



Music Recommendation System

Umang Jain(17074016)
Anuraj Singh(17074003)



What is Recommender System?

Recommender systems are techniques used for providing suggestion for a item related to various decision making process such as what items to buy or what music to listen.



Types

- Content based filtering :
 - Music Genome Project
- Collaborative filtering : relies on past user behaviour
 - More accurate and domain free
- Hybrid recommendation system



Collaborative Filtering

- Neighbourhood based method : relations between users or between items
- Latent factor models : Latent factor models are an alternative approach that tries to explain the ratings by characterizing both items and users.

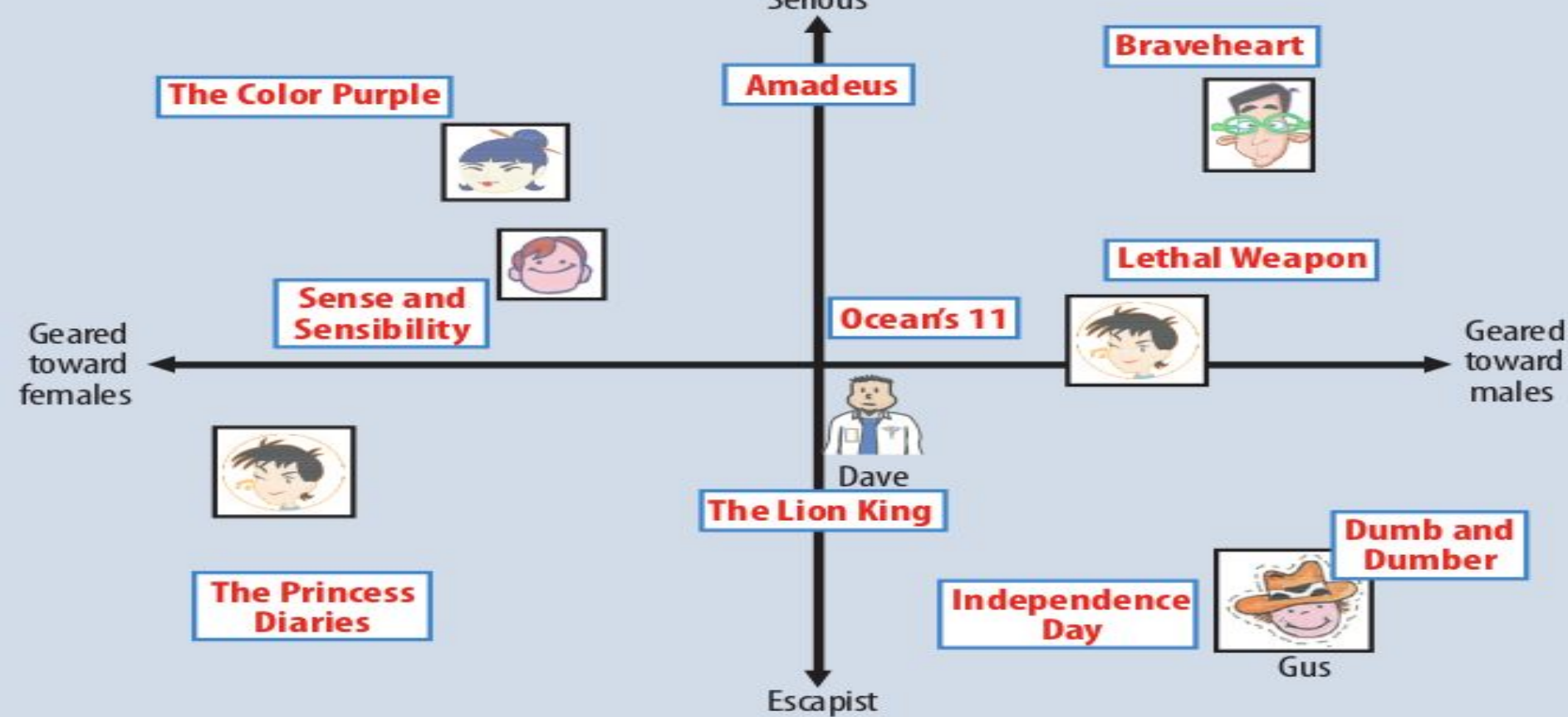


Figure 2. A simplified illustration of the latent factor approach, which characterizes both users and movies using two axes—male versus female and serious versus escapist.



Singular Value Decomposition

SVD, a well-established technique for identifying latent semantic factors in information retrieval. Applying SVD in the collaborative filtering domain requires factoring the user-item rating matrix. This often raises difficulties due to the high portion of missing values caused by sparseness in the user-item ratings matrix. Conventional SVD is undefined when knowledge about the matrix is incomplete. Moreover, carelessly addressing only the relatively few known entries is highly prone to overfitting.



Matrix Factorization

Some of the most successful realizations of latent factor models are based on matrix factorization . In its basic form, matrix factorization characterizes both items and users by vectors of factors inferred from item rating patterns.


$$\hat{r}_{ui} = q_i^T p_u$$

\hat{r}_{ui} = predicted rating by user u for item i

q_i = feature vector for item i

p_u = feature vector for user u



Cost Function

$$\min \sum (r_{ui} - q_i^T p_u)^2 + \lambda(\|q_i\|^2 + \|p_u\|^2)$$

λ = regularisation constant to avoid over-fitting



Adding Bias

$$\hat{r}_{ui} = \mu + b_i + b_u + q_i^T p_u$$

μ = overall average rating

b_i = bias for item i

b_u = bias for user u



Cold Start Problem

When there is no information present for item or user it becomes very hard to recommend something to this user.

A way to relieve this problem is to incorporate additional sources of information about the users.



$N(u)$ denotes the set of items for which user u expressed an implicit preference. This way, the system profiles users through the items they implicitly preferred.

$$\sum_{i \in N(u)} \mathbf{x}_i \quad \mathbf{x}_i = \text{implicit feature vector for item } i$$



Temporal Dynamics

$$\hat{r}_{ui}(t) = \mu + b_i(t) + b_u(t) + q_i p_u(t)$$



Confidence Levels

$$\min \sum C (r_{ui} - \mu - b_u - b_i - q_i^T p_u)^2 + \lambda(\|q_i\|^2 + \|p_u\|^2 + b_u^2 + b_i^2)$$



Future Work

- Playlist generation
- Playlist shuffling



Thank You