open.michigan

Author(s): Rahul Sami, 2009

UNIVERSITY OF MICHIGAN

BY NC SA

si.umich.edu

SCHOOL OF INFORMATION
UNIVERSITY OF MICHIGAN

Citation Key

for more information see: http://open.umich.edu/wiki/CitationPolicy

Use + Share + Adapt

{ Content the copyright holder, author, or law permits you to use, share and adapt. }

PD-GOV **Public Domain – Government**: Works that are produced by the U.S. Government. (USC 17 § 105)

PD-EXP **Public Domain – Expired**: Works that are no longer protected due to an expired copyright term.

PD-SELF **Public Domain – Self Dedicated**: Works that a copyright holder has dedicated to the public domain.

ZERO **Creative Commons – Zero Waiver**

BY **Creative Commons – Attribution License**

BY-SA **Creative Commons – Attribution Share Alike License**

BY-NC **Creative Commons – Attribution Noncommercial License**

BY-NC-SA **Creative Commons – Attribution Noncommercial Share Alike License**

GNU-FDL **GNU – Free Documentation License**

Make Your Own Assessment

{ Content Open.Michigan believes can be used, shared, and adapted because it is ineligible for copyright. }

PD-INEL **Public Domain – Ineligible**: Works that are ineligible for copyright protection in the U.S. (USC 17 § 102(b)) *laws in your jurisdiction may differ

{ Content Open.Michigan has used under a Fair Use determination. }

FAIR USE **Fair Use**: Use of works that is determined to be Fair consistent with the U.S. Copyright Act. (USC 17 § 107) *laws in your jurisdiction may differ

Our determination **DOES NOT** mean that all uses of this 3rd-party content are Fair Uses and we **DO NOT** guarantee that your use of the content is Fair.

To use this content you should **do your own independent analysis** to determine whether or not your use will be Fair.

Lecture 11:
Explanations and Interface Variations

SI583: Recommender Systems

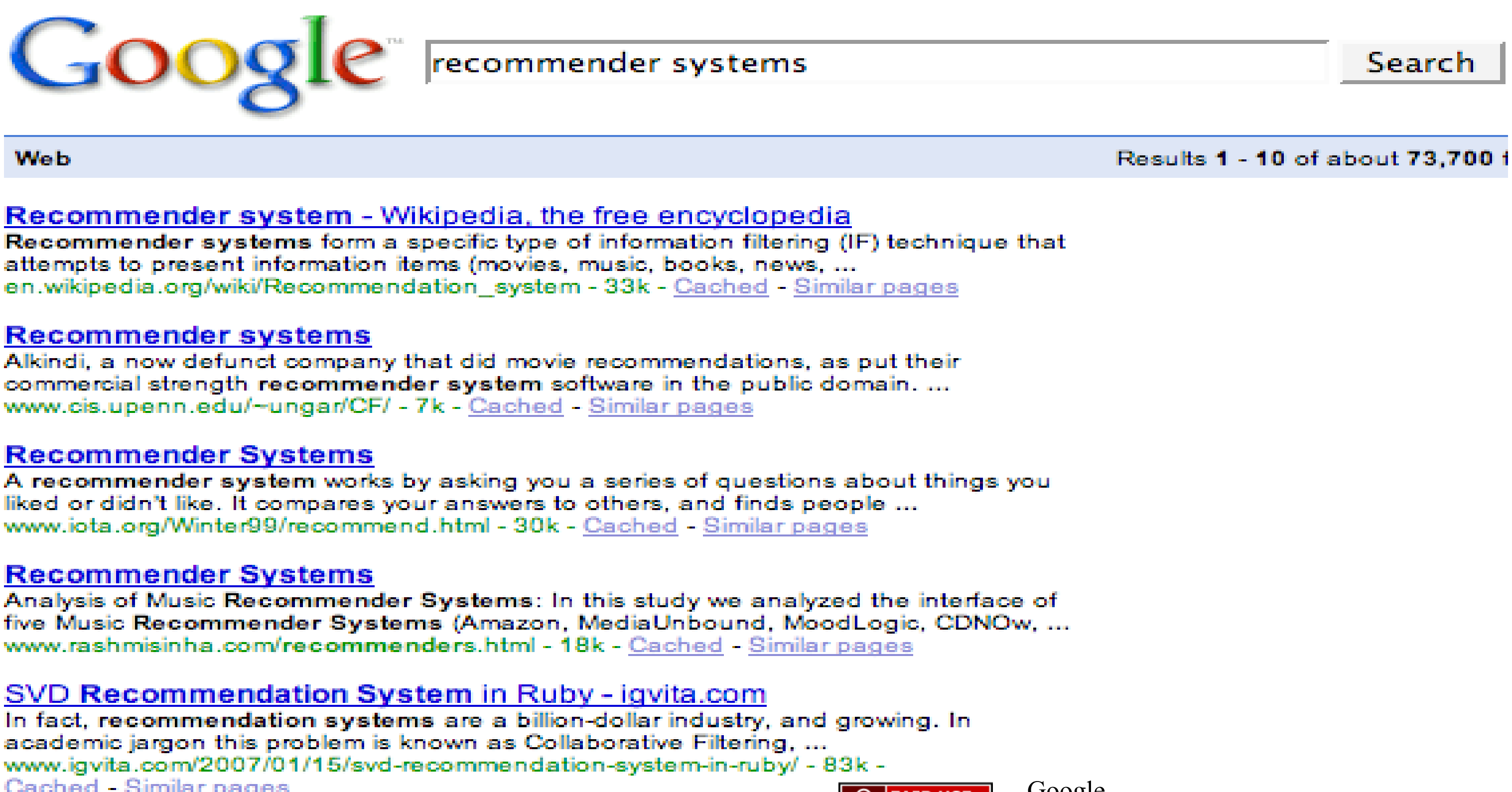
SCHOOL OF INFORMATION
UNIVERSITY OF MICHIGAN

Recap: Evaluation Metrics

- Thresholds
  - – precision, recall, …
- Ranked lists
  - – precision-recall, scores, ..
- Numeric predictions
  - – MAE, RMSE

Are we evaluating the right thing?

- How “good” is this recommender? What factors will you consider?



Rahul's Amazon.com™ › **Recommended for You**
(If you're not Rahul, click here.)

Recommendations by Category

Your Favorites Edit

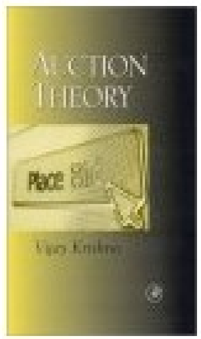
- Books

More Categories

- Apparel & Accessories
- Baby
- Beauty
- Camera & Photo
- Computer & Video Games
- Computers & PC Hardware
- DVD
- Electronics
- Gourmet Food
- Health & Personal Care
- Industrial & Scientific
- Jewelry & Watches

These recommendations are based on items you own and more.

view: **All** | New Releases | Coming Soon   More results

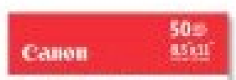
1. 

Auction Theory
by Vijay Krishna
Average Customer Review: ★★★★★
In Stock
Publication Date: March 1, 2002

Our Price: $52.46   **Used & new** from $52.46

Add to cart
Add to Wish List

I Own It   Not interested   x|☆☆☆☆☆ Rate it
Recommended because you purchased Putting Auction Theory to Work and more (edit)

2. 

Canon Matte Photo Paper (8.5x11, 50 Sheets)
by Canon
Average Customer Review: ★★★★☆

Signed by Verisign, I

Why the MAE/RMSE might mislead

- Predictive accuracy doesn't help if it recommends seen items
  - recommenders can get stuck recommending just one small category/cluster
- Users like *diversity* and *serendipity*
- Interface can influence ratings (and thus, measured MSE)
- Trust, confidence important
- Users experience a dialogue/process, not just a *single*, *one-way*, recommendation

Rest of this class

- Impact of interface features on ratings
- Human-Recommender Interaction conceptual model
- Incorporating explanations: why and how

Effect of the interface on ratings

- [Cosley et al, Proceedings of CHI 2003, "Is seeing believing? How recommender Interfaces Affect User Opinions"]

- Studies choices in MovieLens interface:
  - Does the rating scale matter?
  - How consistent are ratings over time? Can recommender prompts affect this?
  - Does the displayed *prediction* affect the submitted *rating*?

- Controlled experiments and survey

Effect of interfaces: Cosley et al findings

- Rating scales:
  - slightly better predictive accuracy with more stars..
  - binary (Like/Dislike) scale results in a positive bias
- Rating consistency
  - Fairly high consistency on rerated movies (60%)
  - Increases when users are prompted with accurate "predicted" value

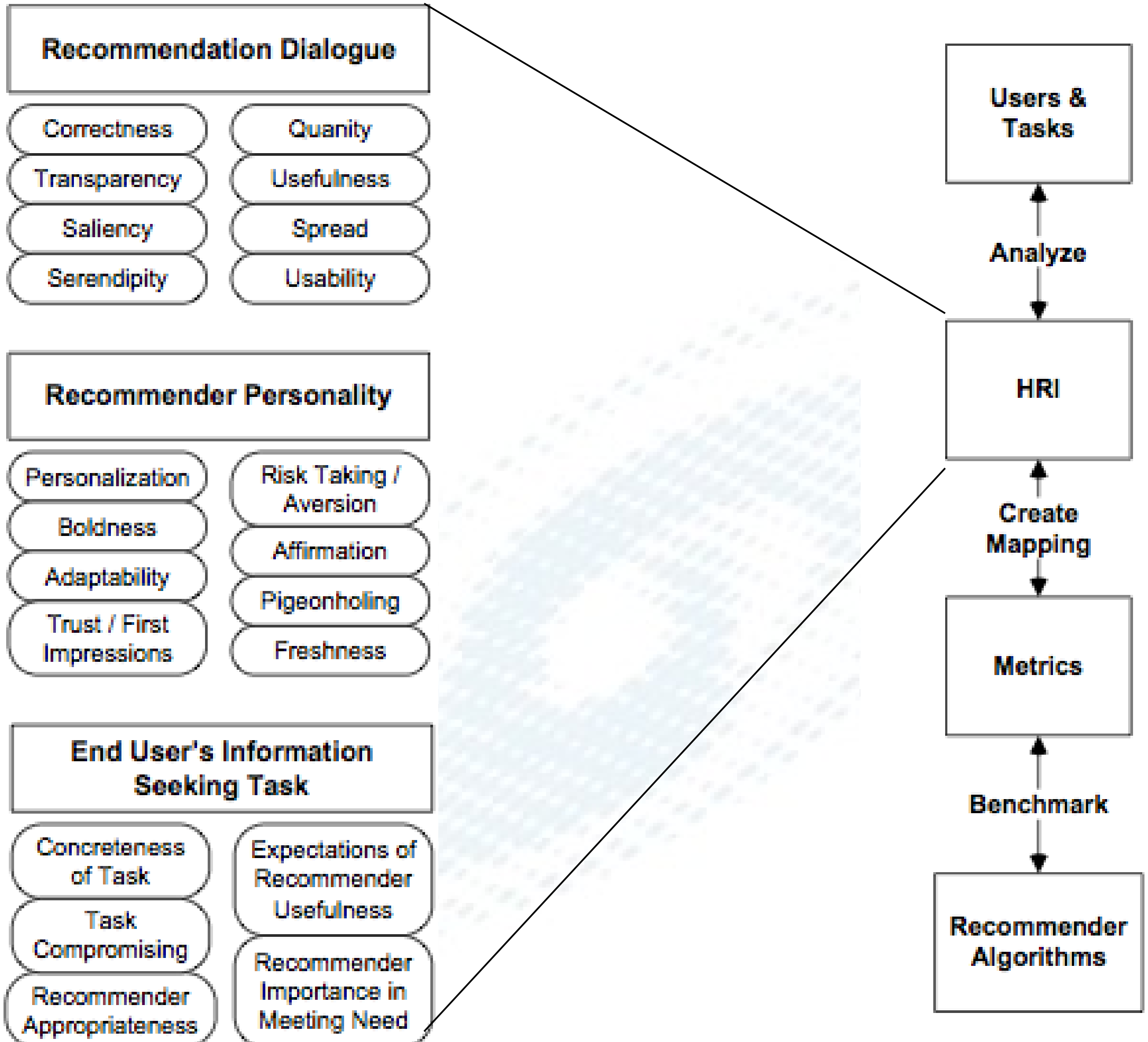
Effect of interfaces: Cosley et al findings

- Effect of displayed predictions:
  - Predictions were randomly perturbed: raised/lowered/left alone
  - Actual ratings were *correlated with the perturbation*
- Implication: Displayed prediction influences users' rating
  - also: manipulation can be (somewhat) self-sustaining

User-centered view

- Consider recommender design *within the context of the users' goals*
- Human-Recommender Interaction model [McNee, Riedl, Konstan]
  - describe/categorize attributes of the context
  - describe attributes/features that influence user satisfaction
  - suggest a design process around these

HRI Model [from McNee et al]



 McNee et al.

HRI model

- Factors describing context
  – concreteness of task
  – expectation of usefulness,etc.

- Different contexts may lead to different evaluation criteria

- Examples?

HRI model

- Factors influencing satisfaction:
  - In one interaction
    - Correctness, usefulness, serendipity (maybe), transparency, diversity of recommended list..
  - Over time
    - Personalization, trust, adaptability, freshness..

Implications

- In studies, users sometime prefer rec. lists that are worse on standard metrics
- Different algorithms better for different goals => recommenders may need multiple CF algorithms
- Interface should provide a way to express context information
- Explaining recommendations can help generate trust, adaptability

Explanations in recommender systems

- Moving away from the black-box oracle model
- *justify* why a certain item is recommended
- maybe also *converse* to reach a recommendation

Why have explanations? [Tintarev & Masthoff]

- Transparency
- "Scrutability": correct errors in learnt preference model
- Trust/Confidence in system
- Effectiveness & efficiency(speed)
- Satisfaction/enjoyment

Example: explanations for transparency and confidence

- “Movie X was recommended to you because it is similar to movie Y, Z that you recently watched”
- “Movie X was recommended to you because you liked other comedies”
- “Other users who bought book X also bought book Y”

Generating explanations

- Essentially, explain the steps of the CF algorithm, picking the most prominent "neighbors"
  - User-user
  - Item-item
- Harder to do for SVD and other abstract model-fitting recommender algorithms

Conversational recommenders

Example transcript: (from [McSherry,“Explanation in Recommender Systems, AI Review 2005]):

- *Top case*: please enter your query
- *User*: Type = wandering, month = aug
- *Top Case*: the target case is “aug, tyrol, ...”
  other competing cases include “....”
- *Top case*: What is the preferred location?
- *User*: why?
- *Top case*: It will help eliminate ... alternatives
- *User*: alps..

Conversational recommenders

- One view: CF using some navigational data as well as ratings
- More structured approach: incremental collaborative filtering
  – similarity metric changes as the query is refined
- e.g., incremental Nearest-Neighbor algorithm [McSherry, AI Review 2005]