



Northeastern University, Khoury College of Computer Science

CS 6220 Data Mining | Assignment 3

Due: February 15, 2023(100 points)

Chang Liu

<https://github.com/Chang-Liu-Harry/6220DataMining>

The script and results can be found in the link and the ipynb file attached with this doc.

For user 11, as required the recommendations should be:

27552, 7785, 27573, 27574, 27589, 27590, 27600, 27617, 27620, 27667

But the result I have for top 10 is not exactly the same. Therefore I printed out the top 20 to see why that happens:

User 11: 27552-4, 27667-3, 27620-3, 33192-3, 32072-3, 27600-3, 27617-3, 27573-3, 7785-3, 27589-3, 27574-3, 27590-3, 27587-2, 27555-2, 27655-2, 27583-2, 35451-2, 25195-2, 10-2, 12-2

It turns out, there are 11 recommendations and all have 3 mutual friends. Hence, I assume there should not be an unique correct answer.

With that said, here is the recommendation for the user set required:

User 9020: 9021,9022,9016,9017,317,9023

User 9992: 9987,9989,35667,9991

User 924: 45881,439,2409,15416,11860,43748,6995

User 9021: 9020,9017,9022,9016,317,9023

User 9993: 9991,13877,34485,37941,13134,13478,34642,34299

User 8941: 8943,8944,8940

User 8942: 8939,8940,8944,8943

User 9022: 9020,9021,9019,9016,9017,9023,317

User 9990: 34299,13134,13478,34642,13877,34485,37941

User 9019: 9022,9023,317