CF : Midsem

-Harkishan Singh (2017233)

Main code file : 2017233_Harkishan_Singh_CF_midsem.py

Explanation :

- ➔ Firstly, I have created a "user_movie_rating.csv" file which contains the rating each user has given to the movie. If the user has not given any rating to the movie, then that rating is set to 0. The file contains : Rating(i , j) ⇒ Rating of user i to movie j.
- ➔ Masking matrix R is created using the "get_r()" method and X0 is initialized randomly using "init_x0()" method. Y = original rating matrix.
- ➔ The main algorithm begins in nuclear_norm_minimization() method :
  - ◆ Iterate in a loop (I have looped 50 times, You can change it in the code)
  - ◆ B = X(i-1) + Y - np.multiply(
  - ◆ U,S,V = Singular value decomposition (B)
  - ◆ Soft threshold S and convert it to sigma. (I have used lamda = 10, you can change it in the code)
  - ◆ Recompute X = U.sigma.V
- ➔ I have completed matrix X and the original matrix Y. To compute the 5 fold mean absolute error ⇒ get the ith fold (0<i<4) and compare each element for the Y and X matrix. If the rating is not given by the user in the original matrix, then just skip it.

Running the code:
Run file "2017233_Harkishan_Singh_CF_midsem.py" and you'll see the output.

Results:
Below is the table of mean absolute error

Fold Number	Mean Absolute Error
1	0.24027305791360393
2	0.2270546175318653
3	0.2254771773538936
4	0.24397242388303508
5	0.24157264545209725

Average MAE = 0.23566998442689907