User-User Collaborative Filtering

Reference for UUCF

- An Algorithmic Framwork for Collaborative Filtering by Herlocker, Konstan, Borchers, and Riedl (Proc. SIGIR 1999)

The following are screenshots the handwritten notes

$$S(u,i) = \frac{\sum_{v \in U} r_{vi}}{|U|}$$

rating

all users

\# users

user item

pred. score

$$S(u,i) = \bar{r}_u + \frac{\sum_{v \in U} (r_{vi} - \bar{r}_v)}{|U|}$$

$$S(u,i) = \frac{\sum_{v \in U} r_{vi} \cdot w_{uv}}{\sum_{v \in U} w_{uv}}$$

similarity

$$S(u,i) = \bar{r}_u + \frac{\sum_{v \in U} (r_{vi} - \bar{r}_v) * w_{uv}}{\sum_{v \in U} w_{uv}}$$

$v \neq u$

limit size

min similarity

Neg.

Pearson Correl

$$w_{uv} = \frac{\sum_{i \in I} (r_{ui} - \bar{r}_u)(r_{vi} - \bar{r}_v)}{\sigma_u \sigma_v}$$

- small overlap
- unary

$$\frac{\sum_{v \in V}}{\sum_{v \in V}}$$

Common Characteristics

- Collection of Ratings
- Measure of Inter-User Agreement
  – Correlation, Vector Cosine
- Personalized Recommendations/Predictions
  – Weighted Combinations of Others' Ratings
- Tweaks to make things work right …
  – Neighborhood limitations
  – Normalization
  – Dealing with limited co-ratings

Let's Formalize This …

- Given a set of items *I*, and a set of users *U*, and a sparse matrix of ratings *R*,
  We compute the prediction s(u,i) as follows:
  - For all users v ≠ u, compute w_{uv}
    - similarity metric (e.g., Pearson correlation)
  - Select a neighborhood of users $V \subset U$ with highest w_{uv}
    - may limit neighborhood to top-k neighbors
    - may limit neighborhood to sim > sim_threshold
    - may use sim or |sim| (risks of negative correlations)
    - may limit neighborhood to people who rated i (if single-use)
  - Compute prediction:

$$s(u,i) = \bar{r}_u + \frac{\sum_{v \in V}(r_{vi} - \bar{r}_v) * w_{uv}}{\sum_{v \in V} w_{uv}}$$

Implementation Issues

- Given m=|U| users and n=|I| items:
  - Computation can be a Bottleneck
    - Correlation between two users is O(n)
    - All correlations for a user is O(mn)
    - All pairwise correlations is O(m^2n)
    - Recommendations at least O(mn)
  - Lots of ways to make more practical
    - More persistent neighborhoods (m->k)
    - Cached or incremental correlations

Core Assumptions/Limitations

- Why does this work?
  - Let's break it down …
- Assumption: Our past agreement predicts our future agreement
  - Base Assumption #1: Our tastes are either individually stable or move in sync with each other
  - Base Assumption #2: Our system is scoped within a domain of agreeement

User-User Collaborative Filtering