Problem 1)

**Predictor.py**, reads the ratings.csv file removing the timestamp column and converts it in to user/item matrix. It then calls user based collaborative filtering and item based collaborating filtering functions from **user\_item\_CF.py file**.

Problem 2)

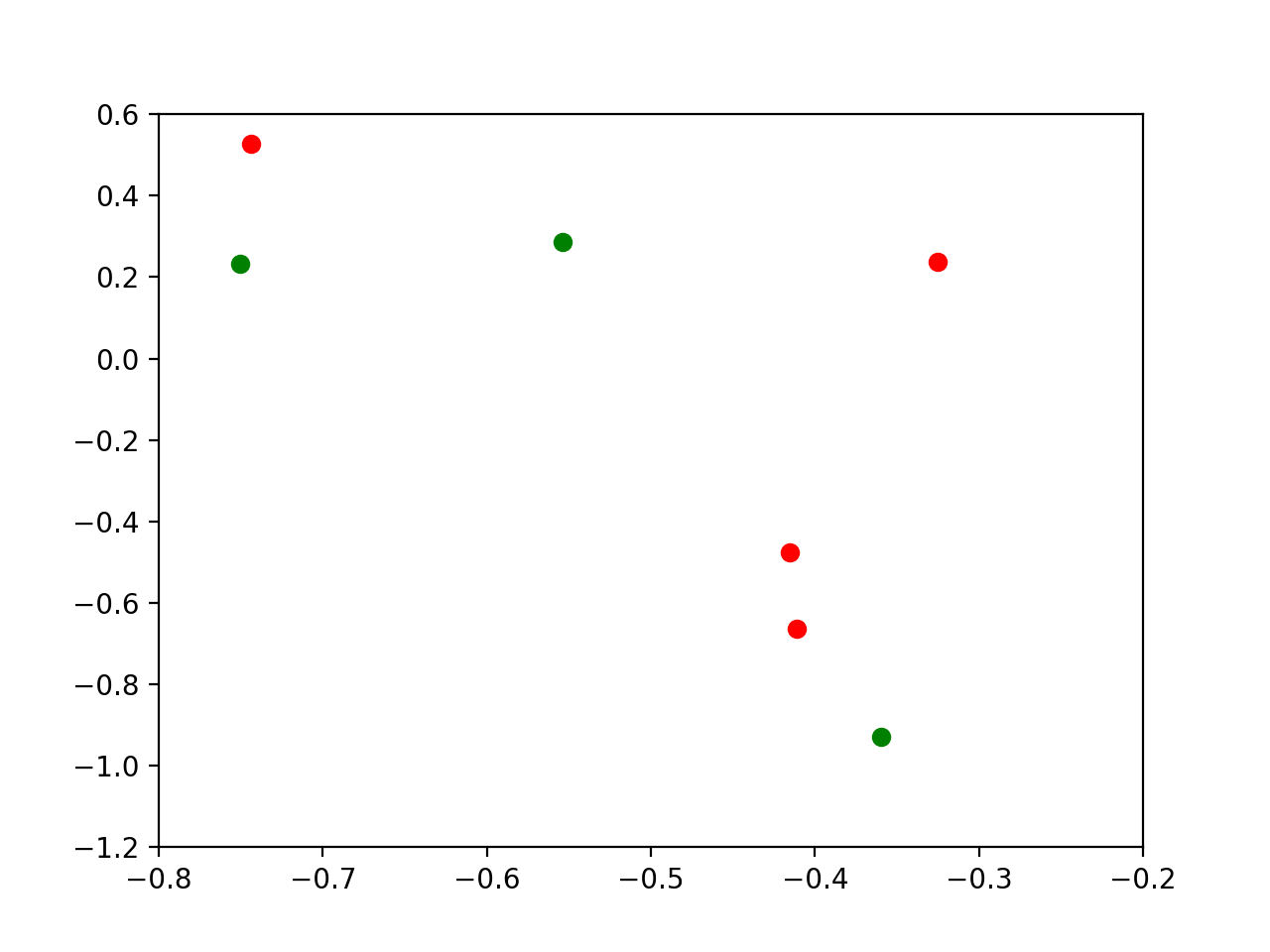
1)

Svd\_plot.py takes the list of movies (top 10 IMDB), extracts the corresponding movie id from movies.csv file. These extracted movie ID is then used to extract the corresponding user ID’s and ratings from ratings.csv file. All the extracted values are stored in IMDB.csv file which later is used to create user item matrix and compute svd for those top 10 movies.

Unfortunately, SVD for IMDB file requires huge memory space which my system could not comprehend.

However, to test the validity of the program I computed svd for the example given in the text book usng samp.csv file.

The plot for this example verifies the correctness of the program

Red dot represents the users and green represents the movies

2) The predicted ratings (without svd) for user 18 and movie 858 is,

User based CF : 1.8564

Item based CF: 1.2261

The file svd.py creates the matrix X from ratings.csv file, computes svd and creates the Xk matrix.

It then calls user based and item based CF programs from user\_item\_CF.py for Xk. The values could not be computed due to memory issue. The program has been verified for the example problem in textbook.