**COMP 7745/8745**

**Spring 2017**

**Project: Recommendation System for Movie Ratings**

**Due Date: April 21, 2017 (submit on Ecourseware)**

**Project Description**

Here you will implement the collaborative filtering algorithm and apply it to a dataset containing movie ratings.

* Read the paper “Empirical Analysis of Predictive Algorithms for Collaborative Filtering”. https://arxiv.org/ftp/arxiv/papers/1301/1301.7363.pdf
* You need to read up to Section 2.1 of the paper, and are encouraged to read further if you have time.
* Implement the collaborative filtering algorithm described in Section 2.1 of the paper (Equations 1 and 2) for making the predictions. You may program in C, C++, C#, Java, Python (or any other programming language of your choice!)
* The dataset we will be using is a subset of the movie ratings data from the Netflix Prize. It contains review-scores, a movies file, a dataset description file, and a README file.
* The dataset description file further describes the dataset, and will help you get started. The README file is from the original set of Netflix files, and has been included to comply with the terms of use for this data.

**Submission**

Provide the source code that you write along with any libraries/references you have used. Your code should implement the following

* **Answering queries:** We are interested in how you would recommend movies to a user. Specifically, the input will be a user-id and an integer K. The output should contain K recommended movies for that user-id (ordered by preference)
* Please provide a readme file that we can use to compile your code, and run it for both the above cases. No GUI is needed, you can use command-line options to run your code for the above two cases.

**Hint**: Pay special attention to complexity in your implementation, i.e., what measures can you compute and store in a lookup table to make it efficient to answer queries, etc. The dataset is quite large, if you want you can reduce the size of the dataset during development.

**Plagiarism Notice**

This is an individual project. Please DO NOT take copy code from the internet or from others in the class. We will check for plagiarism using automated tools.