Appendix

Conner McCloney

#SVMs  
suppressMessages(library(e1071))  
subset <- pulsar\_data[1:1000,]  
pulsar\_data$target\_class <- as.factor(pulsar\_data$target\_class)  
subset$target\_class <- as.factor(subset$target\_class)  
svmfit1=svm(target\_class~.,data=subset,kernel="linear",cost=10,scale=FALSE)  
summary(svmfit1)

##   
## Call:  
## svm(formula = target\_class ~ ., data = subset, kernel = "linear",   
## cost = 10, scale = FALSE)  
##   
##   
## Parameters:  
## SVM-Type: C-classification   
## SVM-Kernel: linear   
## cost: 10   
##   
## Number of Support Vectors: 47  
##   
## ( 25 22 )  
##   
##   
## Number of Classes: 2   
##   
## Levels:   
## 0 1

ypred1=predict(svmfit1,pulsar\_data)  
table(predict=ypred1,pulsar\_data$target\_class)

##   
## predict 0 1  
## 0 16195 314  
## 1 64 1325

svmfit2=svm(target\_class~.,data=subset,kernel="radial",gamma=1,cost=1)  
summary(svmfit2)

##   
## Call:  
## svm(formula = target\_class ~ ., data = subset, kernel = "radial",   
## gamma = 1, cost = 1)  
##   
##   
## Parameters:  
## SVM-Type: C-classification   
## SVM-Kernel: radial   
## cost: 1   
##   
## Number of Support Vectors: 269  
##   
## ( 196 73 )  
##   
##   
## Number of Classes: 2   
##   
## Levels:   
## 0 1

ypred2=predict(svmfit2,pulsar\_data)  
table(predict=ypred2,pulsar\_data$target\_class)

##   
## predict 0 1  
## 0 16228 673  
## 1 31 966

cat("Tree-based - Bagging\n")

## Tree-based - Bagging

suppressMessages(library(randomForest))  
train<-sample(1:nrow(pulsar\_data),1000)  
bag=randomForest(target\_class~.,data=pulsar\_data,subset=train,mtry=8,importance=TRUE)  
bag

##   
## Call:  
## randomForest(formula = target\_class ~ ., data = pulsar\_data, mtry = 8, importance = TRUE, subset = train)   
## Type of random forest: classification  
## Number of trees: 500  
## No. of variables tried at each split: 8  
##   
## OOB estimate of error rate: 2.8%  
## Confusion matrix:  
## 0 1 class.error  
## 0 893 8 0.008879023  
## 1 20 79 0.202020202

cat("Tree-based - RF\n")

## Tree-based - RF

RF=randomForest(target\_class~.,data=pulsar\_data,subset=train,mtry=4,importance=TRUE)  
RF

##   
## Call:  
## randomForest(formula = target\_class ~ ., data = pulsar\_data, mtry = 4, importance = TRUE, subset = train)   
## Type of random forest: classification  
## Number of trees: 500  
## No. of variables tried at each split: 4  
##   
## OOB estimate of error rate: 2.9%  
## Confusion matrix:  
## 0 1 class.error  
## 0 892 9 0.009988901  
## 1 20 79 0.202020202

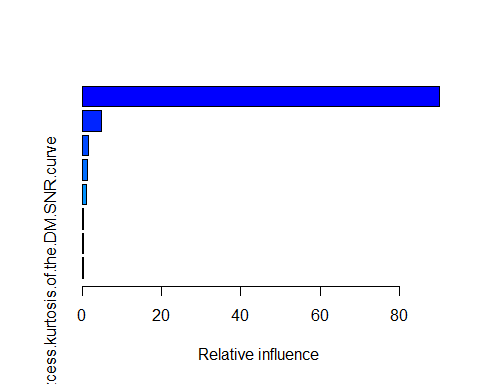
cat("Tree-based - Boosting\n")

## Tree-based - Boosting

suppressMessages(library(gbm))  
Boost=gbm(target\_class~.,data=pulsar\_data[train,],distribution = "gaussian")  
class(pulsar\_data$target\_class)

## [1] "factor"

summary(Boost)



## var  
## Excess.kurtosis.of.the.integrated.profile Excess.kurtosis.of.the.integrated.profile  
## Skewness.of.the.integrated.profile Skewness.of.the.integrated.profile  
## Standard.deviation.of.the.DM.SNR.curve Standard.deviation.of.the.DM.SNR.curve  
## Standard.deviation.of.the.integrated.profile Standard.deviation.of.the.integrated.profile  
## Mean.of.the.integrated.profile Mean.of.the.integrated.profile  
## Mean.of.the.DM.SNR.curve Mean.of.the.DM.SNR.curve  
## Skewness.of.the.DM.SNR.curve Skewness.of.the.DM.SNR.curve  
## Excess.kurtosis.of.the.DM.SNR.curve Excess.kurtosis.of.the.DM.SNR.curve  
## rel.inf  
## Excess.kurtosis.of.the.integrated.profile 90.1402619  
## Skewness.of.the.integrated.profile 4.8414136  
## Standard.deviation.of.the.DM.SNR.curve 1.5989070  
## Standard.deviation.of.the.integrated.profile 1.4075398  
## Mean.of.the.integrated.profile 1.0516628  
## Mean.of.the.DM.SNR.curve 0.4215940  
## Skewness.of.the.DM.SNR.curve 0.2909789  
## Excess.kurtosis.of.the.DM.SNR.curve 0.2476419

yhat.boost <- predict(Boost,newdata=pulsar\_data[-train,],n.trees=100)  
pulsar\_data$target\_class <- as.numeric(pulsar\_data$target\_class)  
boost.test=pulsar\_data[-train,"target\_class"]  
cat("Error rate: ",mean((yhat.boost-boost.test)^2))

## Error rate: 0.0198858