

#### Recommendation Systems

- Predict items a user may be interested in based on information about the user and the items
- An effective way to help people cope with information overload
- Examples: Amazon, Netflix, Tivo, …

# **Problem Definitions**

#### Prediction

- $_{\rm n}$  Given: a user and  ${\bf k}$  items
- n Return: predicted rating for each item
- Recommendation
  - n Given: a user
  - $_{\rm n}$  Return:  ${\bf k}$  items from the database with the highest predicted rating

#### Categories of Filtering Techniques

- Cognitive (content-based) filtering
- Economic filtering
- Collaborative (social) filtering
  Rate items based on the evaluation of other users

#### Collaborative vs. Cognitive

- Support for filtering items whose content is not easily analyzed by automated process
- The ability to filter items based on quality and taste
- The ability to provide serendipitous reocmmendations

# CF – Basic Assumptions

- Items are evaluated by users explicitly or implicitly
  - n Ratings, reviews
  - n Purchases, browsing behaviors
  - · ...
- We may map explicit and implicit evaluations to a rating scale, e.g. 1-5.

So shall we simply recommend the items with the highest average rating??

#### CF – Heuristic

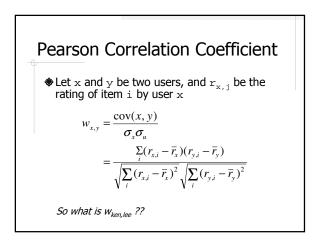
People who agreed in the past are likely to agree in the future

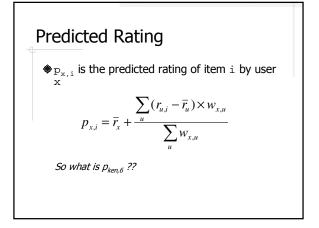
#### **Problem Formulation**

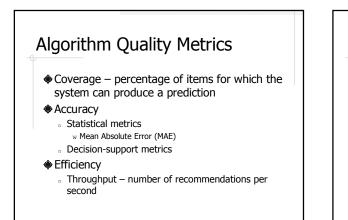
User-Item Matrix

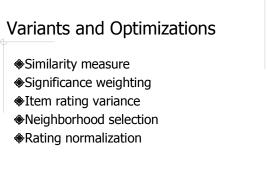
Item	Ken	Lee	Meg	Nan
1	1	4	2	2
2	5	2	4	4
3			3	
4	2	5		5
5	4	1		1
6	??	2	5	

So what would be Ken's rating for Item 6??









# Similarity Measures

- Pearson Correlation
- Spearman Correlation
- Cosine similarity
- Entropy
- Mean-squared-difference
- ۰...

#### Significance Weighting

Weight users in additional to the similarity measure

#### **Item Rating Variance**

Some items are more telling about tastes than others

### Neighborhood Selection

Select a subset of the users for better performance *and accuracy*.

#### Rating Normalization

- No normalization
- Deviation from mean

Z-score

# Other Recommendation Algorithms

- Combine collaborative and contentbased filtering
- Item-item collaborative filtering
- Bayesian networks

#### **CF** Libraries

 Taste – <u>http://taste.sourceforge.net/</u>
 COFE – http://eecs.oregonstate.edu/iis/CoFE/

### Non-personalized Recommendation

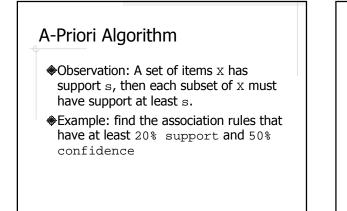
What if the user is new to the site?
What if the site itself is new, i.e. no previous user transactions?



# Association Rule Mining

j

Support: the number of transactions that have  $\{i_1, i_2, ..., i_n\}$  and j



# Item Similarity under Vector-Space Model

- Each unique term is a dimension
- Each document is a vector
- Similarity
  - n Euclidean distance
  - n Cosine similarity measure

#### References

- GroupLens: An Open Architecture for Collaborative Filtering of Netnews by P. Resnick et. al, 1994.
- An Algorithmic Framework for Performing Collaborative Filtering by J. Herlocker et. Al, 1999.
- *E-Commerce Recommendation Applications* by J. B. Schafer et. al, 2001.