CS4065: Multimedia Search & Recommendation - Lab 2

Shashank Rao (4624017), Harikrishnan Manikandan(4613201)

May 17, 2017

1 Effect of low number of users

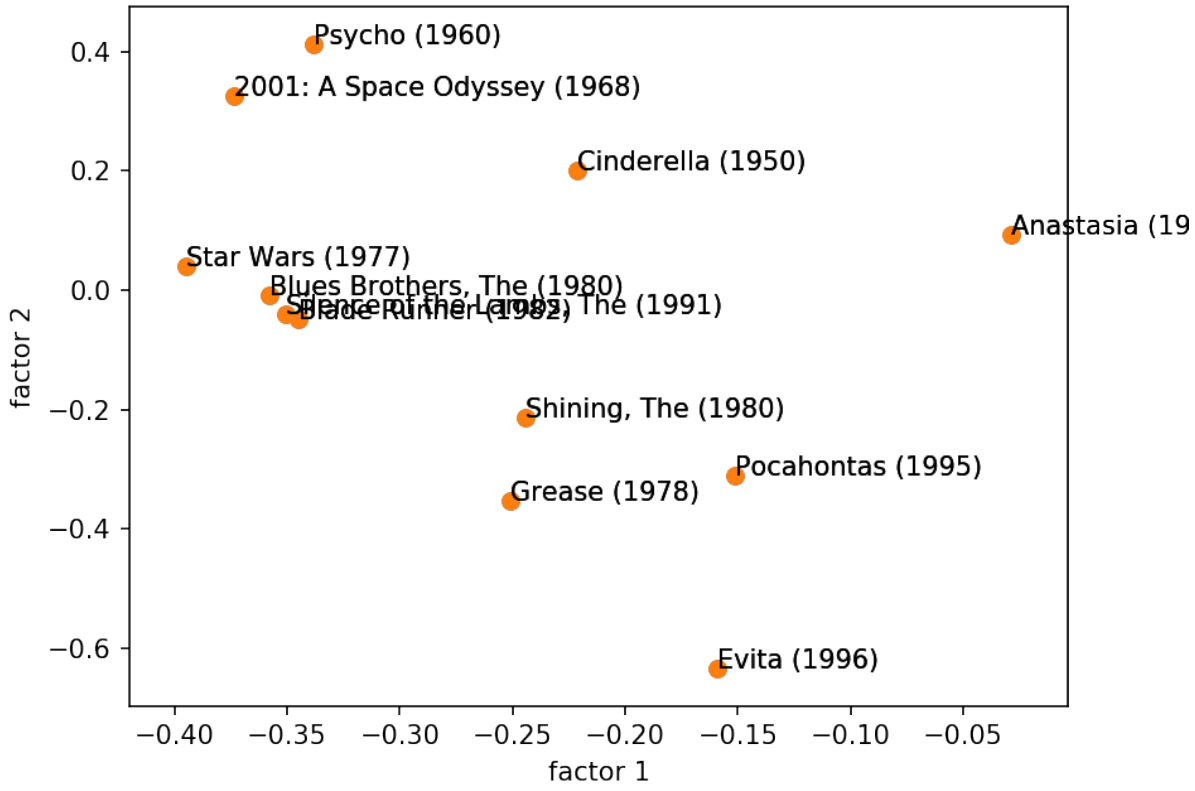

Figure 1: Scatter plot for top 15 users.

Figure 1 shows the scatter plot for the ratings given by **15 users**. It can be seen form the plot that the latent space for just 15 users does not differ much from the 10 users used in the lab exercise. Also, the horror movies or violent movies are well clustered when compared to other category of movies. It could be argued from this observation that **factor 1** captures the "horror" movies well and they are represented by the negative values of factors.

The RMSE value for 15 users was high (0.6). This could be due to the less number of users in the singular matrix to make a prediction of rating for a missing user.

2 Effect of very high number of users

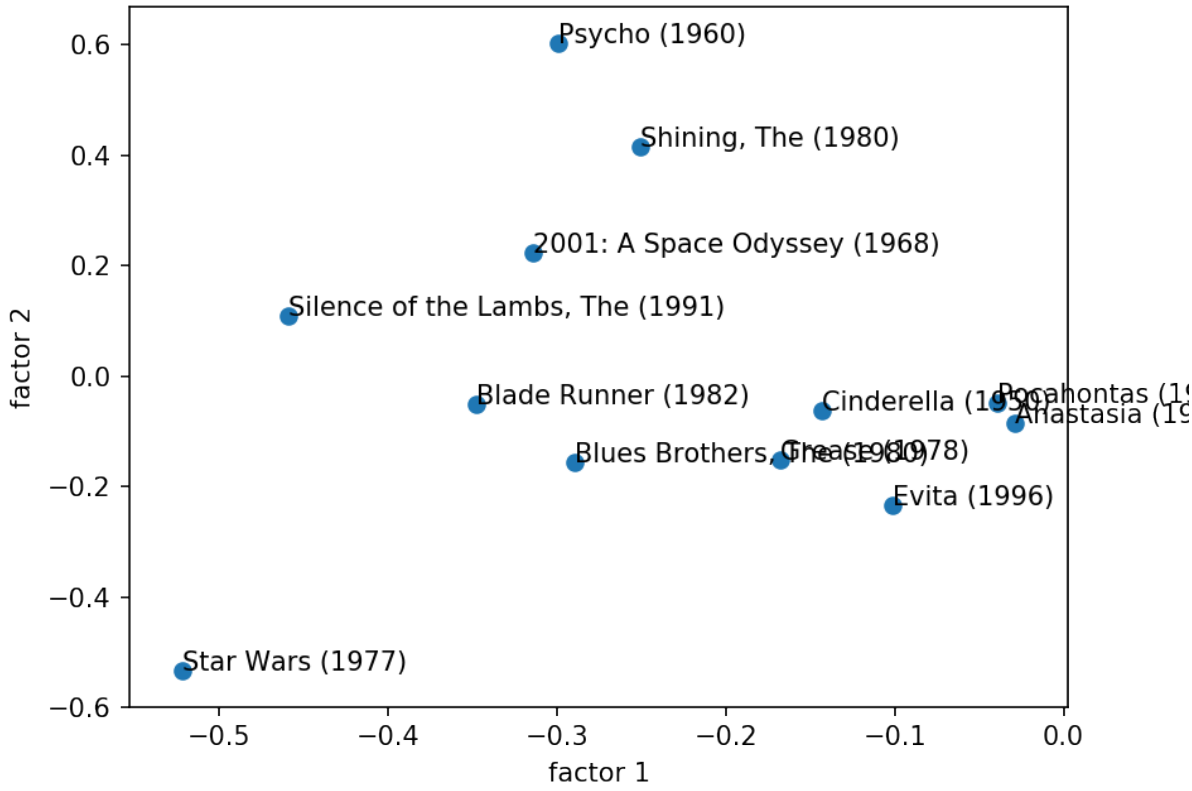

Figure 2: Scatter Plot using 500 users with most unique ratings.

Figure 2 shows the scatter plot for 500 users taken as per the variance in ratings. The users that have rated most different movies were chosen. With more number of users, the plot is more uniformly distributed, especially for "fairy" or "kid" movies as most of the users voted positively for them. The movie "Evita" also comes closer to other musical movies than when compared to 15 users in figure 1.

The factor values in case of 500 users is high with low variance (difference) among the lower factors. The RMSE values is also very low, around 0.25, with even some of the ratings being predicted better than the actual ones (something like overfitting).

3 Effect of different factors

Figure 3 shows the affect of factors value on latent space clustering. This plot consists of 500 users, but using factors 2 and 3. It can be seen that are clear clusters for "Disney" or fairy movies, "Horror" and "Space movies". Factor 2 captures these movies genres well. Musical movies are not captured well using these two factors.

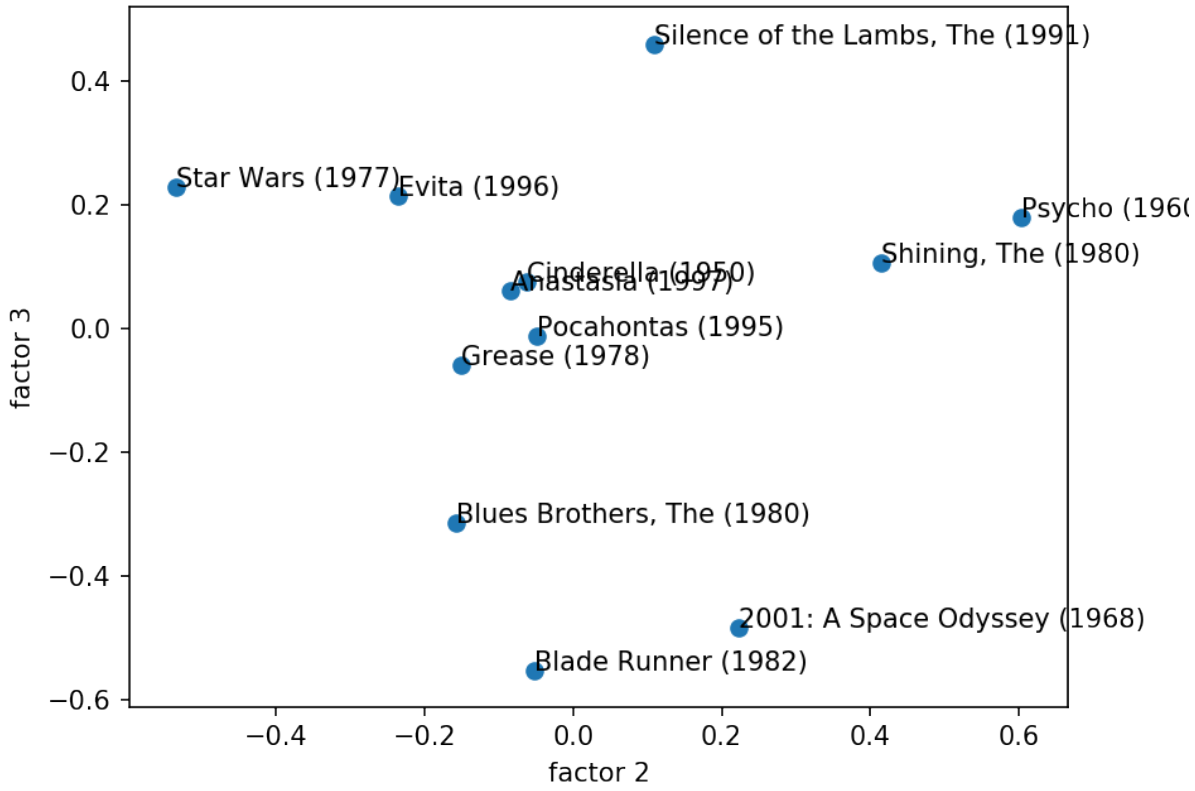

Figure 3: Effect of different factors, factor 2 & factor 3.