**EE5373: Data Modeling Using R**

**Fall, 2017**

Department of Electrical and Computer Engineering

**University of Minnesota**

Lab 8: Final project.

Due date: See the due date shown on the class moodle page.

Goal: To demonstrate that you can integrate what you have learned in this course to study an interesting set of data using appropriate data modeling techniques.

What to do:

You are to conduct an independent project that meets the following minimum requirements:

1. It must show the development of an appropriate regression model.
2. It must use a data set consisting of at least 1000 unique entries.
3. It must demonstrate appropriate data cleaning/sanity checking techniques.
4. It must demonstrate appropriate use of training and testing ideas.
5. It must include an appropriate analysis of the quality of the model.
6. It must demonstrate your mastery of the R language.
7. It must make reasonably good predictions.
8. It should be interesting.
9. It should be fun.
10. It may be useful.

What to turn in for grading:

Write a report (2-4 pages) that describes how you satisfy all of the criteria above. Upload a pdf file with your report to moodle by the due date.

Example project ideas:

1. Solve an interesting problem using data from kaggle.
2. Analyze an interesting data set from your own research field.
3. Determine how you can make better predictions from the CPU DB data.