We built two recommender systems using python **surprise** package.[1]

- SVD model based collaborative filtering
- K-NN model based collaborative filtering

Mean Absolute Error	training error	test error
SVD	0.5803	0.8502
K-NN	0.3038	0.9141

Conclusion:

- Model-based collaborative filtering algorithms (SVD, K-NN) perform better than memory-based collaborative filtering algorithms, decreasing MAE by ***.
- By comparing training error and test error of the two model-based algorithms, we noticed K-NN suffers from an overfitting problem.

Introduction of Singular Value Decomposition Algorithm
Singular value decomposition algorithm is a good method to solve data sparsity issue. It use a latent factor model to capture the similarity between users and items. The unique of SVD method is its ability to discover unobvious dimension of items and users. For example, latent factors not only can reveal explicit dimensions like brand, price, but also can reveal characteristics like geared towards females or male, high-end or low-end.

More specifically, SVD decomposes the utility matrix(users-items rating) into UΣV^T, U denotes the projection of users on latent space, Σdenotes the strength of each latent factors and V^T denotes the projection of items on latent space. The prediction of $\widehat{R_{ui}}$:

$$\widehat{R_{ui}} = \mu + b_u + b_i + q_i^T p_u$$

Our goal is to solve the following minimization:

$$\sum_{\widehat{R_{ui}}} (\widehat{R_{ui}} - R_{ui})^2 + \lambda (b_i^2 + b_u^2 + |q_i|^2 + |p_u|^2)$$

Where bi is the bias of item i and bu is bias of user u, p and q represent the factor coefficients.

[1] http://surpriselib.com/