Homework5 -Kelly Markor

install.package(“ISLR”) install.package(“tree”) library(ISLR) library(tree)

8.3.1 attach(Carseats) High=ifelse(Sales<=8,“No”,“Yes”) Carseats=data.frame(Carseats,High) tree.carseats=tree(High~.-Sales,Carseats) summary(tree.carseats)

plot(tree.carseats) text(tree.carseats,pretty=0) tree.carseats

set.seed(2) train=sample(1:nrow(Carseats), 200) Carseats.test=Carseats[-train,] High.test=High[-train] tree.carseats=tree(High~.-Sales,Carseats,subset=train) tree.pred=predict(tree.carseats,Carseats.test,type=“class”) table(tree.pred,High.test) (86+57)/200

set.seed(3) cv.carseats=cv.tree(tree.carseats,FUN=prune.misclass) names(cv.carseats) cv.carseats

par(mfrow=c(1,2)) plot(cv.carseatsdev,type=“b”) plot(cv.carseatsdev,type=“b”) prune.carseats=prune.misclass(tree.carseats,best=9) plot(prune.carseats) text(prune.carseats,pretty=0)

tree.pred=predict(prune.carseats,Carseats.test,type=“class”) table(tree.pred,High.test) (94+60)/200

prune.carseats=prune.misclass(tree.carseats,best=15) plot(prune.carseats) text(prune.carseats,pretty=0) tree.pred=predict(prune.carseats,Carseats.test,type=“class”) table(tree.pred,High.test)

\*\*8.3.2 install.package(“MASS”) library(MASS) set.seed(1) train = sample(1:nrow(Boston), nrow(Boston)/2) tree.boston=tree(medv~.,Boston,subset=train) summary(tree.boston) plot(tree.boston) text(tree.boston ,pretty=0) cv.boston=cv.tree(tree.boston) plot(cv.bostondev,type=‘b’) prune.boston=prune.tree(tree.boston,best=5) plot(prune.boston) text(prune.boston,pretty=0) yhat=predict(tree.boston,newdata=Boston[-train,]) boston.test=Boston[-train,“medv”] plot(yhat,boston.test) abline(0,1) mean((yhat-boston.test)^2)

\*\*8.3.3

library(randomForest) set.seed(1) bag.boston=randomForest(medv~.,data=Boston,subset=train,mtry=13, ntree=25) yhat = predict(bag.boston,newdata=Boston[-train,]) set.seed(1) rf.boston=randomForest(medv~.,data=Boston,subset=train,mtry=6,importance=TRUE) yhat.rf=predict(rf.boston,newdata=Boston[-train,]) mean((yhat.rf=boston.test)^2) importance(rf.boston)

\*\*8.3.4 library(gbm) set.seed(1) boost.boston=gbm(medv~.,data=Boston[train,],distribution=“gaussian”,n.trees=5000,interaction.depth=4) summary(boost.boston) par(mfrow=c(1,2)) plot(boost.boston,i=“rm”) plot(boost.boston,i=“lstat”) yhat.boost=predict(boost.bostoin,newdata=Boston[-train,] ,n.trees=5000) mean((yhat.boost-boston.test)2) boost.boston=gbm(medv~.,data=Boston[train,],distribution=“gaussian”,n.trees=5000,interaction.depth=4,shrinkage=0.2,verbose=F) yhat.boost=predict(boost.boston,newdata=Boston[-train,],n.trees=5000) mean((yhat.boost=boston.test)2)