Author(s): Rahul Sami, 2009

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.

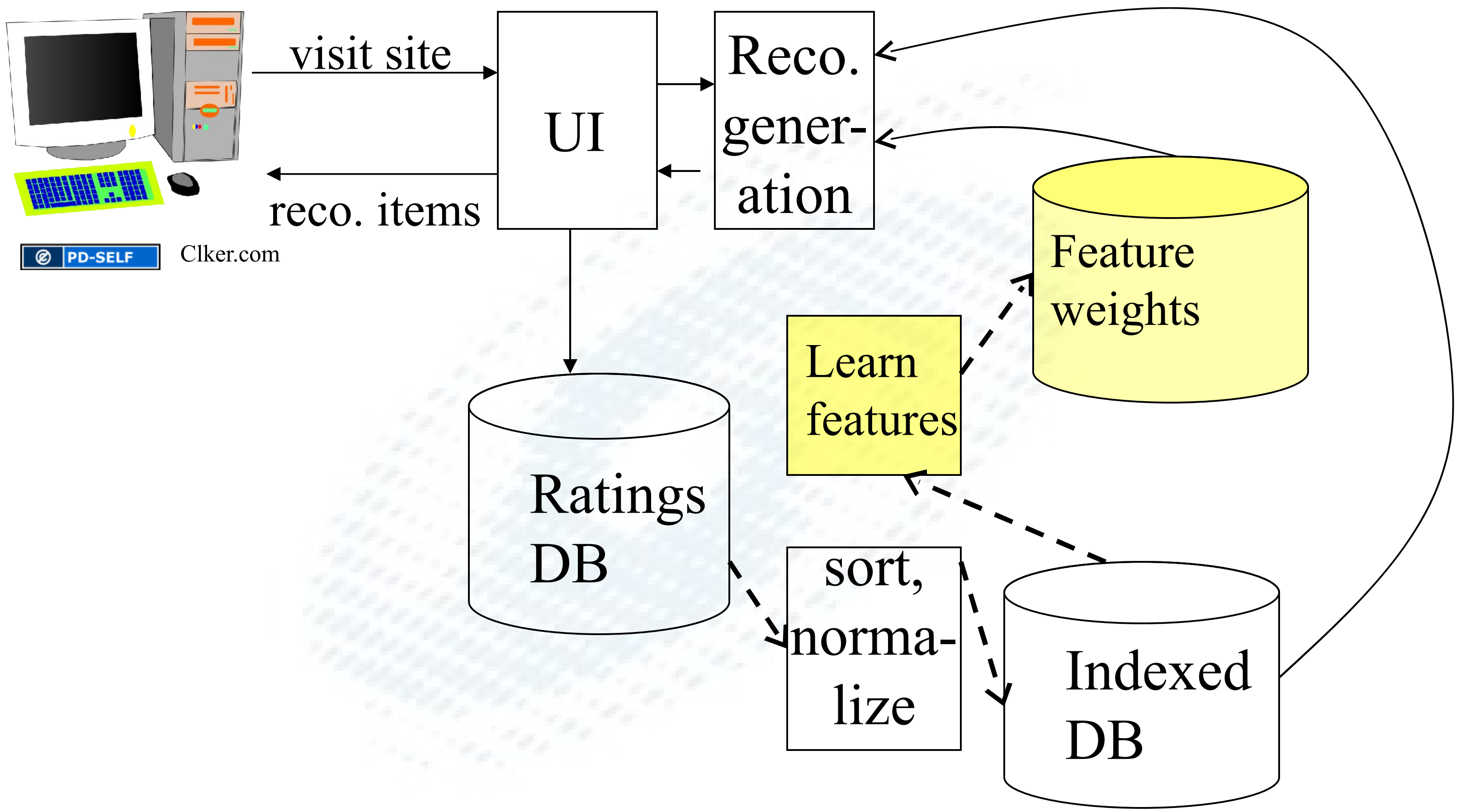
To use this content you should **do your own independent analysis** to determine whether or not your use will be Fair.

Lecture 10:
Singular Value Decomposition;
Evaluation Metrics

SI583: Recommender Systems

SCHOOL OF INFORMATION
UNIVERSITY OF MICHIGAN

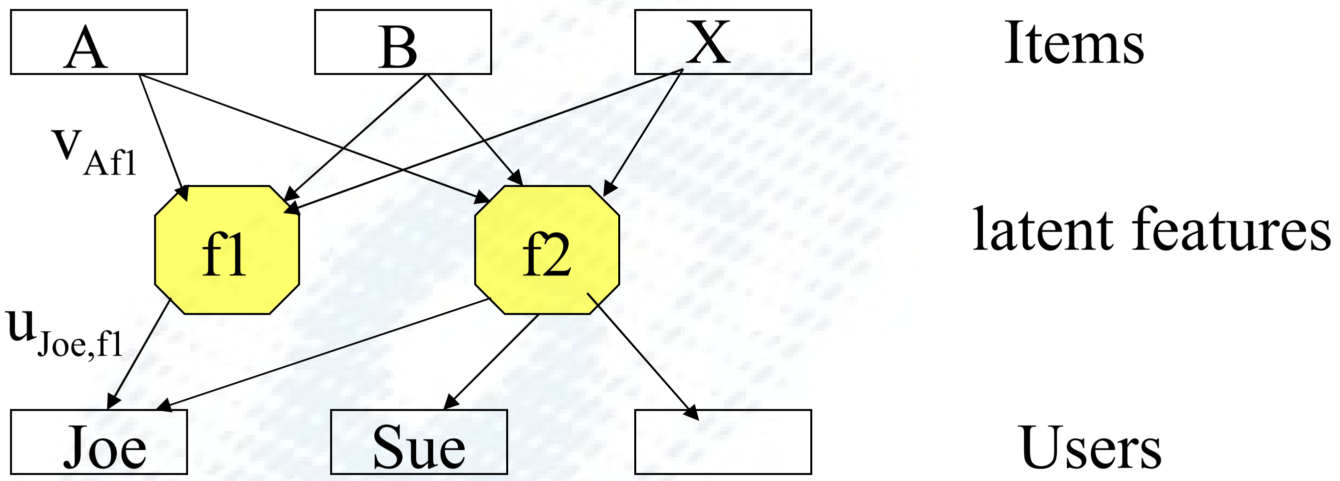
Software modules



SCHOOL OF INFORMATION
UNIVERSITY OF MICHIGAN

Fitting the weights: SVD

- Model weights from SVD (U,S,V):



- Weight (item j, feature f) = $\sqrt{s_{ff}}\, V_{fj}$
- Weight (user i, feature f) = $\sqrt{s_{ff}}\, U_{if}$

Alternative: get software package to calculate weights directly..

SVD-based CF: Summary

- Pick a number of features *k*
- Normalize ratings
- Use SVD to find best fit with *k* features
- Use fitted model to predict value of Joe's normalized rating for item X
- Denormalize (add Joe's mean) to predict Joe's rating for X

SVD Practicalities

- SVD is a common mathematical operation; numerous libraries exist
- Efficient algorithms to compute SVD for the typical case of sparse ratings
- A fast, simple implementation of an SVD-based recommender (by Simon Funk/Brandyn Webb) was shown to do very well on the Netflix challenge

SVD and Content Filtering

- Similar idea: Latent Semantic Indexing used in content-filtering
  - Fit item descriptions and keywords by a set of features
  - Related words map onto the same feature
  - Similar items have the similar feature vectors
- Useful to combine content+collaborative filtering
  - Learn some features from content, some from ratings

Where we are in the course

- Up to this point:
  - Eliciting ratings
  - Using implicit information
  - Software architecture
  - Collaborative filtering algorithms
- Next:
  - Evaluation
  - Scalable software (briefly)
  - Interface extensions
  - Manipulation and defenses
  - Privacy

Evaluation of Recommendation Quality

Recommendation Presentation

- Predicted score
- (Ordered) list of recommended items
- Filter threshold based on score

Fast! (Score:5, Funny)

by bablefisk (115988) on Tuesday February 10, @06:02AM (#26795395)

November 2007 was a bit optimistic, but september 2008 is still a really fast fix!

Reply to This

That's more than just a typo... (Score:5, Funny)

by Arancaytar (966377) <arancaytar.ilyaran@gmail.com> on Tuesday February 10, @06:06AM (#26795419) Hom

That entire news item is outdated. :P

Reply to This

Confusion about Dates (Score:2, Insightful) by Zephiris (788562) The article [bbc.co.uk] apparently fails to

- Re: (Score:2) by harry666t (1062422) ...and you should also always specify whether it's AD or BC, whe
  - Re: (Score:2) by Gandalf_Greyhame (44144) Of course it's AD. You don't have to say it's AD. It's

Assessing Quality of a Threshold

- Many metrics derived from the “confusion matrix”:

		actual value		
		p	*n*	total
prediction outcome	*p'*	True Positive	False Positive	P'
	n'	False Negative	True Negative	N'
	total	P	N	

Assessing Quality of a Threshold

- Precision *p*
  TP/(TP+FP)
- Recall *r*
  TP/(TP+FN)

		actual value		
		p	*n*	total
prediction outcome	*p'*	True Positive	False Positive	P'
	n'	False Negative	True Negative	N'
	total	P	N	

Assessing Quality of a Threshold

- Precision *p*
  TP/(TP+FP)
- Recall *r*
  TP/(TP+FN)
- Combinations, e.g., 2pr/(p+r) {F1-measure}

- Which metric is best?

		actual value		
		p	*n*	total
prediction outcome	*p'*	True Positive	False Positive	P'
	n'	False Negative	True Negative	N'
	total	P	N	

Assessing Quality of a Threshold

- Precision *p*
  TP/(TP+FP)
- Recall *r*
  TP/(TP+FN)
- Combinations, e.g.,
  2pr/(p+r)   {F1-measure}

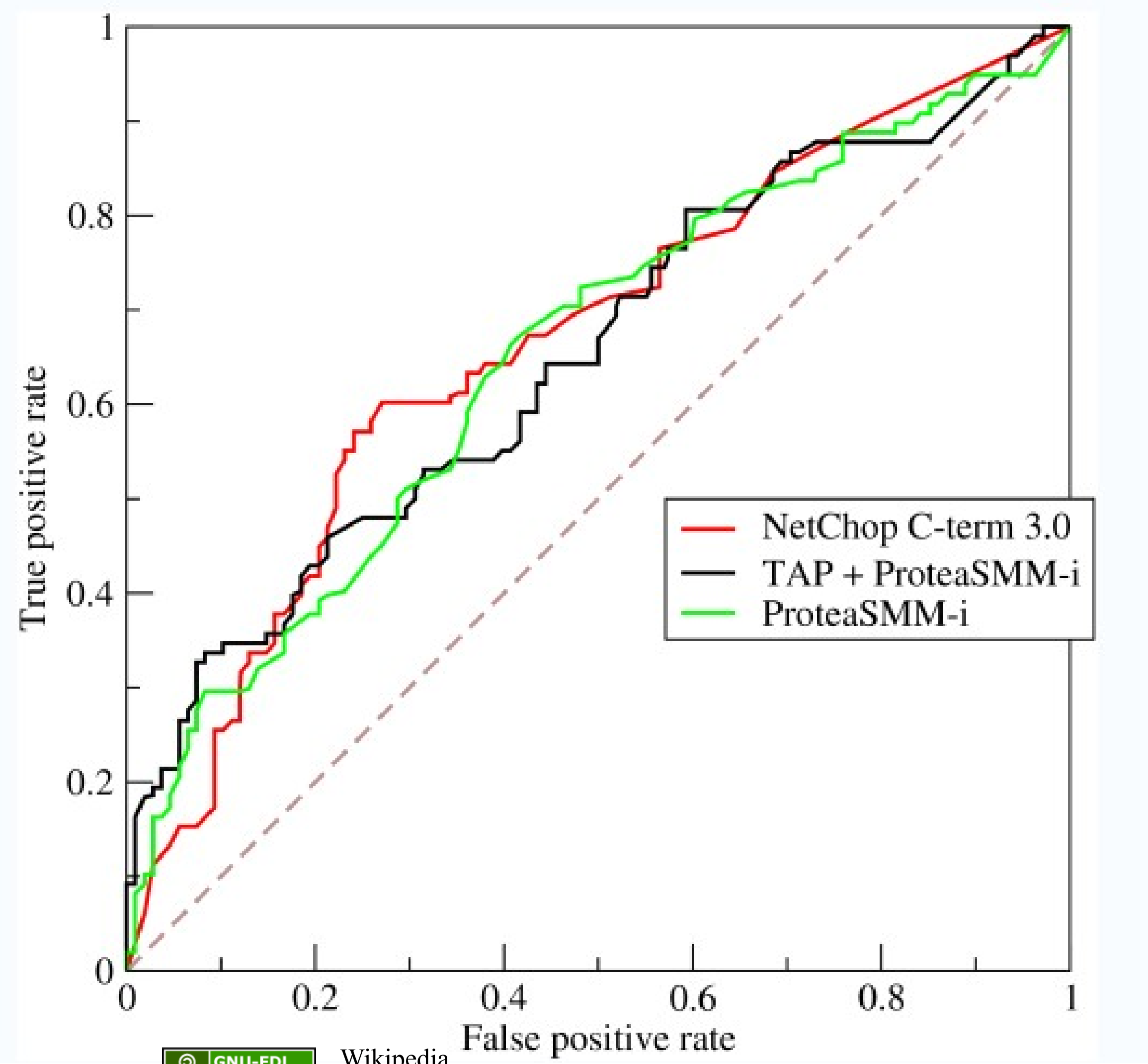
		actual value		
		p	*n*	total
prediction outcome	*p'*	True Positive	False Positive	P'
	n'	False Negative	True Negative	N'
	total	P	N	

- Which metric is best?
- Depends on scenario..
- ultimately, all are special cases of cost-benefit analysis
  - cost of inspecting an item
  - benefit from seeing a good item
  - (perhaps) penalty for missing a good item

Assessing Quality of a Threshold

- Other charts you might see:
  - ROC (receiver operator characteristic) curve
  - precision-recall curve
  - both are different ways of showing how the tradeoff changes with the threshold

Example ROC curve



school of information | Search | Advanced Search | Preferences

Web

[**School of Information** - University of Michigan: The iSchool at ...]
The **School of Information** (SI) at the University of Michigan educates professionals to lead in the **information** age, offering the PhD and master's degree in ...
www.si.umich.edu/ - 27k - Cached - Similar pages

[**School of Information**]
UC Berkeley **School of Information** ... Workshop Unites Ph.D. Students From UC **Information** Schools. First-ever student-run workshop brings together doctoral ...
www.**ischool**.berkeley.edu/ - 34k - Cached - Similar pages

[**School of Information** - University of Texas]
The mission of The University of Texas at Austin **School of Information** is to shape **information** realities for human and social benefit by: 1) Discovering new ...
www.**ischool**.utexas.edu/ - 32k - Cached - Similar pages

[Great Schools - Public and Private **School** Ratings, Reviews and ...]
School information by state (click to expand)**School information** by city (click to expand)**School information** by district (click to expand) ...
www.great**schools**.net/ - 35k - Cached - Similar pages

Assessing quality of a list

- On/off correctness; see previous slide
- Number of swaps necessary to get correct ordering
- Is there anything good on the list?
- Some scoring/point function
  - E.g. 10 points if top choice on the list, etc..

Rating predictions

FAIR USE

www.gizmodo.com

Assessing quality of score predictions

- Mean Absolute Error

$$\frac{\sum |pred - actual|}{N}$$

Assessing quality of score predictions

- Mean Absolute Error

$$\frac{\left| pred - actual \right|}{N}$$

- Mean Squared Error

$$\frac{\text{¥}\left(pred - actual \right)^2}{N}$$

Choice of error metric

- Why did Netflix choose MSE instead of MAE?
- What other metrics could they have used, and what impact would they have had?

Minimizing MAE and MSE

- Given beliefs, probability distribution over ratings
  - E.g., 0, 4, or 5, each with probability 1/3
- What should you predict in order to minimize MAE?
- What should you predict in order to minimize MSE?