Recommendation Assignment Report

2014021358

노민수

1. Test environment

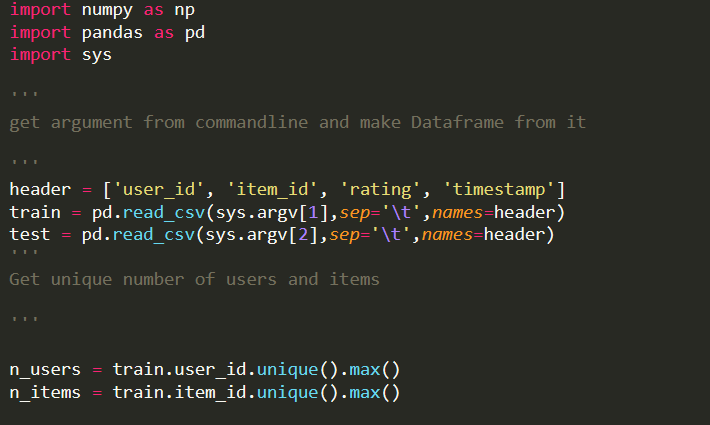
OS : Windows

Language : Python 3.7

2. How to implement : User-based Collaborative Filtering.

1) Get data from u1.base, u1.test and count the number of user\_id, item\_id.

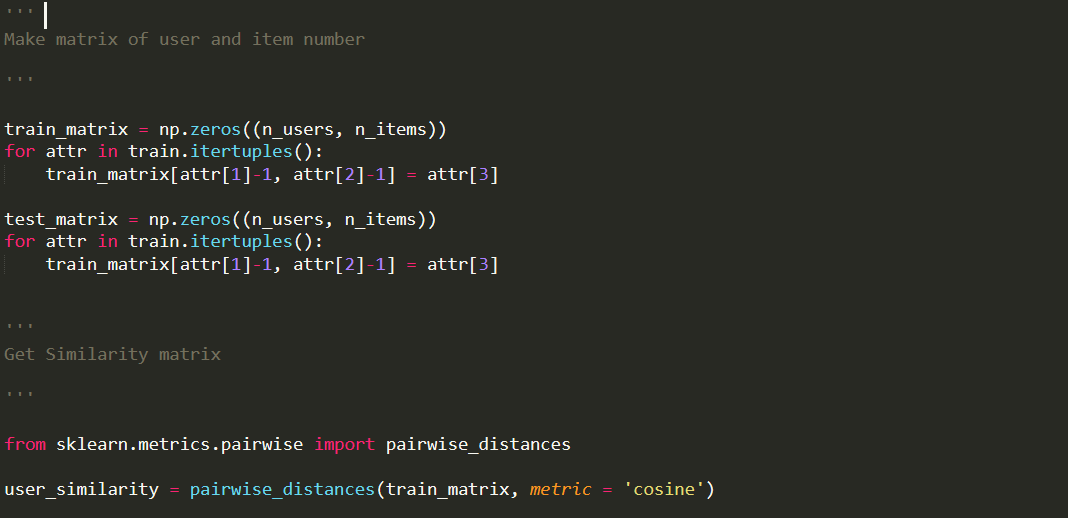
Import numpy to use numpy array , pandas to use dataframe, sys to read argument ( u1.base, u2.test) in command line. I’ve tried to use train.user\_id.unique().shape[0], but somehow it counts 1650 not 1683, therefore found maximum item number.



2) Make matrix for both test and train data and use cosine pairwise distance to get user-similarity matrix.

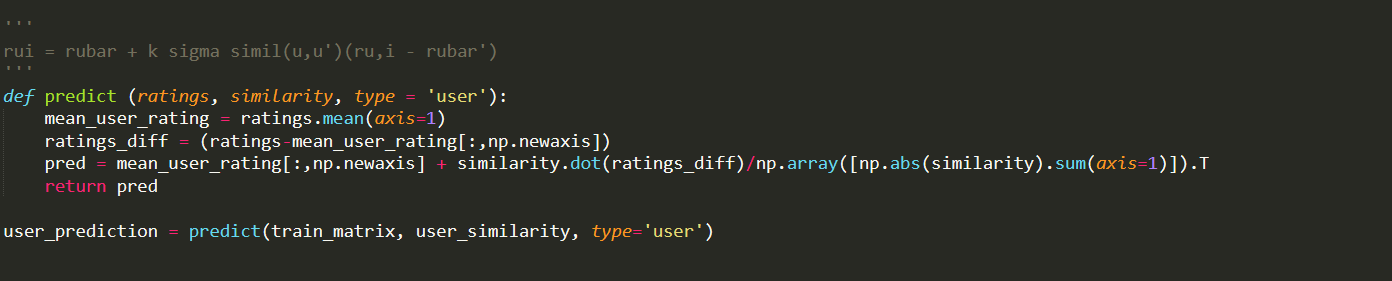
Transform dataframe to numpy array, row becomes user\_id, column becomes item\_id by train\_matrix[attr[1]-1, attr[2]-1] = attr[3]

Use sklearn to get cosine pairwise distance of train matrix

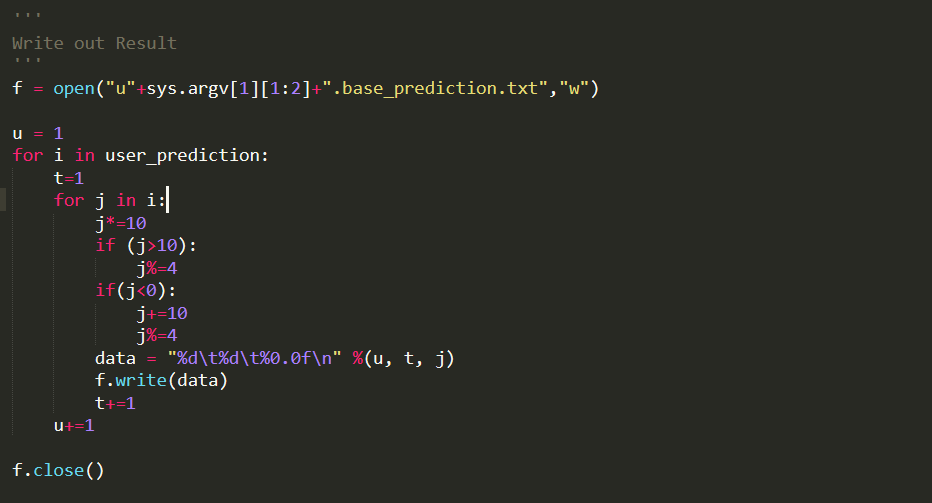


3) Get user-based prediction matrix

Get the difference ‘rui – rubar’ then use it to get pred. Rubar + k sigma similarity(u,u’). Get user\_prediction matrix of train\_matrix by using that function with user\_similarity.



4) Make output file u#.base\_prediction.txt



Get data from numpy matrix user\_prediction. U represent user\_id, t represent item\_id. Due to unexpected numbers, I’ve modified the result by giving those ratings valid value using correction in If statement to make higher correction rate.

3. How to execute : $python recom.py u1.base u1.test

base file and test file should be in the same directory as recom.py

u1 : 2.73

u2 : 2.86

u3 : 2.87

u4 : 2.86

u5 : 2.86

