Collaborative Filtering

Collaborative Filtering (CF)

- The most prominent approach to generate recommendations
  - used by large, commercial e-commerce sites
  - well-understood, various algorithms and variations exist
  - applicable in many domains (book, movies, DVDs, ..)
- Approach
  - Based on ratings & behavior of other users in the system
  - Estimate the preference of active user
- Basic assumption
  - Other users' opinions can be selected and aggregated in such a way as to provide a reasonable prediction of the active user's preference

Pure CF Approaches

- Input
  - Only a matrix of given user–item ratings

				
John 	5	1	3	5
Tom 	?	?	?	2
Alice 	4	?	3	?

Pure CF Approaches

- Output types
  - **Prediction:** A (numerical) prediction indicating to what degree the current user will like or dislike a certain item
  - **Recommendation:** A top-N list of recommended items

Prediction

User	Movie	Predicted Rating
Tom	Argo	5
Tom	Seven	4
Tom	Righteous Kill	3

Top n Recommendations

$$Tom: \{Argo, Seven, Righteous\ Kill\}$$

Analytics Vidhya

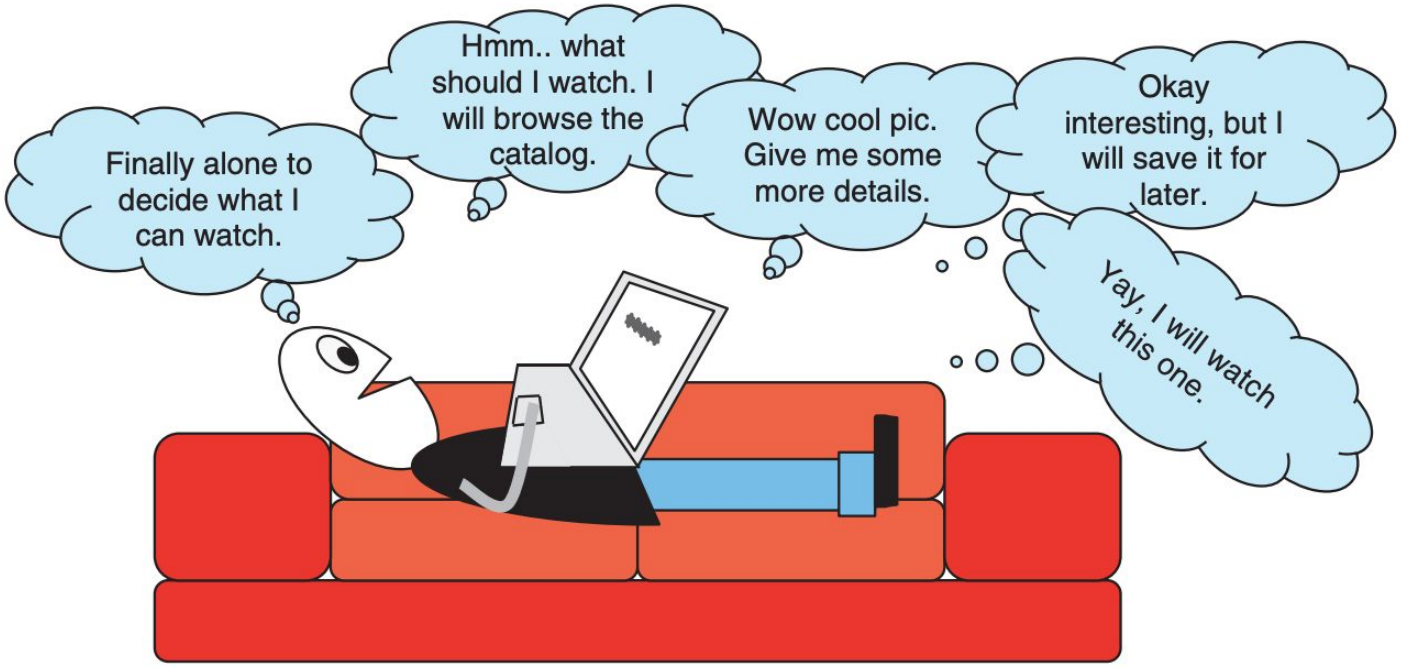
Explicit ratings

Analytics Vidhya

Explicit Ratings

- Most used (1 to 5, 1 to 7 response scales)
- Some Alternates
  - Optimal granularity of scale; indication that 10-point scale is better accepted in movie domain
  - Multidimensional ratings (multiple ratings per movie such as ratings for actors and sound)
- Main problem
  - Users not always willing to rate many items
    - number of available ratings could be too small → sparse rating matrices → poor recommendation quality

Implicit ratings



Analytics Vidhya

Implicit Ratings

- Typically collected by the application in which the recommender system is embedded
- When a customer buys an item, for instance, many recommender systems interpret this behavior as a positive rating
- Clicks, page views, time spent on some page, demo downloads …
- Implicit ratings can be collected constantly & do not require additional efforts from user
- Main problems
  - One cannot be sure whether the user behavior is correctly interpreted
  - For example, a user might not like all the books he or she has bought; the user also might have bought a book for someone else