

Data 621 - Homework #2

Joseph E. Garcia

March 14, 2019

```
library(pROC)

## Warning: package 'pROC' was built under R version 3.5.3
## Type 'citation("pROC")' for a citation.
##
## Attaching package: 'pROC'
##
## The following objects are masked from 'package:stats':
##
##      cov, smooth, var

library(caret)

## Warning: package 'caret' was built under R version 3.5.3
## Loading required package: lattice
## Loading required package: ggplot2
## Warning: package 'ggplot2' was built under R version 3.5.3

library(ggplot2)

data<-read.csv("https://raw.githubusercontent.com/jgarcia71/Data-621-
Homeworks-Spring-2019/master/classification-output-data.csv")
head(data)

##   pregnant glucose diastolic skinfold insulin  bmi pedigree age class
## 1         7     124         70      33     215 25.5    0.161  37     0
## 2         2     122         76      27     200 35.9    0.483  26     0
## 3         3     107         62      13      48 22.9    0.678  23     1
## 4         1      91         64      24       0 29.2    0.192  21     0
## 5         4      83         86      19       0 29.3    0.317  34     0
## 6         1     100         74      12      46 19.5    0.149  28     0
##   scored.class scored.probability
## 1           0      0.32845226
## 2           0      0.27319044
## 3           0      0.10966039
## 4           0      0.05599835
## 5           0      0.10049072
## 6           0      0.05515460
```

```

tab<-table(data$class,data$scored.class)
colnames(tab)<-c("Real Negative","Real Positive")
rownames(tab)<-c("Model Negative","Model Positive")
tab

```

```

##
##           Real Negative Real Positive
## Model Negative         119           5
## Model Positive          30          27

```

```

Confus.Matrix <- function(data) {
  tab <- table(data$class,data$scored.class)
  colnames(tab)<-c("Real Negative","Real Positive")
  rownames(tab)<-c("Model Negative","Model Positive")

  return(tab)
}

```

```

Confus.Matrix(data=data)

```

```

##
##           Real Negative Real Positive
## Model Negative         119           5
## Model Positive          30          27

```

```

Sens<-function(data) {
  tab <- table(data$class,data$scored.class)
  tn<-tab[1,1]
  tp<-tab[2,2]
  fn<-tab[2,1]
  fp<-tab[1,2]

  sens<-tn/(tn+fp)

  return(sens)
}

```

```

Spec<-function(data) {
  tab <- table(data$class,data$scored.class)
  tn<-tab[1,1]
  tp<-tab[2,2]
  fn<-tab[2,1]
  fp<-tab[1,2]

  spec<-tp/(tp+fn)

  return(spec)
}

```

```

CER<-function(data) {
  tab <- table(data$class,data$scored.class)
  tn<-tab[1,1]
  tp<-tab[2,2]
  fn<-tab[2,1]
  fp<-tab[1,2]

  cer<-(fp+fn)/(tp+tn+fn+fp)

  return(cer)
}

PREC<-function(data) {
  tab <- table(data$class,data$scored.class)
  tn<-tab[1,1]
  tp<-tab[2,2]
  fn<-tab[2,1]
  fp<-tab[1,2]

  prec<-tp/(tp+fp)

  return(prec)
}

ACC<-function(data) {
  tab <- table(data$class,data$scored.class)
  tn<-tab[1,1]
  tp<-tab[2,2]
  fn<-tab[2,1]
  fp<-tab[1,2]

  acc<-(tp+tn)/(tp+tn+fn+fp)

  return(acc)
}

F1<-function(data) {
  tab <- table(data$class,data$scored.class)
  tn<-tab[1,1]
  tp<-tab[2,2]
  fn<-tab[2,1]
  fp<-tab[1,2]

  acc<-(tp+tn)/(tp+tn+fn+fp)
  cer<-(fp+fn)/(tp+tn+fn+fp)

```

```

prec<-tp/(tp+fp)
spec<-tp/(tp+fn)
sens<-tn/(tn+fp)
f1<-2*prec*sens/(prec+sens)

return(f1)
}

Metrics<-function(data) {
  tab <- table(data$class,data$score.class)
  tn<-tab[1,1]
  tp<-tab[2,2]
  fn<-tab[2,1]
  fp<-tab[1,2]

  acc<-(tp+tn)/(tp+tn+fn+fp)
  cer<-(fp+fn)/(tp+tn+fn+fp)
  prec<-tp/(tp+fp)
  spec<-tp/(tp+fn)
  sens<-tn/(tn+fp)
  f1<-2*prec*sens/(prec+sens)

  results<-data.frame(list(acc,cer,prec,sens,spec,f1))
  colnames(results) <- c("Accuracy","Classification Error
Rate","Precision","Sensitivity","Specificity","F1 Score")
  results<-data.frame(t(results))
  colnames(results)<-"Result"
  return(results)
}

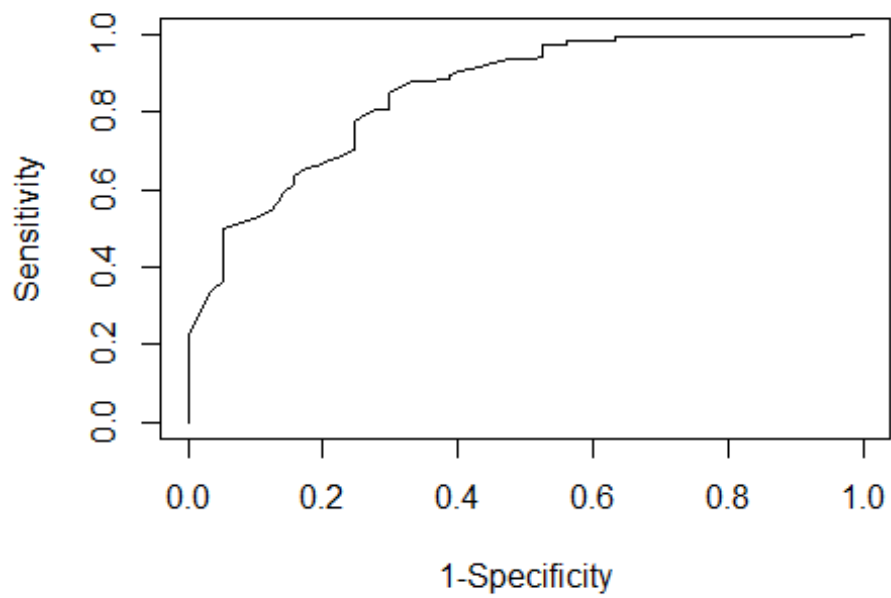
Metrics(data)

##                Result
## Accuracy          0.8066298
## Classification Error Rate 0.1933702
## Precision          0.8437500
## Sensitivity         0.9596774
## Specificity         0.4736842
## F1 Score           0.8979877

ROC_Scott<- function(data,t) {
  se<-0
  sp<-0
  a<-0
  for (i in 1:round(1/t))
  {

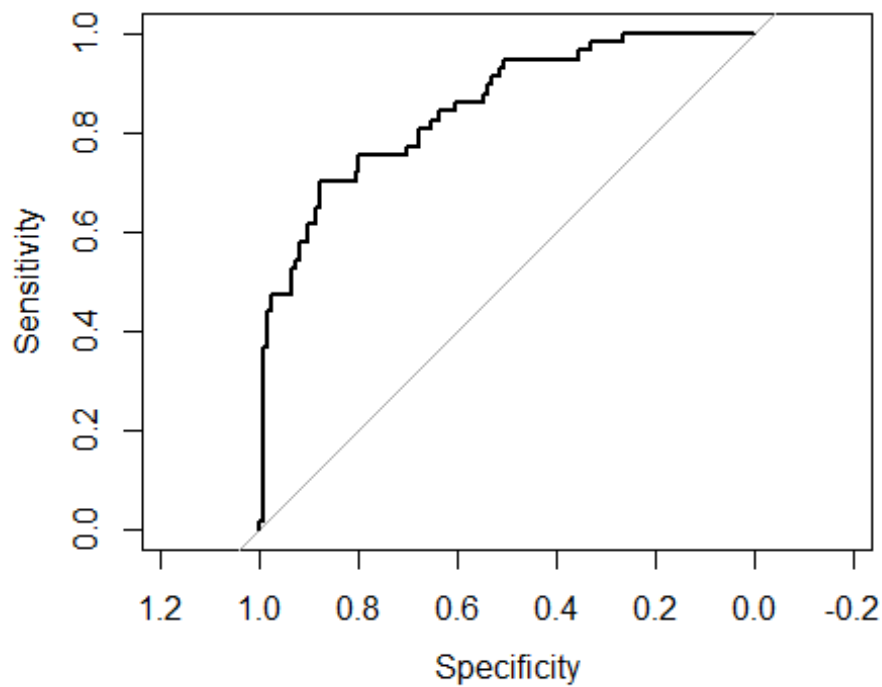
```


Scott ROC



```
## [1] "AUC=0.854"
```

```
roc(data$class,data$scored.probability,plot=TRUE)
```



```
##  
## Call:  
## roc.default(response = data$class, predictor = data$scored.probability,  
plot = TRUE)  
##  
## Data: data$scored.probability in 124 controls (data$class 0) < 57 cases  
(data$class 1).  
## Area under the curve: 0.8503
```