

# MthStat 568/768 – Multivariate Statistical Analysis – Spring 2024

## Homework 4

*Due Wednesday, March 27*

1. Consider the `bupa_liver_disorder` data set, where a number of individuals are classified into two classes according to blood test results (see full description on p. 258 of the book).
  - (a) Split the data set into training and test sets (roughly a 70/30 split). Compute the logistic classifier using the training set.
  - (b) Construct the misclassification table and compute the misclassification rate for the test set.
2. Consider the `ecoli` data set. These data were obtained in a study of protein localization sites for 336 examples of *E. coli*. There are 7 predictor variables and a class variable, `localization_site`, which indicates the protein localization.
  - (a) Split the data set into training and test sets (roughly a 80/20 split). Compute the multinomial logistic classifier using the training set.
  - (b) Construct the misclassification table and compute the misclassification rate for the test set.