RECOMMENDER SYSTEMS

- A specific type of information filtering system technique that attempts to recommend information items (movies, TV programs, music, books, research papers, etc) that are likely to be of interest to the user.
- Many online stores provide recommendations (For example: Amazon, CD now, etc).
- Recommenders have been shown to substantially increase sales at online stores.
- Typically a recommender system compares a user profile to some reference characteristics and seeks to predict the rating or preference that a user would give to an item they had yet not considered.
- User's profile can be built using explicit data collection and implicit data collection.
- **Explicit data collection includes:**
    1. Asking a user to rate an item.
    2. Asking a user to rank a collection of items from favorite to least favorite.
    3. Presenting two items to a user and asking him/her to choose the better one of them.
    4. Asking a user to create a list of items that he/she likes.
- **Implicit data collection includes:**
    1. Observing the items that a user views in an online store.
    2. Analyzing viewed items.
    3. Keeping a record of the items that a user purchases online.
    4. Obtaining a list of items that a user has listened or watched on their computer.
    5. Analyzing the user's social network and discovering similar likes and dislikes.
- Recommender systems are useful alternative to search algorithms, since they help users discover items they might not have found by themselves.
- There are two basic approaches to recommending.
- They are:
    - Collaborative Filtering
    - Content Based Recommending

PERSONALIZATION

- Recommenders are instances of personalization software.
- Personalization concerns adapting to the individual needs, interests and preferences of each user.
- It includes:

1. Recommending
2. Filtering
3. Predicting

COLLABORATIVE FILTERING

- Maintain a database of many users' ratings of a variety of items.
- For a given user, find other similar users whose ratings strongly correlate with the current user.
- Recommend items rated highly by the similar users, but not rated by the current user.
- Collaborative filtering methods are based on collecting and analyzing a large amount of information on user's behavior.
- Types:
  - User based collaborative filtering
  - Items to item collaborative filtering

1. User based collaborative filtering

- Look for users who share the same rating patterns with the active user (the user who the prediction is for).
- Use the ratings from those likeminded users found in above step to calculate the prediction for the active user.

2. Items to item collaborative filtering

- People who buy X also buy Y.
- Association rule.

User 6

- Typically Pearson correlation coefficient is used between ratings for active user, 'a' and another user, 'u'.
- $\mathbf{C_{a,u}} = \frac{covar(r_a, r_u)}{\sigma_{ra}\sigma_{ru}}$

- r_a and r_u are the rating vectors for the m items rated by both a and u.
- $r_{i,j}$ is users i's rating for item j.

i.e. $$covar(r_a, r_u) = \frac{\sum_{i=1}^{m} (r_{a,i} - \overline{r_a})(r_{u,i} - \overline{r_u})}{m}$$

$$\overline{r_x} = \frac{\sum_{i=1}^{m} r_{x,i}}{m}$$

$$\sigma_{ra} = \sqrt{\frac{\sum_{i=1}^{m} (r_{x,i} - \overline{r}_x)^2}{m}}$$

PROBLEMS WITH COLLABORATIVE FILTERING

1. Cold start:

   There needs to be enough other users already in the system to find a match.

2. Sparsing:

   If there are many items to be recommended, even if there are many users, the user/rating matrix is spare and it is hard to find users that have rated the same items.

3. First rater:

   It cannot recommend on item that has not previously rated.

4. Popularity basis:

   It cannot recommend items to someone with unique taste.

SIGNIFICANCE WEIGHTING

$W_{a,u} = S_{a,u} C_{a,u}$

$S_{a,u}$ = 1, if m>50

m, if m<=50

RATING PREDICTION

$$P_{a,i} = r_a + \frac{\sum_{u=1}^{n} W_{a,u}(r_{u,i} - r_u)}{\sum_{u=1}^{n} W_{a,u}}$$

CONTENT BASED RECOMMENDING

- Recommendations are based on information on the content of items rather than on other user's opinions.
- Content based filtering methods are based on the information about the items that are going to be recommended.
- Try to recommend the items similar to those that a user liked in the past.
- In particular, various candidate items are compared with items previously rated by the user and the best matching items are recommended.
- Some previous applications are:
    1. News weeder (1995)
    2. Syskill and Webert (1996)
- Example of content based recommending: LIBRA (Learning Intelligent Book Recommending Agent).
- **Advantages**
    - No need for data on other users. (No cold start or sparsity problems.)
    - Able to recommend to users with unique tasks.
    - Able to recommend new and unpopular items. (No first rater problem.)
    - Can provide explanations of recommended items by listing content features that caused an item to be recommended.
- **Disadvantages**
    - Requires content that can be encoded as meaningful features.
    - User's taste must be represented as a learnable function of these content features.
    - Unable to exploit quality judgments of other users.