open.michigan

Author(s): Rahul Sami, 2009

License: Unless otherwise noted, this material is made available under the terms of the **Creative Commons Attribution Noncommercial Share Alike 3.0 License**: http://creativecommons.org/licenses/by-nc-sa/3.0/

We have reviewed this material in accordance with U.S. Copyright Law **and have tried to maximize your ability to use, share, and adapt it.** The citation key on the following slide provides information about how you may share and adapt this material.

Copyright holders of content included in this material should contact **open.michigan@umich.edu** with any questions, corrections, or clarification regarding the use of content.

For more information about **how to cite** these materials visit http://open.umich.edu/education/about/terms-of-use.

UNIVERSITY OF MICHIGAN

BY NC SA

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.

Implicit ratings

SI583: Recommender Systems

Recap: Eliciting Contribution of Ratings/Feedback

Learning objective:

- Learn how motivating factors are evaluated, what factors have been found to influence people's contribution, and the design implications of these results.

- Two sets of studies:
  - Slashdot commenting
  - MovieLens research on movie rating contribution

This class: Use of Implicit Ratings/Information

- Learning goal:
  - ways in which implicit information has been used
  - a framework to think about different categories of information [Oard&Kim]
  - high-level operation of recommenders using implicit information

Why use implicit information?

- Implicit information: information about users' preferences and/or item qualities that are inferred by monitoring user's activities
- .. *not* derived by asking user how much she liked an item

- advantages and disadvantages?

Useful information in Netnews[Morita and Shinoda]

- Idea: read times of news articles may reflect preference/quality

- Methodology:
  - volunteers use modified software to record read times
  - Later asked to rate articles
  - Check how closely they are correlated

Useful Information in NetNews

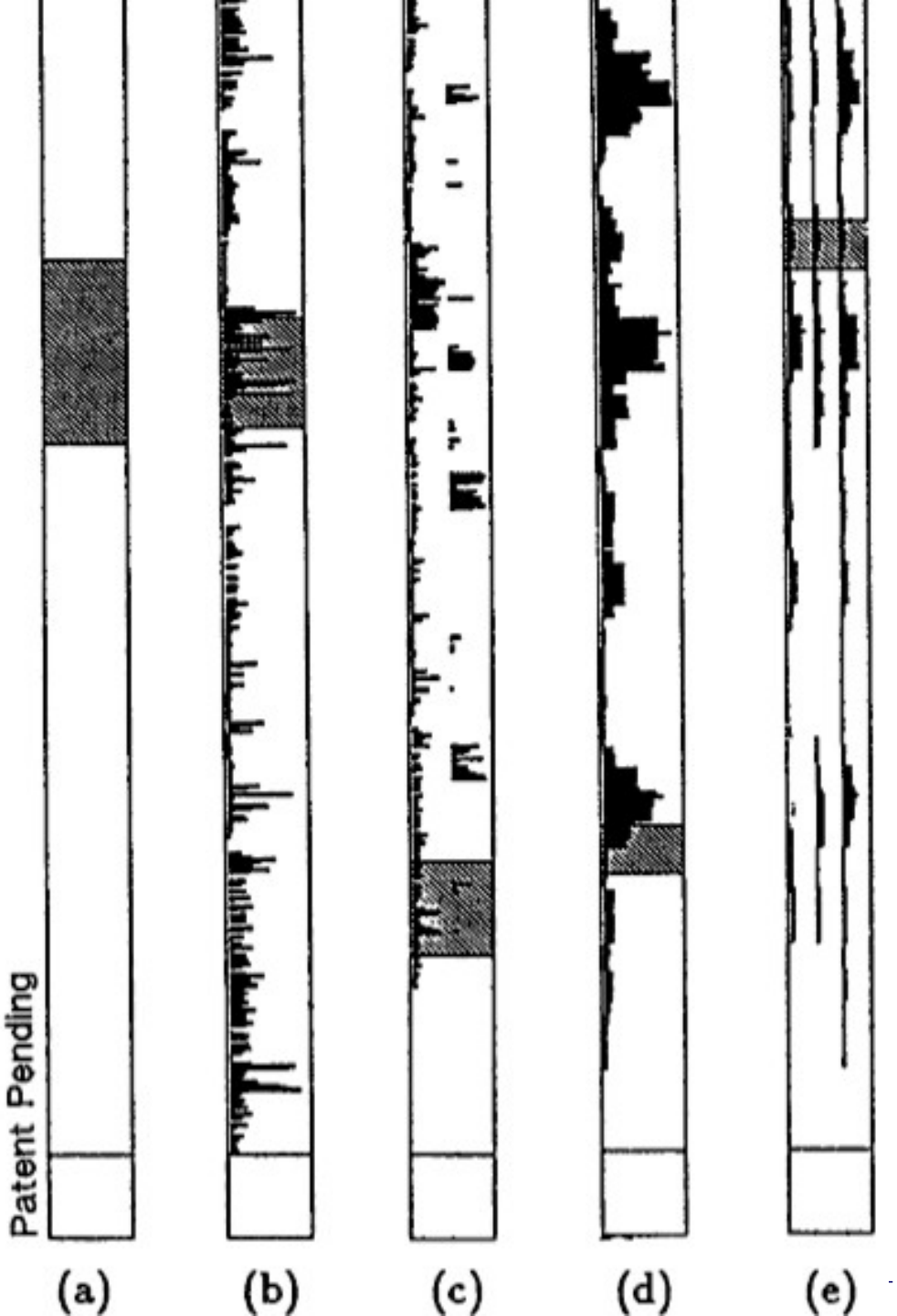
- Result: Read times are very highly correlated with stated preference

- Later confirmation by [Konstan et al, CACM '97]
  - Recommenders built to use read time information are almost as accurate as recommenders using reported preferences.

Edit Wear and Read Wear [Hill et al]

- Augmenting an editor to
  - track read times for each segment of text
  - show where the most frequently edited/read pieces of information are.

Wear Indicators: Scroll Bar



PD-INEL Source: Undetermined

Edit and Read Wear: Insights

- Innovative interface for guiding users towards interesting content segments
- Physical media have built in “behavior-based recommenders”
  - goal: reconstruct this in the digital environment
- A slightly different design goal: summarization rather than recommendation

Applications and information examples

A systematic framework[Oard and Kim]

- Types of observable information categorized along two dimensions
  - Purpose
  - Scope/granularity
- More structured than the “implicit/explicit” categorization

Framework [Oard and Kim]

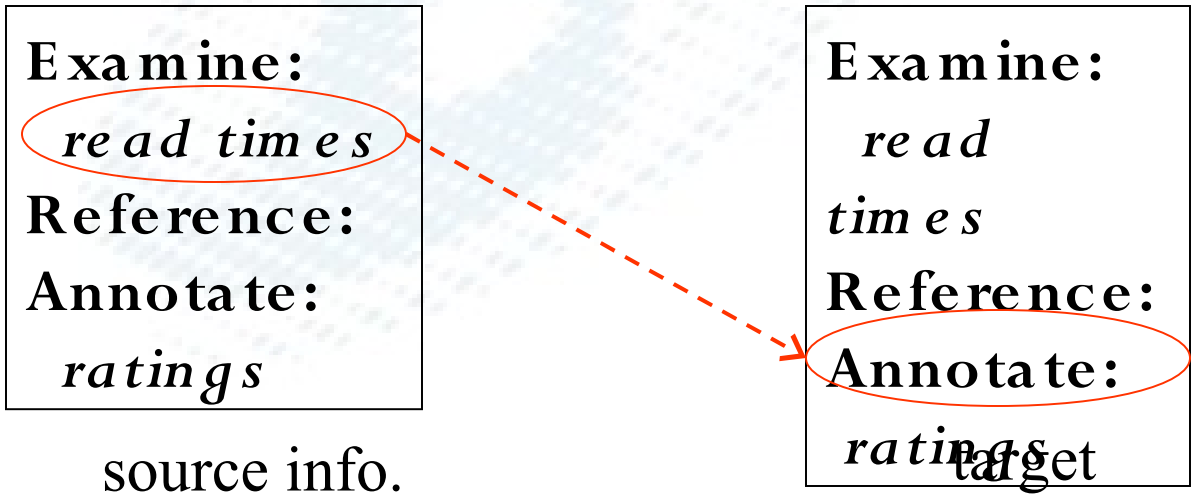
Behavior Category	Minimum Scope: Segment	Minimum Scope: Object	Minimum Scope: Class
Examine	View Listen	Select	
Retain	Print	Bookmark Save Delete Purchase	Subscribe
Reference	Copy-and-paste Quote	Forward Reply Link Cite	
Annotate	Mark up	Rate Publish	Organize

Table 2. Potentially observable behaviors.

PD-INEL Source: Oard and Kim

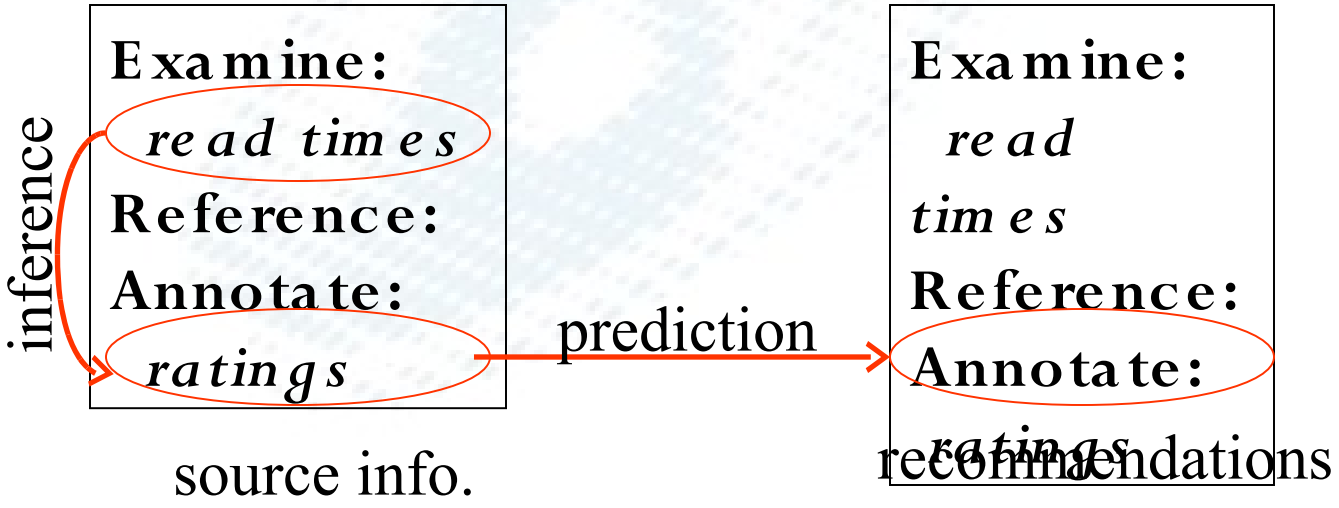
Prediction and Inference

- Recommendations are based on predictions of particular behaviors, which can also be categorized in this way
- May be in a different category from the information used to make recommendations



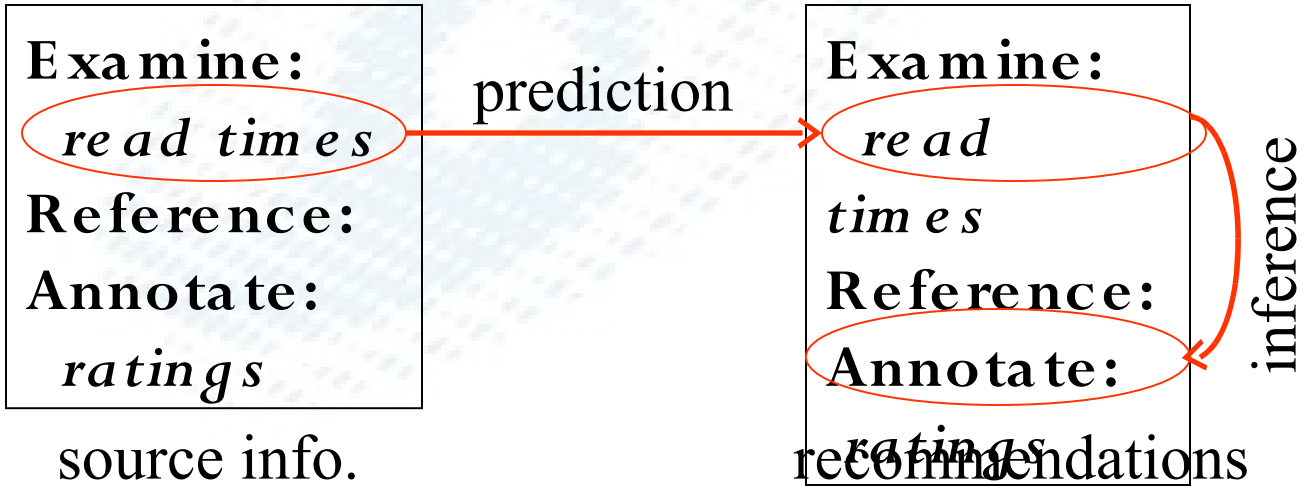
Prediction and inference

- Can use
  - statistical inference between categories
    - "long read times tend to be correlated with high ratings"
  - collaborative filtering prediction algorithms



Prediction and inference

- Alternative: flip the order
  - collaborative filtering prediction algorithms
  - statistical inference between categories
    - "long read times tend to be correlated with high ratings"



Advantages of each method

- Prediction-inference allows for common prediction, followed by personalized inference
  - e.g., Slashdot score is a "prediction" of average reader's rating
  - A user who liked comments the average reader found bad could adjust the inference made
- Inference-prediction may require less communication
  - assuming many observations go into one inference

Implicit information

- easier to obtain lots of data
- can't choose format, so requires good inference procedures
- can be built around collaborative filtering algorithms for explicit ratings