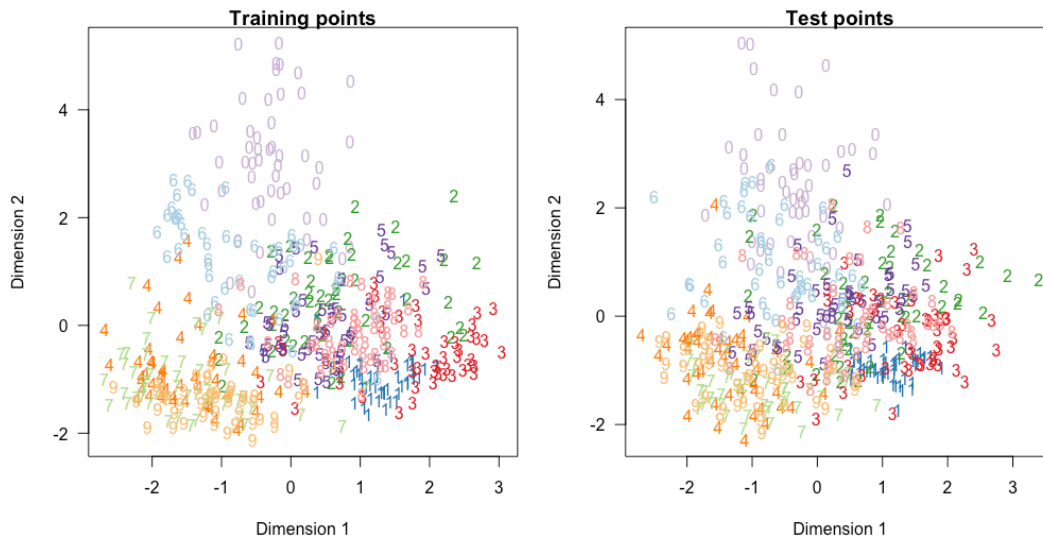


## COMP/INDR 421/521 HW06: Linear Discriminant Analysis

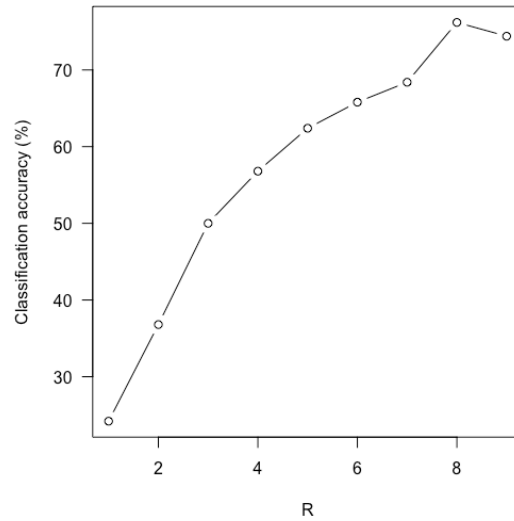
Deadline: December 15, 2017, 11:59 PM

In this homework, you will implement the linear discriminant analysis algorithm in R, Matlab, or Python. Here are the steps you need to follow:

1. You are given a data set, which contains 500 training and 500 test data points in the files `hw06_mnist_training_digits.csv` and `hw06_mnist_test_digits.csv` together with the labels in the files `hw06_mnist_training_labels.csv` and `hw06_mnist_test_labels.csv`.
2. Implement the linear discriminant analysis algorithm discussed in the lecture. (Hint: If the within-class scatter matrix is not invertible, you can, for example, add  $1e-10$  to its diagonal entries to get rid of the singularity.)
3. Calculate the two-dimensional projections of the training and test data points using the linear discriminant analysis by setting  $R$  to 2. Draw the two-dimensional projections of the training and test data points. Your figures should be similar to the following figures.



4. Calculate the low-dimensional projections of the training and test data points using the linear discriminant analysis by changing  $R$  from 1 to 9. For each low-dimensional representation, learn a 5-nearest neighbor classifier using the projections of training data points and calculate the classification accuracy on the projections of test data points. Draw the classification accuracy as a function of the dimensionality of the projected subspace. Your figure should be similar to the following figure.



**What to submit:** You need to submit your source code in a single file (.R file if you are using R, .m file if you are using Matlab, or .py file if you are using Python) and a short report explaining your approach (.doc, .docx, or .pdf file). You will put these two files in a single zip file named as ***STUDENTID.zip***, where ***STUDENTID*** should be replaced with your 7-digit student number.

**How to submit:** E-mail the zip file you created to [mehmetgonen@ku.edu.tr](mailto:mehmetgonen@ku.edu.tr) with the subject line ***Intro2MachineLearningHW06***. Please follow the exact style mentioned for the subject line and do not send a zip file named as ***STUDENTID.zip***. Submissions that do not follow these guidelines will not be graded.

**Late submission policy:** Late submissions will not be graded.

**Cheating policy:** Very similar submissions will not be graded.