4-1: Introduction to User-User Collaborative Filtering

Learning Objectives

- To understand the intuition and history of the user-user collaborative filtering algorithm
- To review the basic ideas and assumptions (and therefore limitations) behind the algorithm.

Historical Reflection …

- 1992: Tapestry and seeds of ACF
- 1994-1995 Early Automated CF Systems
  - GroupLens
  - Ringo/HOMR
  - Video Recommender

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

Implementation Issues

- Given m users and n 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

So What Happened?

- GroupLens -> Net Perceptions -> GroupLens
- RINGO -> Agents Inc. -> Firefly Networks
- Industry Accceptance of ACF
  - Pressure to innovate more efficient algorithms

Moving Forward

- Next Lectures
  - Breaking down the core algorithm
  - Tuning and tweaks: normalization, neighborhood size, and more
- Later this Module
  - Explanations, trust, reptuation
  - Programming user-user CF

4-1: Introduction to User-User Collaborative Filtering