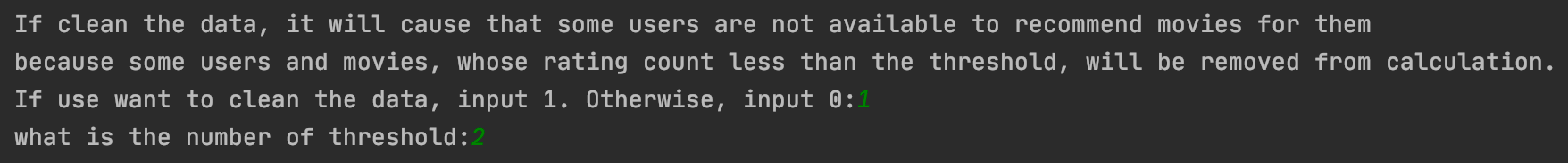
Instructions for use of Item-based and User-based RS

The clearlyItemBasedRS.py and the clearlyUserBasedRS.py is the main program of our RS. Others are just for testing the program. The dataPreprocessor.py is for data preprocessing. The data sets are paired-use. The IMDb data set is “./data/begin/ratings7.csv” and “./data/begin/movies7.csv”. And the Douban movie data set is “./ez\_douban/ratings.csv” and “./ez\_douban/movies.csv”.

Take clearlyItemBasedRS.py as example:

First, if you want to use IMDb dataset, input 0; if use Douban dataset, input 1.

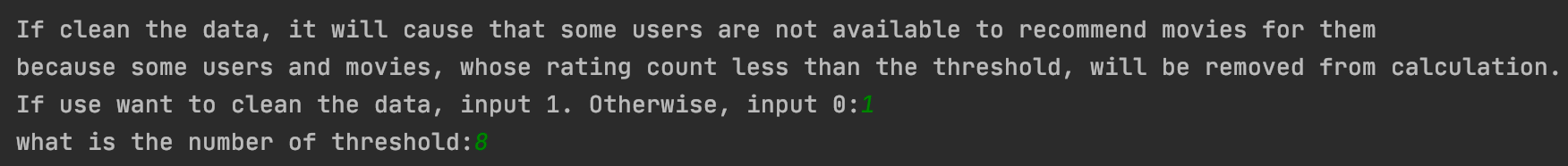
Then, the program will ask you to input some questions about the parameters. You can just follow the picture shown.



图片包含 游戏机, 电话, 键盘

描述已自动生成

However, if you choose to clean the data with a high threshold such as 8.



Then it will filter a lot of movies and users and only few users and movies left. You only can input one of the remaining user-id with a small neighbor size because there are no other neighbors left for calculate the recommendation. This situation will be better when you switch to use the Douban data set.

