**CSCI 4155/6508 Machine Learning 2016: Assignment 2**

Please note that late assignments are not accepted. Also, I recommend that you start with the assignments early as I expect that you have questions. Past experience has shown that starting assignments the night before due date is usually problematic.

**1. Group project:**

This small project is an exercise where you should apply a binary machine learning classification algorithm to a problem of your choice. You might have an own dataset or you could choose a dataset from the UCI machine learning repository at http://archive.ics.uci.edu/ml. Note that many problems are not binary, but you can usually change these data sets to a binary problem domain, for example by changing a multivalued label such as a size measure into two classes such as small and large. You should modify the programs form the Gaussian example in class to suit your problem, and you should evaluate the performance of some classifiers with cross validation.

Please form groups of two and choose your topic. Please inform me of your topic and group by email to dalhousieml2016@gmail.com by Wednesday, Sept 21 before class. On Wednesday I will provide sign up sheets for a 5min presentation the following week on Monday and Wednesday. A short 1 page write-up will be due on Friday, Sept 30.

**2. Theoretical limit of Gaussian classification example**
   (Only Grad Students in CSCI6505)

We followed an example of classifications with two Gaussian classes in section 2.4. In this assignments you should calculate analytically the theoretical limit of the optimal accuracy for the parameters used in the printed program. Provide your answer with a brief outline of the calculation by email to dalhousieml2016@gmail.com by Friday, Sept 23. You can scan a hand written sheet if you like.