Assignment 2: User-Group Recommendations

How to run the code:

The program file is a .py file. When running the code, one requires to give the user ID of the interested user, as the input to the program. Then it'll provide answers to all questions provided in the assignment.

The assumptions made while completing the assignment:

- Groups with higher inner group similarity were selected for this assignment. In that case, use a function to select users (3 users) whose correlation between each other is higher than 0.25.
- For part b, we used the below function [1] $\mathcal{F}(\mathcal{G},i) = w_1 \times \mathtt{rel}(\mathcal{G},i) + w_2 \times (1 \mathtt{dis}(\mathcal{G},i))$; Where $w_1 + w_2 = 1$ To obtain the values for w_1 and w_2 , it's required to implement an experiment by iterating through different value combinations and obtain the accuracy of the results. Without doing such an experiment here we assume the value of $w_1 = 0.2$ and $w_2 = 0.8$
- For Individual recommendation It's assumed that obtaining **50 similar users** when making the predictions will provide a reasonable recommendation.

Other important things about the code:

- When recommending movies to the user group, the movies that no one from the group has rated were selected.
- When selecting 50 similar users for making the individual predictions, the top 50 users who have rated for a given particular movie were selected.
- When calculating the Pearson correlation between 2 users, only the commonly rated movies were selected. This is done as there are many missing values in the dataset. Also, correlation is calculated if the commonly rated movie count is greater than 2. This is done because we can't say 2 users are similar if they have similar ratings for a very small number of movie like 1,2.
- [1] S. R. S. B. C. A. D. G. &. Y. C. Amer-Yahia, "Group recommendation: semantics and efficiency.," *Proceedings of the VLDB Endowmen*, vol. 2(1), p. 754–765, 2009.