Project 1: Linear Classification Methods and QDA

The aim of this project is to practice linear classification methods and QDA and study basic techniques of dimension reduction. You are strongly encouraged to use datasets related to your research. You only need to choose one dataset to work on.

Five datasets are available at
`http://www.stat.psu.edu/~jiali/stat557/material.html`. In the directory of each dataset, one file provides general information about the dataset; and the other contains data. For the UCI data files, each line corresponds to one sample (or case). All numbers except the last one in each line are values of the features. The last number, or sometimes letter, indicates the class identity. To work on the aerial images, you need background on preliminary image processing.

Requirements for the project are listed below.

1. Apply LDA, QDA and logistic regression to your dataset and compare them.

2. Study whether dimension reduction improves classification. You can use principle components, a certain feature selection procedure, etc.

3. You are encouraged but not required to use cross-validation to evaluate your classification algorithm. If you decide not to use cross-validation, you can divide the data randomly into a training set and a testing set.

4. Write a report. In the report, you are required to explain the contribution of each individual group member.