Bruce Campbell ST-617 Homework 4

Tue Jul 19 09:37:50 2016

```
rm(list = ls())
set.seed(7)
```

Chapter 8

Problem 5

Suppose we produce ten bootstrapped samples from a data set containing red and green classes. We then apply a classification tree to each bootstrapped sample and, for a specific value of X, produce 10 estimates of $P(ClassisRed|X)$

```
data <- c(0.1, 0.15, 0.2, 0.2, 0.55, 0.6, 0.6, 0.65, 0.7, 0.75)
library(pander)
pander(data)
```

0.1, *0.15*, *0.2*, *0.2*, *0.55*, *0.6*, *0.6*, *0.65*, *0.7* and *0.75*

There are two common ways to combine these results together into a single class prediction. One is the majority vote approach discussed in this chapter. The second approach is to classify based on the average probability. In this example, what is the final classification under each of these two approaches?

```
Red <- data > 0.5

sum_red <- sum(Red)

is_red_by_vote <- sum_red >= 5

mean_probability <- mean(data)

is_red_by_mean <- mean_probability > 0.5

results <- data.frame(method = c("voting", "mean"), is_red = c(is_red_by_vote,
    is_red_by_mean))

pander(results)
```

method	is_red
voting	TRUE
mean	FALSE