ECE 6254 Final Project

The final project is worth 30% of your grade, and will consist either of an independent investigation into a topic of your choosing, or alternatively, the implementation of a recommendation system (described below). As before, the project can be done individually or in groups of two (I would recommend groups – if you have trouble finding a partner, let me know and I will try to help you find one).

Independent investigation

For the final project, you will have a choice of two formats. If you choose to pursue an independent investigation, your project will involve an in-depth study of a topic of your choice. It can focus on an in-depth theoretical study of some particular topic, the design and implementation of a cutting-edge algorithm, a novel application of something we've discussed in class, or on some combination of these. It will be treated as a "research project," so some type of original synthesis is a required component.

If you select this format for your project, I would like you to notify me of your intended subject by **Monday, April 14**. You may also optionally arrange a short meeting with me to discuss your project.

For this format, the main deliverable for the project will consist of a short in-class presentation on April 23 and a writeup describing the results of your investigation (10 pages or (perferrably) less, no smaller than 10 point font). Turn in your writeup as a .pdf file named "Lastname1Firstname1-Lastname2Firstname2.pdf". The writeup must be submitted via e-mail by **5pm on Friday, May 2** (the end of the finals period). Earlier submission is encouraged, but not required.

Your grade for this format will be assigned as follows:

1. Originality of the project. (Did you simply summarize someone else's paper or did you do something truly innovative?) (35 points)
2. Clarity/quality of your writeup. (35 points)
3. Clarity/quality of your in-class presentation. (30 points)

Finally, I will also award what I deem to the "most original/innovative" project an automatic 100%.

Recommendation system

The alternative format for your project is to implement a movie recommendation system. On the course website you can download a data set in (`movieRatings.zip`) that consists a dataset containing $80,000$ ratings (1-5) from 943 users on 1682 movies. This data is organized as a matrix which is mostly zeros (corresponding to user/movie combinations for which we have no ratings) but which has $80,000$ nonzeros corresponding to the observed ratings. The file also contains some

simple demographic data about each user and a few pieces of data about each movie. See the included README file for more details.

Your goal in this project is to implement a "recommendation system" that will take the observed ratings and form predictions for the unobserved ratings, using whatever methods you wish. The end result should be a "complete" matrix that has predicted ratings for every user/movie combination.

This format for the project will have two main deliverables:

1. A .zip file that contains a file `predictedRatings.mat` containing your predicted ratings as well as the code you used to obtain this prediction. Your code can be organized however you like, but the file `predictedRatings.mat` must contain a single 943×1682 matrix that contains predicted ratings for each movie. Please name your .zip file "Lastname1Firstname1-Lastname2Firstname2.zip", where the names correspond to the names of each member of your group. This must be submitted via e-mail by **11:59pm on Tuesday April 22**.

2. A short in-class presentation on April 23.

3. A writeup (10 pages or (preferrably) less, no smaller than a 10 point font) that describes the approach you took. Turn in your writeup as a .pdf file named "Lastname1Firstname1-Lastname2Firstname2.pdf". The writeup must be submitted via e-mail by **5pm on Friday, May 2** (the end of the finals period). Earlier submission is encouraged, but not required.

When you turn in your predicted ratings, I will compare them to some test data that is not included as part of the training data. **Note that it might be possible to find such testing data on the internet somewhere, if you are very very enterprising. However, I would consider this a violation of the GT Honor Code – please only use the training data provided.** I will not tell you the exact performance metric that I will use in advance, but it will be a relatively straightfoward metric that compares your predicted ratings to the actual ones.

As with the mini-project, the group that achieves the best performance on the test data set will receive an automatic 100% on the project. For the remaining groups, the grade will be assigned as follows:

1. How accurate are your recommendations (based on my performance metric)? (10 points)

2. Originality/apprpopriateness of your approach. (25 points)

3. Clarity/quality of your writeup. (35 points)

4. Clarity/quality of your in-class presentation. (30 points)