Aspect-based active learning for user preference elicitation in recommender systems

María Hernández Rubio (presenting author)
Alejandro Bellogín
Iván Cantador
Aspect-based active learning for user preference elicitation in recommender systems

Recommender systems

![Netflix recommendation interface](image)
Aspect-based active learning for user preference elicitation in recommender systems

User Preferences

Ratings

Intex 86" x 23" Rectangular Frame Above Ground Outdoor Child Safe Splash Swimming Pool

Available from these sellers.

Size:
- 7' 2" x 11" x 23"
- 8' 6" x 5' 2" x 25"
- 10' 8" x 7' 3" x 43"

SUMMERTIME FUN! Escape the summer heat and help keep your little ones cool with the Intex Rectangular Frame Baby Pool. This pool can fit up to 6 onlookers and has a water capacity of 430 gallons of water.

Categorical

Thumbs up / down

Reviews

jwood1964

★★★★★ PERFECT FOR BACK PATIO
Reviewed in the United States on May 3, 2019
Size: 8' 6" x 5' 3" x 25" Verified Purchase

Easy to set up, Instructions were clear, I am seriously thinking about buying another one for Camping! Enough space for water exercise. We added a pump/filter and it creates a perfect back yard oasis. We used foam (for under laminate flooring) as underlay. very comfortable. An adult can sit in the pool and water is up to shoulders. There is a cover sold on amazon which is a perfect fit! Bestway 58105 Frame Pool Cover, 102 by 67-Inch and the Intex filter/pump is Intex Krystal Clear Cartridge Filter Pump for Above Ground Pools, 1000 GPH Pump Flow Rate, 110-120V with GFCI also sold on amazon!
Aspect-based active learning for user preference elicitation in recommender systems

Aspect Information

LG Nexus 5X H791 32GB Carbon Black, Factory Unlocked EU GSM Smartphone, International Model, 1 Year Warranty

Price: $300.00 & FREE Shipping—see free Two-Day Shipping with Amazon Prime

In Stock. Want it Wednesday, Sept. 17? .FREE 2-day delivery within 380 miles and choose One-Day Shipping at checkout. .Sold by Amazon.com and Fulfilled by Amazon.

<table>
<thead>
<tr>
<th>Color</th>
<th>Compatibility Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>Carbon</td>
<td>Fully Compatible</td>
</tr>
<tr>
<td>Black</td>
<td>Fully Compatible</td>
</tr>
</tbody>
</table>

- Fully Compatible

Honest, unbiased, and straight to the point
By Michael on December 4, 2016
Color: Quartz

Verified Purchase

Makes you satisfied and glad you bought it. It's snappy processor can handle most anything you through at it.

Pros- even with plastic backing it still feels premium, the design leaves you breathless, camera is sub par, and it's stock Android which is a great experience.

Cons- lack of as compatibility, no wireless charging, comes with travel charger with an adapter

Overall- definitely a deal and bang for your buck. Highly recommended and after time using it I've fallen in love with no regrets.

CIRCLE2020, July 6-9, 2020, Samatan, Gers, France
User preferences acquisition

Preference elicitation: how to model user’s preferences

Active Learning (AL): ask users to rate items smartly
Aspect-based active learning for user preference elicitation in recommender systems

Our work

Build an **AL algorithm** based on **aspect opinions** extracted from **reviews**.

**Objective:** get similar recommendation metrics with fewer items asked.
Table of contents

- Introduction and Motivation
- Active Learning Methods
  - SoA item-based methods
  - Proposal: aspect based method
- Experiments
  - Datasets
  - Evaluation
  - Results
- Conclusions and Future Work
Table of contents

- Introduction
- Active Learning Methods
  - SoA item-based methods
  - Proposal: aspect based method
- Experiments
  - Datasets
  - Evaluation
  - Results
- Conclusions and Future Work
Aspect-based active learning for user preference elicitation in recommender systems

SoA item-based methods

- **Non-Personalized vs Personalized**
  - Active Learning:
    - Take into account users' previously expressed ratings
    - Request all the users to rate the same items

- **Single- vs combined-heuristics**
  - Single: implements a unique item selection rule
  - Combined: hybridize several single-heuristics strategies

Mehdi et al. TIST (2013)
• **variance**: items with highest rating variance
• **popularity**: items with highest number of ratings
• **entropy**: items with highest rating dispersion
• **log(pop)*entropy**

• **item-item**: items more similar to user’s previously rated items
• **binary-pred**: items with highest probability of being rated by the user
Table of contents

- Introduction

- Active Learning Methods
  - SoA item-based methods
  - Proposal: aspect based method

- Experiments
  - Datasets
  - Evaluation
  - Results

- Conclusions and Future Work
Aspect-based active learning for user preference elicitation in recommender systems

Active Learning Methods

Aspect-based Active Learning method

Exploiting the rich information that can be extracted from reviews: **item aspects** mentioned and the **opinion or sentiment** associated to them.

- Help user to find items that share characteristics with previously interacted items
- Item aspects (vs other content or collaborative information) should alleviate the cold-start problem
Exploiting the rich information that can be extracted from reviews: item aspects mentioned and the opinion or sentiment associated to them.

- Hybrid recommendation approach (Frolov & Oseledets, RecSys 2019): aspect-based item-item similarity matrix plus collaborative information.

- Similarity between item $i_n$ and $i_m$ is computed as the cosine similarity over the item profile $i_n = \{w_{na}\}^{K}_{a=1}$ built on the $K$ aspect opinions, where $w_{na}$ is the weight assigned to aspect $a$ for item $i_n$.

\[
\text{sim}(i_n, i_m) = \frac{\sum w_{na}w_{ma}}{\sqrt{\sum w_{na}^2w_{ma}^2}}
\]

- item-item (personalized and single heuristic)
Table of contents

- Introduction
- Active Learning Methods
  - SoA item-based methods
  - Proposal: aspect based method
- Experiments
  - Dataset
  - Evaluation
  - Results
- Conclusions and Future Work
Aspect-based active learning for user preference elicitation in recommender systems

Experiments

### Dataset

- **Product Dataset:** Movies & TV Amazon product reviews dataset (McAuley, WWW (2016))

- **Aspect method:** vocabulary (voc) (Hernández-Rubio et al. UMUAI (2019))

<table>
<thead>
<tr>
<th></th>
<th>Ratings</th>
<th>Users</th>
<th>Items</th>
<th>Annotations</th>
<th>Aspects</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Initial</strong></td>
<td>1,697,533</td>
<td>123,960</td>
<td>50,052</td>
<td>369,175</td>
<td>23</td>
</tr>
<tr>
<td><strong>Items with aspects</strong></td>
<td>1,683,190</td>
<td>123,960</td>
<td>48,074</td>
<td>369,175</td>
<td>23</td>
</tr>
<tr>
<td><strong>Users with &gt;= 20 ratings</strong></td>
<td>819,148</td>
<td>14,010</td>
<td>47,506</td>
<td>367,750</td>
<td>23</td>
</tr>
</tbody>
</table>

Table 1: dataset and aspects statistics

* for this work we have sample to 1500 users for computational reasons
Evaluation

Methodology:

| Training set 2% | Candidate set (68%) | Test set (30%) |
Evaluation

Methodology:

<table>
<thead>
<tr>
<th>Training set 2%</th>
<th>Candidate set (68%)</th>
<th>Test set (30%)</th>
</tr>
</thead>
</table>

AL algorithm

\[ i_1, i_2, \ldots, i_N \]
Evaluation

Methodology:

<table>
<thead>
<tr>
<th>Training set 2%</th>
<th>Candidate set (68%)</th>
<th>Test set (30%)</th>
</tr>
</thead>
</table>

AL algorithm

\[ [i_1, i_2, \ldots, i_N] \]

metrics

SVD

Experiments

Aspect-based active learning for user preference elicitation in recommender systems

CIRCLE2020, July 6-9, 2020, Samatan, Gers, France
Evaluation

Methodology:

<table>
<thead>
<tr>
<th>Training set 2%</th>
<th>Candidate set (68%)</th>
<th>Test set (30%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>AL algorithm</td>
<td></td>
<td>metrics</td>
</tr>
<tr>
<td>$[i_1, i_2, ... , i_N]$</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

N = 10
iter = 170
CV = 3

Aspect-based active learning for user preference elicitation in recommender systems

Experiments
Evaluation

- Metrics:
  - Rating: MAE, RMSE
  - Ranking: P@1, P@5, P@10

- Baselines:
  - random
  - variance
  - popularity
  - entropy
  - log-pop-entropy
Results

- Aspect-based method is not able to find all known items for the user

Figure 1: Evolution on the number of ratings correctly elicited by each strategy (zoomed in on the first 50 iterations)
Results

- Aspect-based method gets the highest improvement in error.

Figure 2: Evolution on the error accuracy (the lower, the better) under the effect of six elicitation strategies.
Results

- Aspect-based method is the best performing method throughout most of the elicitation process.

![Graph showing ranking accuracy measured as P@5 (the higher, the better) under the effect of six elicitation strategies, smoothed values taking the average of the last 3 points.](image)

**Figure 3:** Ranking accuracy measured as P@5 (the higher, the better) under the effect of six elicitation strategies, smoothed values taking the average of the last 3 points.
## Table of contents

- Introduction
- Active Learning Methods
  - SoA item-based methods
  - Proposal: aspect based method
- Experiments
  - Datasets
  - Evaluation
  - Results
- Conclusions and Future Work
Conclusions

- Novel active learning approach based on opinions about item aspects.
- Tested on a real world dataset
- Outperforms AL strategies on rating prediction error and ranking precision metrics.
Conclusions and Future Work

Future Work

- More exhaustive experiments:
  - more sophisticated aspect extraction methods
  - several recommender systems
  - datasets from several domains

- Analyze the behaviour of our method on different cold-start settings

- Online evaluation with real users to confirm offline results

- Integrate into a conversational agent or chatbot
Questions?

Thank you!