STAT 388

Charles Hwang

Professor Matthews

STAT 388-001

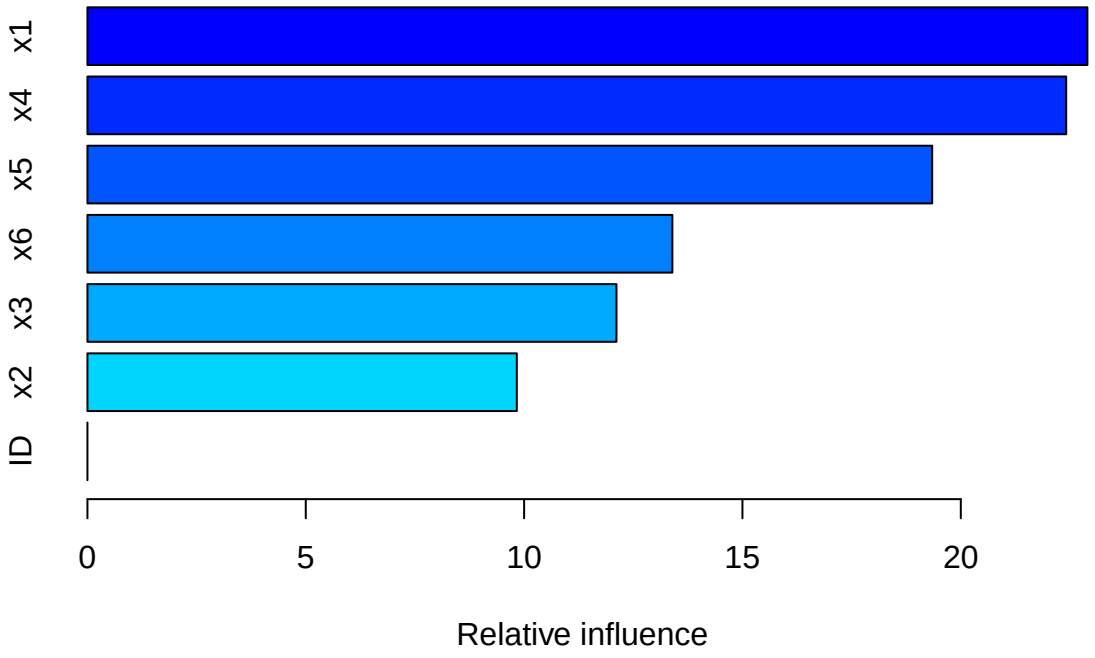
22 November 2019

Problem 1

```
rm(list=ls())
# Boosting and bagging are similar because they both use the same process of bootstrapping on the data
```

Problem 2

```
library(gbm)
data <- read.csv(file="/Users/newuser/Desktop/Notes/Undergraduate/STAT 338 - Predictive Analytics/HW6.cs
names(data)[1] <- "ID"
gbm <- gbm(Y~.,distribution="gaussian",data=data,shrinkage=.01,cv.folds=10)
summary(gbm)
```




```
##    var   rel.inf
## x1  x1 22.899136
## x4  x4 22.413446
## x5  x5 19.344956
## x6  x6 13.394936
```

```
## x3  x3 12.114763
## x2  x2  9.832762
## ID  ID  0.000000
```

```
gbm01 <- gbm(Y~.,distribution="gaussian",data=data,shrinkage=.001,cv.folds=10)
gbm05 <- gbm(Y~.,distribution="gaussian",data=data,shrinkage=.005,cv.folds=10)
gbm15 <- gbm(Y~.,distribution="gaussian",data=data,shrinkage=.015,cv.folds=10)
gbm2 <- gbm(Y~.,distribution="gaussian",data=data,shrinkage=.02,cv.folds=10)
gbm25 <- gbm(Y~.,distribution="gaussian",data=data,shrinkage=.025,cv.folds=10)
gbm3 <- gbm(Y~.,distribution="gaussian",data=data,shrinkage=.03,cv.folds=10)
gbm35 <- gbm(Y~.,distribution="gaussian",data=data,shrinkage=.035,cv.folds=10)
gbm4 <- gbm(Y~.,distribution="gaussian",data=data,shrinkage=.04,cv.folds=10)
gbm45 <- gbm(Y~.,distribution="gaussian",data=data,shrinkage=.045,cv.folds=10)
gbm5 <- gbm(Y~.,distribution="gaussian",data=data,shrinkage=.05,cv.folds=10)
gbm1 <- gbm(Y~.,distribution="gaussian",data=data,shrinkage=.1,cv.folds=10)
plot(c(.001,.005,.01,.015,.02,.025,.03,.035,.04,.045,.05,.1),c(sum(data$Y-predict(gbm01,n.trees=100))^2
```