

## ENGR 421 / DASC 521: Introduction to Machine Learning

### Homework 07: Linear Discriminant Analysis

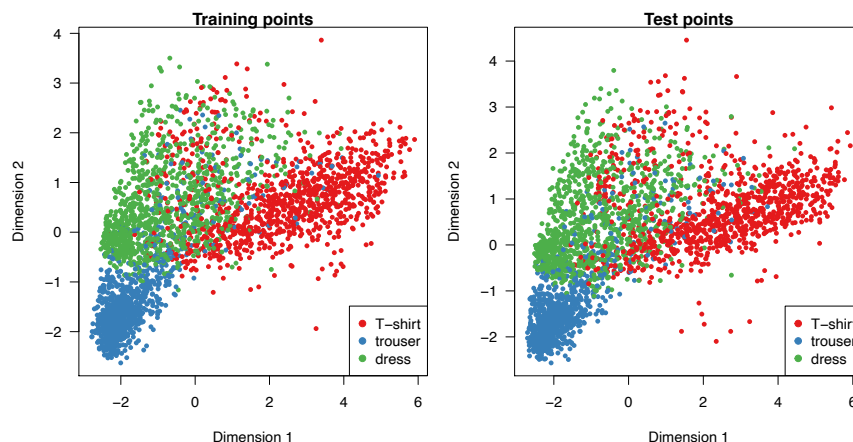
Deadline: January 15, 2021, 11:59 PM

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

1. Read Section 6.8 from the textbook.
2. You are given a data set, which contains 3000 training and 3000 test images in the files `hw07_training_images.csv` and `hw07_test_images.csv` together with the labels in the files `hw07_training_labels.csv` and `hw07_test_labels.csv`. These data points are clothing images of size 28 pixels  $\times$  28 pixels (i.e., 784 pixels). These images are from three distinct classes, namely, T-shirt, trouser, and dress. The figure below shows five sample images from each class.



3. Implement the linear discriminant analysis algorithm. (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.)
4. 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.



5. For two-dimensional representation, learn a five-nearest neighbor classifier using the projections of training data points and calculate the confusion matrices on the projections of training and test data points, respectively. Your results should be similar to the following confusion matrices.

```
      y_predicted
y_train  1    2    3
      1 886   14 112
      2  40 845   89
      3  87  47 880
```

```
      y_predicted
y_test   1    2    3
      1 834    8 158
      2  49 851 100
      3 133   61 806
```

**What to submit:** You need to submit your source code in a single file (.m file if you are using Matlab, .py file if you are using Python, or .R file if you are using R) 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:** Submit the zip file you created to Blackboard. Please follow the exact style mentioned 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.