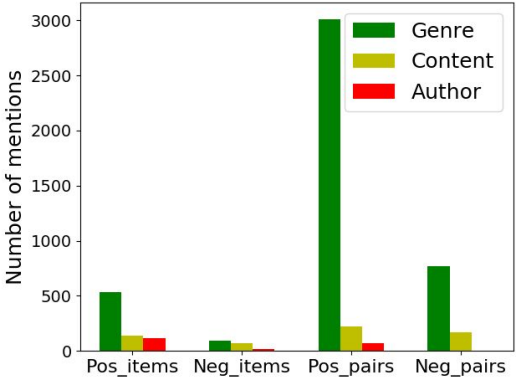
-Books-

Some insights

- ★ Users give much more positive feedback than negative.
- ★ User behavior carries over from items to pairs (pearson correlation ≥ 0.5)
- ★ Users mostly give feedback on shared genres/content between *rec* and *exp*



-Movies-



-Books-

Summary

ELIXIR

- ★ A human-in-the-loop framework that elicits lightweight user feedback on pairs of recommendation and explanation items.
- ★ Learns user-specific item representations based on user feedback.
- ★ Instantiated with recommenders based on random walks with restart.
- ★ Major gains in recommendation quality over item-level feedback, shown with real user studies.

Summary

ELIXIR

- ★ A human-in-the-loop framework that elicits lightweight user feedback on pairs of recommendation and explanation items.
- ★ Learns user-specific item representations based on user feedback.
- ★ Instantiated with recommenders based on random walks with restart.
- ★ Major gains in recommendation quality over item-level feedback, shown with real user studies.

Thanks!

Explanations in Action

- ★ The role of explanations is mostly limited to improving user trust and satisfaction.
- ★ Critiqued-enabled recommenders [Jin et al., CIKM'19], [Luo et al., SIGIR'20]
  - Feedback on ***explicit*** and ***coarse-grained*** item properties
  - ***Negative*** feedback
- ★ ELIXIR
  - elicits ***lightweight*** user feedback on ***similarity*** of recommendation and explanation items.
  - supports both ***positive*** and negative ***feedback***.

Jin et al., *MusicBot: Evaluating Critiquing-Based Music Recommenders with Conversational Interaction*, CIKM'19
Luo et al., *Deep Critiquing for VAE-based Recommender Systems*, SIGIR'20

Explanations in Action

- ★ The role of explanations is mostly limited to improving user trust and satisfaction.
- ★ Critiqued-enabled recommenders [Jin et al., CIKM'19], [Luo et al., SIGIR'20]
  - Feedback on ***explicit*** and ***coarse-grained*** item properties
  - ***Negative*** feedback

✓ I want to stop seeing this item and the like.

✗ I want to see more items like this.

- ★ ELIXIR
  - elicits ***lightweight*** user feedback on ***similarity*** of recommendation and explanation items.
  - supports both ***positive*** and negative ***feedback***.

✓ I want to stop seeing this item and the like.

✓ I want to see more items like this.

Jin et al., *MusicBot: Evaluating Critiquing-Based Music Recommenders with Conversational Interaction*, CIKM'19
Luo et al., *Deep Critiquing for VAE-based Recommender Systems*, SIGIR'20

Anecdotal Examples

	History h	Recommendations at time T *(recs)*	Feedback on *recs*	Explanations *exp* from h	Feedback on *(recs, exp)* pairs	*(recs, exp)* Similarity	Recommendation at time $T+1$ using **only item-level feedback**	Recommendation at time $T+1$ using **both item and pair-level feedback**
Movies	User	Divergent		Doctor Strange		Genre: Science Fiction	Space	The BFG
		Harry Potter and the Deathly Hallows: Part 1		Charlie and the Chocolate Factory		Content: Based on a book, fantasy		
Books	User	Ultimate Guide to Google AdWords		Hooked: How to Build Habit-Forming Products		Genre: Technology	DotCom Secrets	Mindset: The New Psychology of Success
		The Economic Naturalist		Influence: The Psychology of Persuasion		Genre/Content: Self-Help, Psychology		

Setup [Movies]	**P@3**	**P@5**	**P@10**
Item-level	0.253	0.368	0.484
Pair-level(Exp-1)	0.506*	0.592*	0.58*
Pair-level(Exp-3)	0.506*	0.568*	0.596*
Pair-level(Exp-5)	0.48*	0.512*	0.568*
Pair-level(Rand-1)	0.453*	0.504*	0.56*
Item+Pair-level (Exp-5)	0.533*	0.544*	0.596*
Item+Pair-level (Rand-5)	0.453*	0.488*	0.572*
Setup [Books]	**P@3**	**P@5**	**P@10**
Item-level	0.253	0.336	0.436
Pair-level(Exp-1)	0.506*	0.528*	0.54*
Pair-level(Exp-3)	0.506*	0.536*	0.58*
Pair-level(Exp-5)	0.586*	0.6*	0.656*
Pair-level(Rand-1)	0.48*	0.504	0.504
Item+Pair-level (Exp-5)	0.493*	0.456*	0.56*
Item+Pair-level (Rand-5)	0.386*	0.416*	0.516*

Results

- ★ Metrics: P@k, MAP@k, nDCG@k at time $T+1$
- ★ Key findings:
  - ○ Pair-level feedback improves recommendation.

Setup [Movies]	P@3	P@5	P@10
Item-level	0.253	0.368	0.484
Pair-level (Exp-1)	0.506*	**0.592***	0.580*
Pair-level (Exp-3)	0.506*	0.568*	**0.596***
Pair-level (Exp-5)	0.480*	0.512*	0.568*
Pair-level (Rand-5)	0.453*	0.504*	0.560*
Item+Pair-level (Exp-5)	**0.533***	0.544*	**0.596***
Item+Pair-level (Rand-5)	0.453*	0.488*	0.572*
Setup [Books]	**P@3**	**P@5**	**P@10**
Item-level	0.253	0.336	0.436
Pair-level (Exp-1)	0.506*	0.528*	0.54*
Pair-level (Exp-3)	0.506*	0.536*	0.58*
Pair-level (Exp-5)	**0.586***	**0.600***	**0.656***
Pair-level (Rand-5)	0.480*	0.504*	0.504
Item+Pair-level (Exp-5)	0.493*	0.456*	0.560*
Item+Pair-level (Rand-5)	0.386*	0.416*	0.516*

Results

- ★ Metrics: P@k, MAP@k, nDCG@k at time $T+1$
- ★ Key findings:
  - Pair-level feedback improves recommendation.
  - Pair-level feedback is more discriminative than item-level.

Setup [Movies]	**P@3**	**P@5**	**P@10**
Item-level	0.253	0.368	0.484
Pair-level (Exp-1)	0.506*	**0.592***	0.580*
Pair-level (Exp-3)	0.506*	0.568*	**0.596***
Pair-level (Exp-5)	0.480*	0.512*	0.568*
Pair-level (Rand-5)	0.453*	0.504*	0.560*
Item+Pair-level (Exp-5)	**0.533***	0.544*	**0.596***
Item+Pair-level (Rand-5)	0.453*	0.488*	0.572*
Setup [Books]	**P@3**	**P@5**	**P@10**
Item-level	0.253	0.336	0.436
Pair-level (Exp-1)	0.506*	0.528*	0.54*
Pair-level (Exp-3)	0.506*	0.536*	0.58*
Pair-level (Exp-5)	**0.586***	**0.600***	**0.656***
Pair-level (Rand-5)	0.480*	0.504*	0.504
Item+Pair-level (Exp-5)	0.493*	0.456*	0.560*
Item+Pair-level (Rand-5)	0.386*	0.416*	0.516*

Results

- ★ Metrics: P@k, MAP@k, nDCG@k at time $T+1$
- ★ Key findings:
  - Pair-level feedback improves recommendation.
  - Pair-level feedback is more discriminative than item-level.
  - Using explanations for pair-level feedback is essential.

Setup [Movies]	**P@3**	**P@5**	**P@10**
Item-level	0.253	0.368	0.484
Pair-level (Exp-1)	0.506*	**0.592***	0.580*
Pair-level (Exp-3)	0.506*	0.568*	**0.596***
Pair-level (Exp-5)	0.480*	0.512*	0.568*
Pair-level (Rand-5)	0.453*	0.504*	0.560*
Item+Pair-level (Exp-5)	**0.533***	0.544*	**0.596***
Item+Pair-level (Rand-5)	0.453*	0.488*	0.572*
Setup [Books]	**P@3**	**P@5**	**P@10**
Item-level	0.253	0.336	0.436
Pair-level (Exp-1)	0.506*	0.528*	0.54*
Pair-level (Exp-3)	0.506*	0.536*	0.58*
Pair-level (Exp-5)	**0.586***	**0.600***	**0.656***
Pair-level (Rand-5)	0.480*	0.504*	0.504
Item+Pair-level (Exp-5)	0.493*	0.456*	0.560*
Item+Pair-level (Rand-5)	0.386*	0.416*	0.516*

ELIXIR: Feedback Densification

★ Label propagation [Zhu and Ghahramani, 2002] on pairs of items to mitigate sparsity $\Rightarrow F_u^d$
- defining pseudo-items, selecting unlabeled pairs, using locality-sensitive hashing (LSH) for efficiency, ...

Zhu and Ghahramani, *Learning from labeled and unlabeled data with label propagation*, CMU Technical Report (2002)

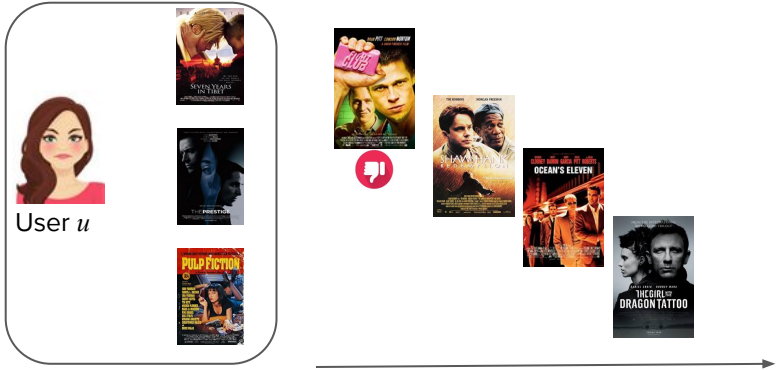
ELIXIR: Feedback Incorporation

The Idea: Learn user-specific parameters $\overrightarrow{w_u}$ of a mapping function g to produce user-specific item representations

$$\min_{\overrightarrow{w_u}} \frac{1}{m} \sum_{v_i, v_j} F_u^d(v_i, v_j) \cdot (sim(\overrightarrow{v_i}, \overrightarrow{v_j}) - sim(g(\overrightarrow{v_i}, \overrightarrow{w_u}), g(\overrightarrow{v_j}, \overrightarrow{w_u}))) + \gamma ||\overrightarrow{w_u}||^2$$

No. feedback pairs

similarity function



Distance of items to user's profile

ELIXIR: Feedback Incorporation

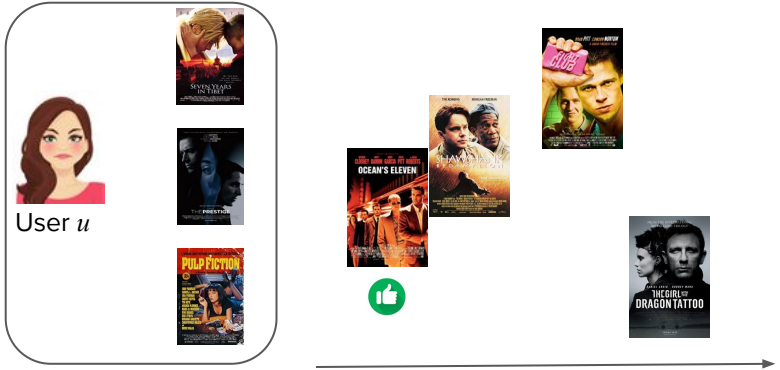
The Idea: Learn user-specific parameters $\overrightarrow{w_u}$ of a mapping function g to produce user-specific item representations

$$\min_{\overrightarrow{w_u}} \frac{1}{m} \sum_{v_i, v_j} F_u^d(v_i, v_j) \cdot (sim(\overrightarrow{v_i}, \overrightarrow{v_j}) - sim(g(\overrightarrow{v_i}, \overrightarrow{w_u}), g(\overrightarrow{v_j}, \overrightarrow{w_u}))) + \gamma ||\overrightarrow{w_u}||^2$$

No. feedback pairs

similarity function

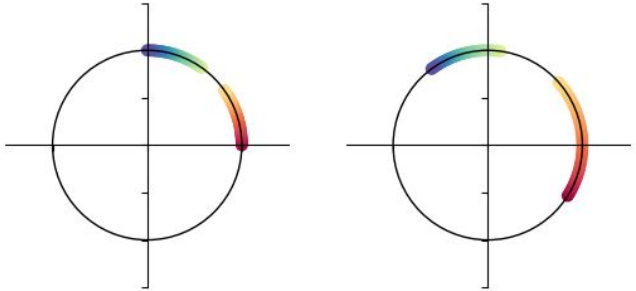
After feedback incorporation



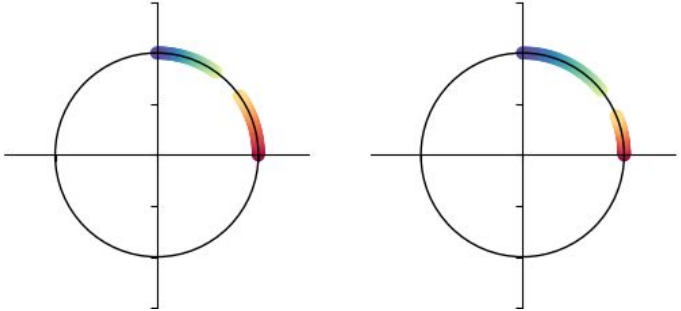
Distance of items to user's profile

Translation vs. Scaling

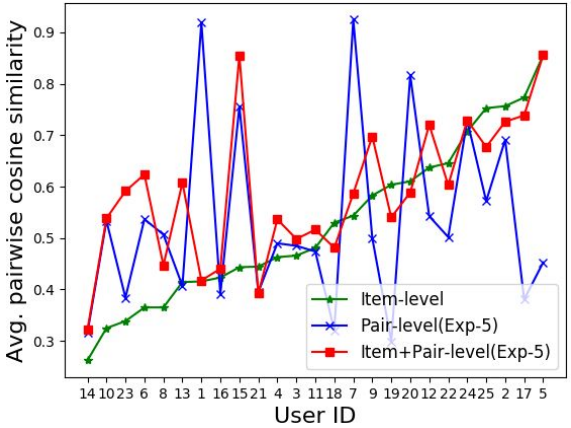
Translation

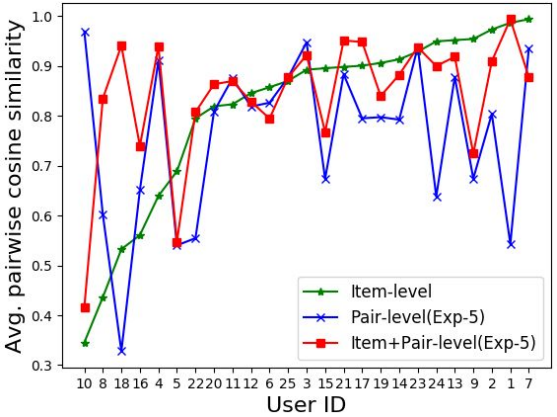


Scaling



Diversity


Avg. pairwise cosine similarity
User ID
Item-level
Pair-level(Exp-5)
Item+Pair-level(Exp-5)


Avg. pairwise cosine similarity
User ID
Item-level
Pair-level(Exp-5)
Item+Pair-level(Exp-5)