

Math 583 HW 11 Fall 2017. Due MONDAY, Dec. 4.

Exam 3, FRIDAY, Dec. 1. Use 10 sheets of notes.

Quiz 11 on WEDNESDAY, Dec. 6 is similar to HW 11. Use 4 sheets of notes.

Final Wednesday, December 13, 10:15-12:15. Cumulative, 25 sheets of notes. Or do a project due final date.

If you do not use the Math lab, then you need to download some libraries. You did this for homework 1 with libraries *glmnet*, *leaps*, and *pls*. Use the `install.packages("e1071")` and `install.packages("randomForest")` commands near the top of the *slrhw* file, pick a mirror from Iowa or Indiana, and follow the directions. (Lab 1 explained how to download a package.)

Problem numbers and example numbers are from the Olive text. Do the two source commands to get the data into *R*.

A) 5.16 This is the pottery data of Problem 5.11, but the 28 cases were classified as Arrentine for $y = -1$ and nonArrentine for $y = 1$.

a) Copy and paste the commands for this part into *R*. These commands make the data and do bagging. Copy and paste the truth table into *Word*. What is the AER?

b) Copy and paste the commands for this part into *R*. These commands do random forests. Copy and paste the truth table into *Word*. What is the AER?

c) Copy and paste the commands for this part into *R*. These commands do SVM with a fixed cost. Copy and paste the truth table into *Word*. What is the AER?

d) Copy and paste the commands for this part into *R*. These commands do SVM with a cost chosen by 10-fold CV. Copy and paste the truth table into *Word*. What is the AER?

B) This problem uses the Gladstone brain weight data and classifies gender (F for $y = -1$, M for $y = 1$) using various predictors including head measurements, brain weight, and height. Some outliers were removed and the data set was divided into a training set with $n = 200$ cases and a test set with $m = 61$ cases.

a) Copy and paste the commands for this part into *R*. These commands make the data and do bagging. Copy and paste the truth table into *Word*. What is the AER?

b) Copy and paste the commands for this part into *R*. These use bagging on the training data and validation set. Copy and paste the truth table into *Word*. What is the bagging validation error rate?

c) Copy and paste the commands for this part into *R*. These commands do random forests. Copy and paste the truth table into *Word*. What is the AER?

d) Copy and paste the commands for this part into *R*. These use random forests on the training data and validation set. Copy and paste the truth table into *Word*. What is the random forests validation error rate?

e) Copy and paste the commands for this part into *R*. These commands do SVM with a cost chosen by 10-fold CV. Copy and paste the truth table into *Word*. What is the AER?

f) Copy and paste the commands for this part into *R*. These commands do SVM with a cost chosen by 10-fold CV on the training data and validation set. Copy and paste the truth table into *Word*. What is the SVM validation error rate?