CS5600
Assigment 1
Expert Recommendation System

Group 3
Members

Name	Roll Number
Kritik Agarwal	CS23MTECH11009
Darpan Gaur	C021BTECH11004
Hari Priyanka Allam	CC23M24P100001
Maloth David	CS21BTECH11035

Exercise 1

The top 3 answerers and the top 3 tags, along with their answer count and annotation count, are listed below:

Top 3 Answerers

OwnerUserId	**AnswerCount**
9113	2839
177980	2326
1204	2043

Top 3 Tags

Tags	**TagCount**
design	5162
c#	4931
java	4928

Exercise 2

The dimension of the Expert utility matrix is: **(1163, 974)**

Exercise 3

Utility Matrix

- Summation value: **41403**
- Highest row sum: **1164**
- Highest column sum: **1403**

Train and Test Data

- Train matrix summation value: **28713**
- Test matrix dimension: **(174, 146)**

- Test matrix summation value: **899**

Exercise 4

Performance of Item-Item and User-User Recommendation Systems

Method	Rating Prediction Function	Metric	N=2	N=3	N=5
Item-Item	Simple average	RMSE	0.3078	0.3078	0.3078
	Weighted average	RMSE	0.3078	0.3078	0.3078
User-User	Simple average	RMSE	0.3078	0.3078	0.3078
	Weighted average	RMSE	0.3078	0.3078	0.3078

Exercise 5

Performance of Item-Item and User-User Recommendation Systems

Method	Metric	K=2	K=5	K=10
Without Regularization	RMSE	0.874779	0.734253	0.591676
Without Regularization $\lambda_1 = 0.001$, $\lambda_2 = 0.003$	RMSE	0.880756	0.739713	0.595585
$\lambda_1 = 0.05$, $\lambda_2 = 0.05$	RMSE	0.883248	0.742098	0.589883
$\lambda_1 = 0.50$, $\lambda_2 = 0.75$	RMSE	0.878045	0.748843	0.594022

Exercise 6

KNN-Baseline surprise library vs Exercise 4

Algorithm	Method	RMSE for N=2	RMSE for N=3	RMSE for N=5
Item-Item	Your method	0.3078	0.3078	0.3078
	Surprise library	0.3058	0.3058	0.3058
User-User	Your method	0.3078	0.3078	0.3078
	Surprise library	0.3058	0.3058	0.3058

- The RMSE values for the methods implemented by us and the surprise library are almost similar.
- For surprise library, the RMSE values are slightly lower than the values obtained by us.
- This is becasue KNN Baseline also do bais correction, while our implementation does not.

SVD surprise library vs Exercise 5

Method	RMSE for K=2	RMSE for K=5	RMSE for K=10
Your Method	0.874779	0.734253	0.591676
Surprise	0.3058	0.3058	0.3058

- Did grid serach on hyperparameters, n_epochs, reg_pu and reg_qi.
- Taken lr_all = 0.0005 as mentined in exercise 5, n_epochs = 1000, reg_pu = 0.05, and reg_qi=0.05.