Recommender Systems, Evaluation, and Human-Computer Interaction

Joseph A. Konstan
University of Minnesota
konstan@cs.umn.edu
http://www.grouplens.org

The Right Stuff

Finding lost Shekels …

Measuring “Error”

Yes, we have no Bananas!
It’s Worse than That!

• Cosley et al. 2003 showed that showing predictions leads to skewed ratings!

Seeing Matters

![Graph showing the effect of showing predictions on ratings.](June 27, 2005)
Accuracy Matters

Does Seeing Predictions Affect User Ratings?

RERATE: 212 users rerate 40 movies
- 10 with no shown prediction
- 30 with shown predictions (random order):
  10 accurate, 10 up a star, 10 down a star
- 8% change from showing predictions
- 12% change from altering predictions

UNRATED movies 137 users 1599 ratings
- 8% change from showing predictions
- 12% change from altering predictions
But Metrics Evolve!

Decision Support Metrics
- ROC
- Precision/Recall

User-Centered Studies and Metrics

Where Are We Going?

Novelty and Serendipity
- Did the recommender tell you what you already knew?
- Did it *surprise* and *delight* you?

Confirmability and Building Trust
- How to develop them
Ziegler et. al 2005 reported on the effects of taxonomy-based topic diversification

- Worse recommendations
- Better recommendation lists
- Greater Happiness

Fun!

A system I wish I’d built!!
Systems I use all the time

- Zagat.com (*)
- Launch.yahoo.com (*)
- Ticketmaster.com
- Priceline Hotels

(*) I even pay to subscribe!
Finding User Value

Understanding user needs

- Context
- Task or Activity
- Deep user knowledge

These are not algorithmic ideas – these are human ideas from HCI (and marketing!)

Not all Engineering

Serious Social Science

- Studies on preferences, participation, contribution

- E.g., 5-year study on eliciting participation
Eliciting Participation: An Initial Study

A study of participation in discussions with two factors controlled:
- Similarity of tastes
- Awareness of own uniqueness

Results
- Dissimilarity increased contribution
- Awareness of own uniqueness increased contribution
- Active discussants were not highly-active raters
- Participants rated more than a control group

Motivational Follow-Up

Class projects at CMU used an e-mail campaign to elicit ratings to discover:
- Making users aware of their uniqueness increased rating
- Giving users specific, achievable goals increased rating
- But...
  - Reminding users of their self-benefit or benefit to others actually decreased the number of ratings!
Other Projects

Self-Maintaining Communities
- What happens when you let the masses maintain the database?

Social Preference
- Using economic models to study user behavior

Reinventing Conversation

Goal: More comprehensive engagement
- Full-factorial design around entities, people, comments/discussion entries
<table>
<thead>
<tr>
<th>Name</th>
<th>Message</th>
</tr>
</thead>
<tbody>
<tr>
<td>MaxTarter</td>
<td>I'm curious to hear users' thoughts about our movie recommender system. Does Movielens help you find interesting, good movies to watch? How much do you trust Movielens predictions? Also, how does Movielens compare to other, commercial ventures out there (e.g., Netflix, Yahoo Movies, Amazon) in terms of the quality of recommendations you receive?</td>
</tr>
</tbody>
</table>
| eirwatcher | Hi Max!
My wife and I have rated 263 movies so far. I tell people that ML is dead-on accurate about 85% of the time. The remaining 14% (ML isn't quite perfect) are rarely more than 1/2 star off. In other words, ML is amazingly accurate. We love and use ML frequently. Let's keep up the good work :) |
| Deadmarch | I'm very happy with Movielens :D
Been using it for 3-4 years now (I think). The recommendations are usually right on the money or maybe half a star off.
1632 movies rated so far and counting :P |

---

**movielens**

helping you find the right movies

Welcome konstan@cs.umn.edu (Log Out)
You've rated 99 movies.
You're the 16th visitor in the past hour.

---

Hotel Rwanda (2004)

Your prediction: ★★★★☆
Rate this movie: [Not seen]
Add this movie to my wishlist: [ ]

Movie Information

*Starring:* Don Cheadle, Nick Nolte, Sophie Okonedo, Joaquin Phoenix
*Directed by:* Terry George
*Genre:* Drama, War
*Language:* English

Ratings Information

*Average rating:* ★★★☆☆ (4.2 stars)
*Rated by:* 561 users

External Links

- Internet Movie Database (IMDb)
- Rotten Tomatoes

---

10 Movies with Similar Ratings

1. Defiant Ones, The (1958)
2. Spookhouse (The Vanishing) (1988)
4. Petrified Forest, The (1936)
5. Tale of Two Cities, A (1935)
6. Thing From Another World, The (1951)
7. Kind Hearts and Coronets (1949)
8. Children of Paradise (Les enfants du paradis) (1945)
9. Mr. Deeds Goes to Town (1936)
10. Shoot the Piano Player (Tirez sur le pianiste) (1960)
Not all Engineering

Serious Computer Science
- Properties of different algorithms
- Tunability of algorithms
- Custom construction of recommenders to meet users needs
- E.g., TechLens (demos later)

Results Thus Far

Off-line studies showed promise
On-line study showed different algorithms met different needs
- General positive attitude from users
- Typically one or two useful recommendations in a set of five
  ➔ That’s enough to be useful

Specific value for hybrid content/collab algorithms

Current work: ACM Digital Library
TechLens ACM DL Demo

Please enter an author’s name

Please give us the name of an author whose work you are familiar with, and enter it in
the form it is usually written for a paper. Looking up all the papers a person has
published from our database may take a little while. Please be patient.

Author name: [Georg Lausen]

This research is being conducted by Simulans Research at the University of Minnesota.
As a part of this research, Simulans many monitor and record how you use this recommender
demonstration.

Questions or comments? Please contact us.

Select one of these papers

Maximal convexity by locking
Georg Lausen, Ulrich Priss, and Peter Widmer
[view abstract of this paper]

Taxonomy-driven computation of product recommendations
Can-Ming Lee, Georg Lausen, Lars Schmidt-Thieme
[view abstract of this paper]

F logic: a higher-order language for reasoning about objects,
inheritance, and schemes
Michael Rifkin, Georg Lausen
[view abstract of this paper]

Parallelising Dataflow programs by generalised pivoting
Jurgen Deb, Georg Lausen
[view abstract of this paper]

Referential assertion re logic rules

(C) 2000 John Riedl and Joseph A. Konstan  
December 3, 2000
Directions

Application-Focused Research

- Awareness service
- Paper and proposal-writing support
- Find people (reviewers/experts)
- Overview of a field
Conclusions

- No more frogs telling stories …

- Need for human-centered deeper approaches to developing recommenders that address real needs (including entertainment)

- Need to measure what matters, not what’s easy to measure

Acknowledgements

- The GroupLens research project is supported by grants from the National Science Foundation, and has been supported by grants from Net Perceptions, Inc.

- Many people have contributed ideas, time, and energy to this project.
Acknowledgements

• Thanks to the organizers of this workshop!

Recommender Systems, Evaluation, and Human-Computer Interaction

Joseph A. Konstan
University of Minnesota
konstan@cs.umn.edu
http://www.grouplens.org