

1 Description

In this programming assignment we will build a movie recommendation engine. You are provided a file, critics.csv. It is a small dataset but good enough for this exercise. The file contains critics (individuals) and their ratings for different movies. It is a comma separated file. Each line in the file represents the ratings of a critic (individual) and can be described as follows: first field is the critic's name, followed by the movie name and its rating. Since the number of movies that each individual has seen may vary, so not all lines contain the same number of fields and the order of the ratings may vary as well.

We want to create a system such that we can recommend movies to each of these individuals based on their preferences and the similarities with other users in the dataset. A higher rating for a movie means the individual has preferred that genre of the movies. So someone else who also rates such movies higher would be more similar to the first individual. Something akin to what Netflix would do, albeit on a very tiny dataset.

You are provided a starter Jupyter notebook with some pre-filled code and comments. You are supposed to fill the code in the appropriate locations. This file also contains cells that should print the output of various functions in the notebook

2 What to submit

Submit a hwk2_<your directory id>.ipynb file. You may add comments in the file to describe your code. The homework is to be turned in on ELMS. Do not use any in-built Python function, such as Pearson's coefficient etc., to directly any part of the code.