CIT 594 – Homework 4

Milestones

This homework has three milestones with the following points and deadlines:

~~Milestone 1 – Design (30 points)~~
~~Due Monday, Mar. 21, at 12pm~~

Milestone 2 – Implementation (70 points)
Due Monday, Mar. 28, at 12pm

Milestone 3 – Changing Requirements (30 points)
Due Friday, Apr. 1, at 12pm

Note: Late submissions are not allowed for Milestone 2.

Recommender Systems – User-User Collaborative Filtering

Milestone 2 – Implementation

For milestone 2, the goal is to implement your recommender system based on your initial design. More details follow:

1. The data file provided is the MovieLens 10M Dataset. It contains 10 million ratings for 10000 movies by 72000 users. More details on the dataset are here: http://grouplens.org/datasets/movielens/10m/
2. The most relevant file will be the ratings.dat file, but take a look at the README for details on file formats, file structure, etc.
3. The main functionality that your system should support is the following:
    a. Given a user u and item i, what is the system's prediction for the user's likely preference of that item?
    b. Given a user u and a threshold n, what are the n-highest predicted preferences for that user?

Grading Criteria for Milestone 2

80% for functionality – Does the code work as required? Does it crash while running? Are there bugs? …
10% for design – Does your design match the one from Milestone 1?
10% for style – Do you have good comments in the code? Are your variables named appropriately? ...

Programming – General Comments

Here are some guidelines with respect to programming style.

Please use Javadoc-style comments.

For things like naming conventions, please see Appendix I (Page A-79) of the Horstmann book. You can also install the Checkstyle plugin (http://eclipse-cs.sourceforge.net/) in Eclipse, which will automatically warn you about style violations.

Submission Instructions

You should also submit a text file titled readme.txt. That is, write in plain English, instructions for using your software, explanations for how and why you chose to design your code the way you did. The readme.txt file is also an opportunity for you to get partial credit when certain requirements of the assignment are not met. Think of the readme as a combination of instructions for the user and a chance for you to get partial credit.

If your design deviated from the one you submitted during milestone 1, please explain why and what changes you made in the readme file.

Please create a folder called YOUR_PENNKEY. Places all your files inside this – the Java files, the readme.txt file. Zip up this folder. It will thus be called YOUR_PENNKEY.zip. So, e.g., my homework submission would be swapneel.zip. Please submit this zip file via canvas.