

# MelanomavsSeborreic

Jose Tamez

4/24/2024

- 1 Melanoma vs Seborreic\_Keratosis
  - 1.1 Loading data sets
  - 1.2 Data Conditioning
  - 1.3 The Heatmap
  - 1.4 Learning Melanoma with KNN
  - 1.5 KNN plot performance
  - 1.6 SVM
  - 1.7 SVM plot performance
  - 1.8 Learning Melanoma with Naive Bayes with PCA
  - 1.9 NB performance
  - 1.10 Learning Melanoma with LASSO
  - 1.11 LASSO performance

## 1 Melanoma vs Seborreic\_Keratosis

### 1.1 Loading data sets

```
# /Users/joseantoniopalaciosdorbecker/Documents/Carrera/Semestre 6/Materias/Procesamiento/R/  
  
MelanomaFeatures <- read.csv("/Users/joseantoniopalaciosdorbecker/Documents/Carrera/Semestre 6/Materias/Procesamiento/R/Evidencia1/MelanomaLesionFeatures.csv", header=FALSE)  
MelanomaControlFeatures <- read.csv("/Users/joseantoniopalaciosdorbecker/Documents/Carrera/Semestre 6/Materias/Procesamiento/R/Evidencia1/MelanomaControlFeatures.csv", header=FALSE)  
  
SeborrheicFeatures <- read.csv("/Users/joseantoniopalaciosdorbecker/Documents/Carrera/Semestre 6/Materias/Procesamiento/R/Evidencia1/SeborrheicLesionFeatures.csv", header=FALSE)  
SeborrheicControlFeatures <- read.csv("/Users/joseantoniopalaciosdorbecker/Documents/Carrera/Semestre 6/Materias/Procesamiento/R/Evidencia1/SeborrheicControlFeatures.csv", header=FALSE)  
  
NevusFeatures <- read.csv("/Users/joseantoniopalaciosdorbecker/Documents/Carrera/Semestre 6/Materias/Procesamiento/R/Evidencia1/NevusLesionFeatures.csv", header=FALSE)  
NevusControlFeatures <- read.csv("/Users/joseantoniopalaciosdorbecker/Documents/Carrera/Semestre 6/Materias/Procesamiento/R/Evidencia1/NevusControlFeatures.csv", header=FALSE)  
  
#LesionFeatureNames <- read_excel("FeatureNames.xlsx")  
#ControlFeatureNames <- read_excel("FeatureNames.xlsx", sheet = "Control")  
#FeatureNames <- read_excel("FeatureNames.xlsx", sheet = "PerChannel")
```

## 1.2 Data Conditioning

```
op <- par(no.readonly = TRUE)

colnames(MelanomaControlFeatures) <- paste("C", colnames(MelanomaControlFeatures), sep = "")
colnames(SeborrheicControlFeatures) <- paste("C", colnames(SeborrheicControlFeatures), sep = "")
colnames(NevusControlFeatures) <- paste("C", colnames(NevusControlFeatures), sep = "")

sum(is.na(MelanomaFeatures))
sum(is.na(SeborrheicFeatures))
sum(is.na(NevusFeatures))
sum(is.na(MelanomaControlFeatures))
sum(is.na(NevusControlFeatures))

channel1 <- c(1:32)
channel2 <- c(33:64)
channel3 <- c(65:96)

MMelanomaFeatures <- ( MelanomaFeatures[,channel1] +
                         MelanomaFeatures[,channel2] +
                         MelanomaFeatures[,channel3] )/3.0
SMelanomaFeatures <- ( abs(MelanomaFeatures[,channel1] - MMelanomaFeatures) +
                         abs(MelanomaFeatures[,channel2] - MMelanomaFeatures) +
                         abs(MelanomaFeatures[,channel3] - MMelanomaFeatures) )/3.0

colnames(MMelanomaFeatures) <- paste("M", colnames(MMelanomaFeatures), sep = "")
colnames(SMelanomaFeatures) <- paste("S", colnames(SMelanomaFeatures), sep = "")

MelanomaFeatures <- cbind(MelanomaFeatures, SMelanomaFeatures/(0.01+abs(MMelanomaFeatures)))

MNevusFeatures <- ( NevusFeatures[,channel1] +
                      NevusFeatures[,channel2] +
                      NevusFeatures[,channel3] )/3.0
SNevasFeatures <- ( abs(NevusFeatures[,channel1] - MNevusFeatures) +
                      abs(NevusFeatures[,channel2] - MNevusFeatures) +
                      abs(NevusFeatures[,channel3] - MNevusFeatures) )/3.0

colnames(MNevusFeatures) <- paste("M", colnames(MNevusFeatures), sep = "")
colnames(SNevusFeatures) <- paste("S", colnames(SNevusFeatures), sep = "")

NevusFeatures <- cbind(NevusFeatures, SNevusFeatures/(0.001+abs(MNevusFeatures)))

MSeborrheicFeatures <- ( SeborrheicFeatures[,channel1] +
                           SeborrheicFeatures[,channel2] +
                           SeborrheicFeatures[,channel3] )/3.0
SSeborrheicFeatures <- ( abs(SeborrheicFeatures[,channel1] - MSeborrheicFeatures) +
                           abs(SeborrheicFeatures[,channel2] - MSeborrheicFeatures) +
```

```
abs(SeborrheicFeatures[,channel3] - MSeborrheicFeatures))/3.0
colnames(MSeborrheicFeatures) <- paste("M", colnames(MSeborrheicFeatures), sep="")
colnames(SSeborrheicFeatures) <- paste("S", colnames(SSeborrheicFeatures), sep="")

SeborrheicFeatures <- cbind(SeborrheicFeatures, SSeborrheicFeatures/(0.001+abs(MSeborrheicFeatures)))

MMelanomaControlFeatures <- ( MelanomaControlFeatures[,channel1] +
                               MelanomaControlFeatures[,channel2] +
                               MelanomaControlFeatures[,channel3] )/3.0

SMelanomaControlFeatures <- (abs(MelanomaControlFeatures[,channel1] - MMelanomaControlFeatures) +
                               abs(MelanomaControlFeatures[,channel2] - MMelanomaControlFeatures) +
                               abs(MelanomaControlFeatures[,channel3] - MMelanomaControlFeatures))/3.0
colnames(MMelanomaControlFeatures) <- paste("M", colnames(MMelanomaControlFeatures), sep="")
colnames(SMelanomaControlFeatures) <- paste("S", colnames(SMelanomaControlFeatures), sep="")

MelanomaControlFeatures <- cbind(MelanomaControlFeatures, SMelanomaControlFeatures/(0.001+abs(MMelanomaControlFeatures)))

MNevusControlFeatures <- ( NevusControlFeatures[,channel1] +
                            NevusControlFeatures[,channel2] +
                            NevusControlFeatures[,channel3] )/3.0
SNevusControlFeatures <- (abs(NevusControlFeatures[,channel1] - MNevusControlFeatures) +
                           abs(NevusControlFeatures[,channel2] - MNevusControlFeatures) +
                           abs(NevusControlFeatures[,channel3] - MNevusControlFeatures))/3.0
colnames(MNevusControlFeatures) <- paste("M", colnames(MNevusControlFeatures), sep="")
colnames(SNevusControlFeatures) <- paste("S", colnames(SNevusControlFeatures), sep="")

NevusControlFeatures <- cbind(NevusControlFeatures, SNevusControlFeatures/(0.001+abs(MNevusControlFeatures)))

MSeborrheicControlFeatures <- ( SeborrheicControlFeatures[,channel1] +
                                 SeborrheicControlFeatures[,channel2] +
                                 SeborrheicControlFeatures[,channel3] )/3.0
SSeborrheicControlFeatures <- (abs(SeborrheicControlFeatures[,channel1] - MSeborrheicControlFeatures) +
                                 abs(SeborrheicControlFeatures[,channel2] - MSeborrheicControlFeatures) +
                                 abs(SeborrheicControlFeatures[,channel3] - MSeborrheicControlFeatures))/3.0
```

```
colnames(MSeborrheicControlFeatures) <- paste("M", colnames(MSeborrheicControlFeatures), sep="")
colnames(SSeborrheicControlFeatures) <- paste("S", colnames(SSeborrheicControlFeatures), sep="")

SeborrheicControlFeatures <- cbind(SeborrheicControlFeatures,
                                    SSeborrheicControlFeatures/(0.001+abs(MSeborrheicControlFeatures)))

CtrDiff <- MelanomaFeatures[,1:ncol(MelanomaControlFeatures)] - MelanomaControlFeatures;
colnames(CtrDiff) <- colnames(MelanomaControlFeatures)
MelanomaFeatures <- cbind(MelanomaFeatures, CtrDiff)

CtrDiff <- NevusFeatures[,1:ncol(MelanomaControlFeatures)] - NevusControlFeatures;
colnames(CtrDiff) <- colnames(NevusControlFeatures)
NevusFeatures <- cbind(NevusFeatures, CtrDiff)

CtrDiff <- SeborrheicFeatures[,1:ncol(MelanomaControlFeatures)] - SeborrheicControlFeatures;
colnames(CtrDiff) <- colnames(SeborrheicControlFeatures)
SeborrheicFeatures <- cbind(SeborrheicFeatures, CtrDiff)

MelanomaFeatures <- MelanomaFeatures[complete.cases(MelanomaFeatures),]
NevusFeatures <- NevusFeatures[complete.cases(NevusFeatures),]
SeborrheicFeatures <- SeborrheicFeatures[complete.cases(SeborrheicFeatures),]

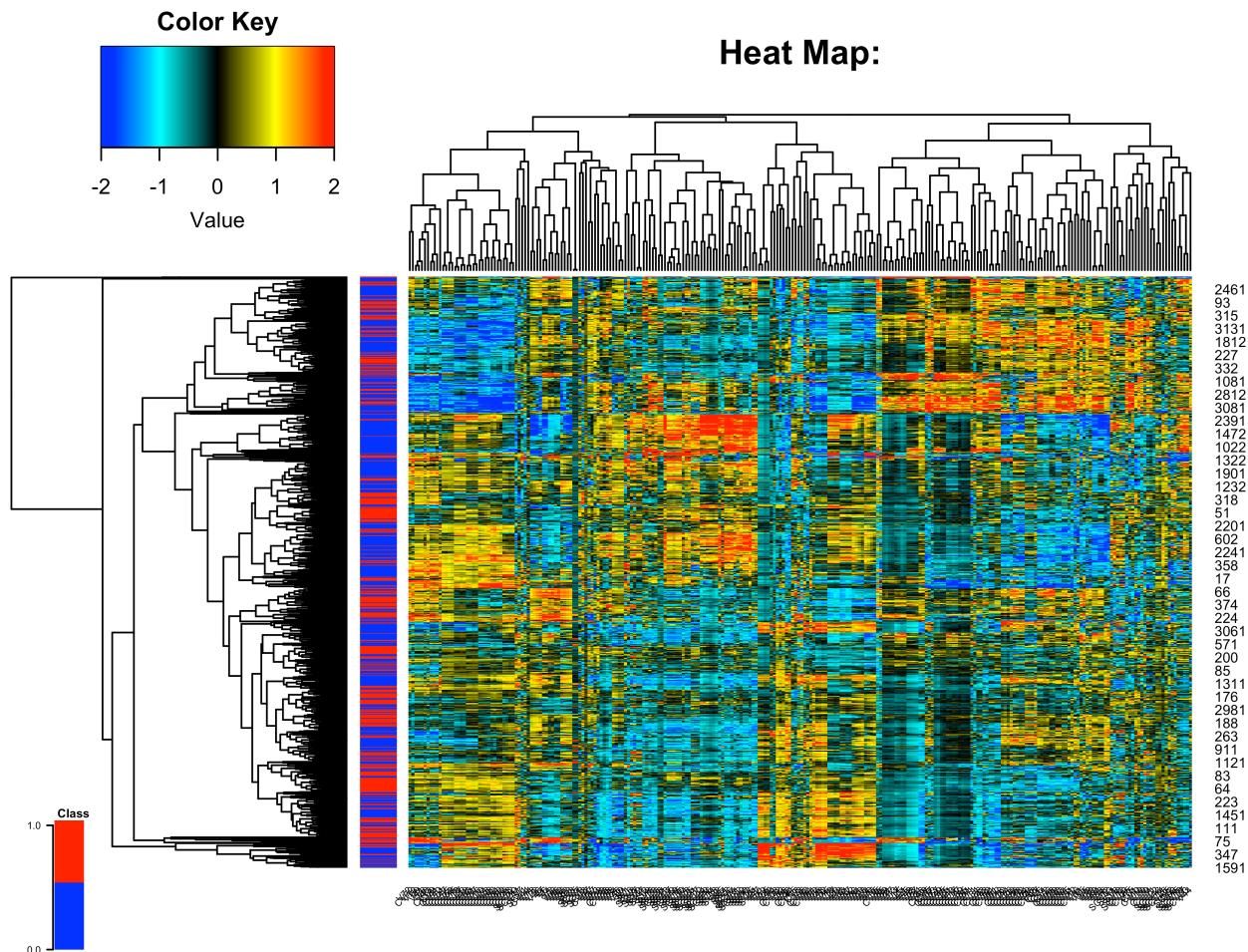
MelanomaFeatures$Class <- rep(1,nrow(MelanomaFeatures))
NevusFeatures$Class <- rep(0,nrow(NevusFeatures))
SeborrheicFeatures$Class <- rep(0,nrow(SeborrheicFeatures))

MelanomaSeborrheicNevus <- rbind(MelanomaFeatures, SeborrheicFeatures, NevusFeatures)

table(MelanomaSeborrheicNevus$Class)
```

## 1.3 The Heatmap

```
hm <- heatMaps(Outcome = "Class",
                 data = MelanomaSeborrheicNevus[1:900,],
                 title = "Heat Map:", Scale = TRUE,
                 cexRow = 0.75, cexCol = 0.5, srtCol = 45)
```



```
par(op)
```

## 1.4 Learning Melanoma with KNN

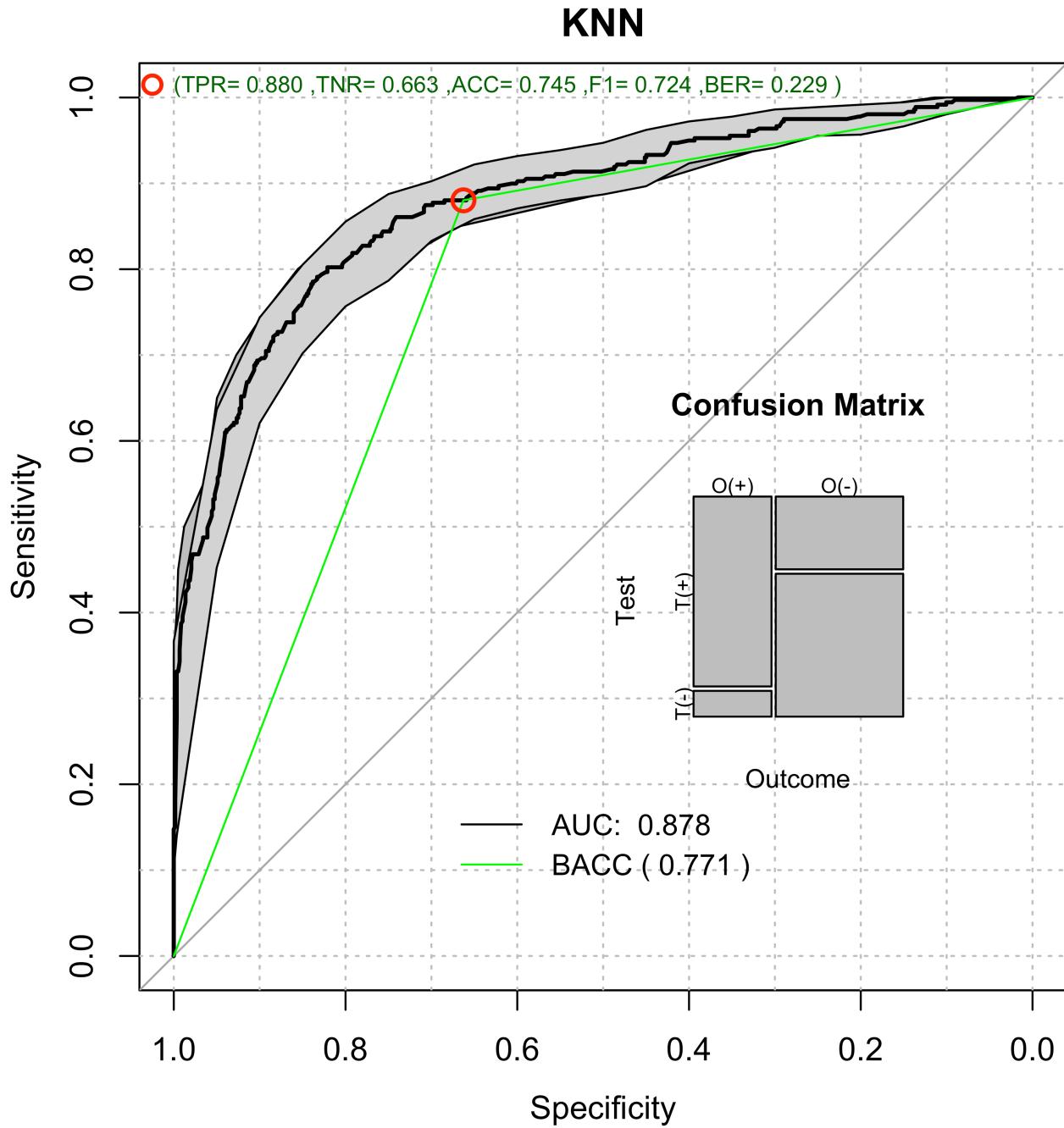
```
cvKNN <- randomCV(MelanomaSeborrheicNevus,"Class",
  KNN_method,
  trainFraction = 0.70,
  repetitions = 10,
  classSamplingType = "Ba",
  featureSelectionFunction = univariate_KS,
  featureSelection.control = list(limit= 30), # que nos regrese los 30 m
as significativos
)
#> .....10 Tested: 946 Avg. Selected: 28.7 Min Tests: 1 Max Tests: 10 Mean Test
s: 4.630021 . MAD: 0.3478308
#>

# si se entrena balanceado, el punto donde hay cancer o no es de 0.5
# el criterio de decisión del sistema depende de como fue entrenado el sistema
# es mejor entrenar con todos los datos

#MAD – mean absolute difference
#Area Under Curve (AUC) representa el intervalo de confianza
# BACC – Balance Accuracy : Sensibility + Specificity / 2, area bajo curva verde
```

## 1.5 KNN plot performance

```
performace <- predictionStats_binary(cvKNN$medianTest,"KNN")
```



```
par(op)
pander::pander(performace$aucs)
```

est	lower	upper
0.8783	0.8547	0.9018

```
pander::pander(performace$accc)
```

est	lower	upper

<b>est</b>	<b>lower</b>	<b>upper</b>
0.7452	0.7162	0.7727

```
pander::pander(performace$CM.analysis)
```

- **detail:**

<b>statistic</b>	<b>est</b>	<b>lower</b>	<b>upper</b>
ap	0.5433	0.511	0.5754
tp	0.3795	0.3485	0.4113
se	0.8802	0.8421	0.9119
sp	0.6627	0.6228	0.7009
diag.ac	0.7452	0.7162	0.7727
diag.or	14.44	10.06	20.73
nndx	1.842	1.632	2.151
youden	0.5429	0.4649	0.6128
pv.pos	0.6148	0.5712	0.6571
pv.neg	0.9005	0.8683	0.927
lr.pos	2.61	2.315	2.941
lr.neg	0.1807	0.1357	0.2407
p.rout	0.4567	0.4246	0.489
p.rin	0.5433	0.511	0.5754
p.tpdn	0.3373	0.2991	0.3772
p.tndp	0.1198	0.08805	0.1579
p.dntp	0.3852	0.3429	0.4288
p.dptn	0.09954	0.07298	0.1317

- **tab:**

	<b>Outcome +</b>	<b>Outcome -</b>	<b>Total</b>
<b>Test +</b>	316	198	514
<b>Test -</b>	43	389	432
<b>Total</b>	359	587	946

- **method:** exact

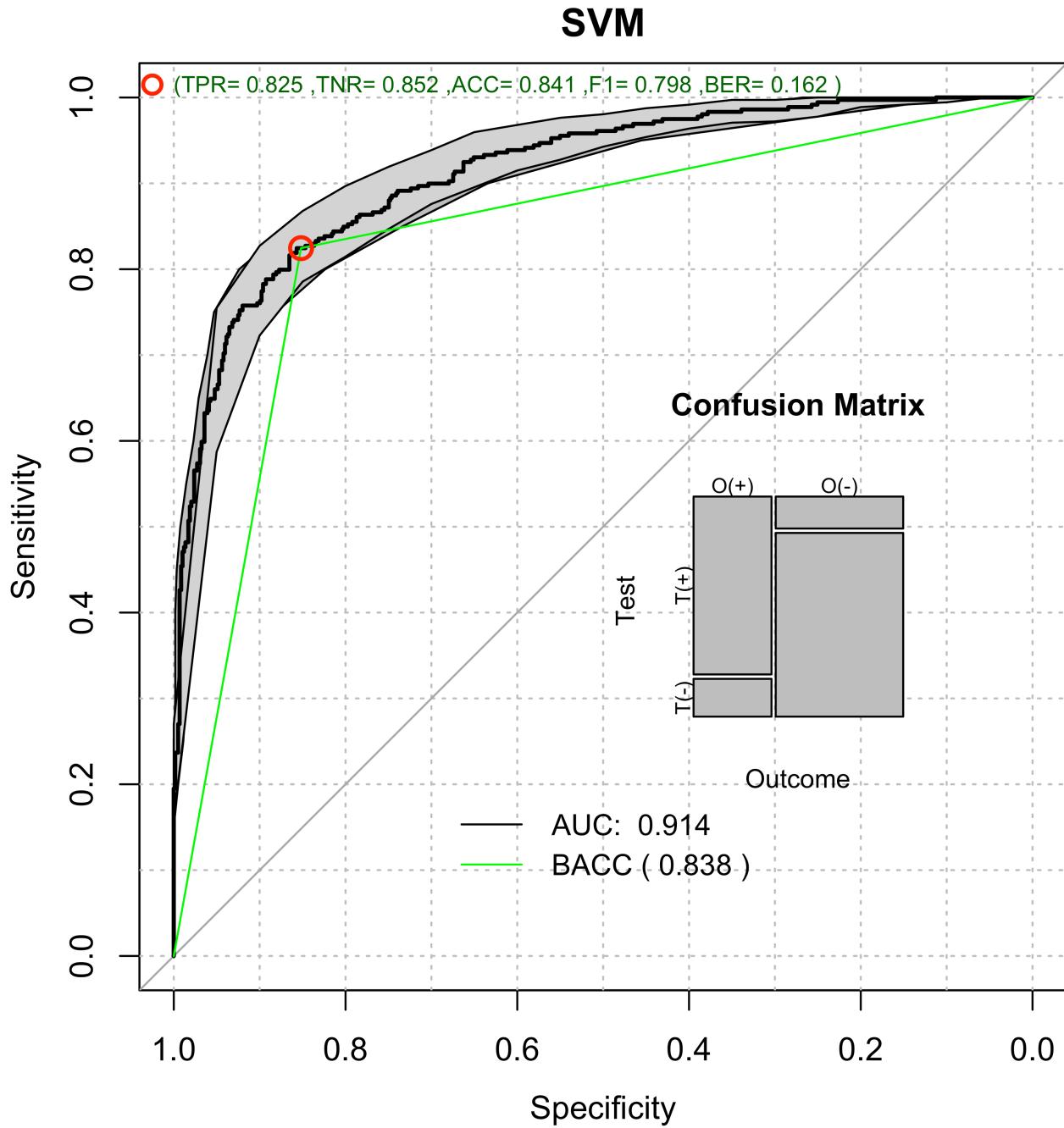
- **digits:** 2
- **conf.level:** 0.95

## 1.6 SVM

```
cvSVM <- randomCV(MelanomaSeborrheicNevus,"Class",
                     e1071::svm,
                     asFactor = TRUE,
                     trainSampleSets= cvKNN$trainSamplesSets,
                     featureSelectionFunction = cvKNN$selectedFeaturesSet,
                     probability=TRUE
)
#> .....10  Tested: 946 Avg. Selected: 28.7 Min Tests: 1 Max Tests: 10 Mean Test
s: 4.630021 . MAD: 0.2446622
#>
```

## 1.7 SVM plot performance

```
performace <- predictionStats_binary(cvSVM$medianTest,"SVM")
```



```
par(op)
pander::pander(performace$aucs)
```

est	lower	upper
0.9139	0.8952	0.9326

```
pander::pander(performace$accc)
```

est	lower	upper

<b>est</b>	<b>lower</b>	<b>upper</b>
0.8414	0.8166	0.8642

```
pander::pander(performace$CM.analysis)
```

- **detail:**

<b>statistic</b>	<b>est</b>	<b>lower</b>	<b>upper</b>
ap	0.4049	0.3734	0.4369
tp	0.3795	0.3485	0.4113
se	0.8245	0.7811	0.8624
sp	0.8518	0.8204	0.8796
diag.ac	0.8414	0.8166	0.8642
diag.or	27	18.94	38.5
nndx	1.479	1.348	1.662
youden	0.6763	0.6016	0.742
pv.pos	0.7728	0.7276	0.8139
pv.neg	0.8881	0.8591	0.9129
lr.pos	5.563	4.556	6.793
lr.neg	0.206	0.1642	0.2585
p.rout	0.5951	0.5631	0.6266
p.rin	0.4049	0.3734	0.4369
p.tpdn	0.1482	0.1204	0.1796
p.tndp	0.1755	0.1376	0.2189
p.dntp	0.2272	0.1861	0.2724
p.dptn	0.1119	0.08707	0.1409

- **tab:**

	<b>Outcome +</b>	<b>Outcome -</b>	<b>Total</b>
<b>Test +</b>	296	87	383
<b>Test -</b>	63	500	563
<b>Total</b>	359	587	946

- **method:** exact

- **digits:** 2
- **conf.level:** 0.95

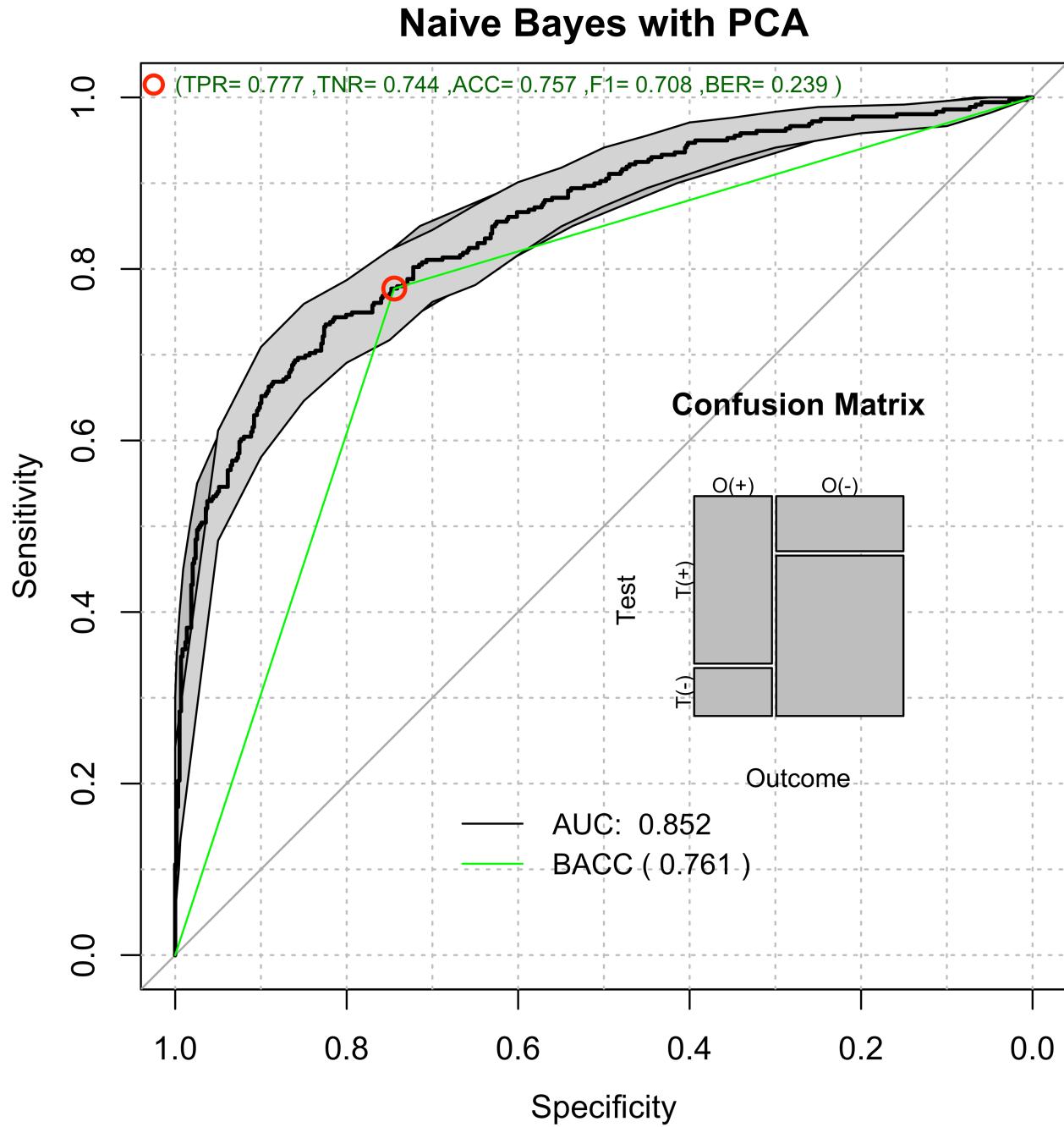
```
# curva roc  
  
# sensitivity: Ppred / Preal  
# Especificidad = Npred / Neg_real  
  
# circulo rojo es un predictor continuo, curva rock
```

## 1.8 Learning Melanoma with Naive Bayes with PCA

```
cvNB <- randomCV(MelanomaSeborrheicNevus,"Class",  
                    NAIVE_BAYES,  
                    trainSampleSets= cvKNN$trainSamplesSets,  
                    featureSelectionFunction = cvKNN$selectedFeaturesSet  
                    )  
#> .....10 Tested: 946 Avg. Selected: 28.7 Min Tests: 1 Max Tests: 10 Mean Test  
s: 4.630021 . MAD: 0.2874876  
#>
```

## 1.9 NB performance

```
performace <- predictionStats_binary(cvNB$medianTest,"Naive Bayes with PCA")
```



```
par(op)
pander::pander(performace$aucs)
```

est	lower	upper
0.8519	0.8261	0.8777

```
pander::pander(performace$accc)
```

est	lower	upper

<b>est</b>	<b>lower</b>	<b>upper</b>
0.7569	0.7282	0.7839

```
pander::pander(performace$CM.analysis)
```

- **detail:**

<b>statistic</b>	<b>est</b>	<b>lower</b>	<b>upper</b>
ap	0.4535	0.4214	0.4858
tp	0.3795	0.3485	0.4113
se	0.7772	0.7305	0.8192
sp	0.7445	0.7072	0.7793
diag.ac	0.7569	0.7282	0.7839
diag.or	10.16	7.451	13.85
nndx	1.917	1.671	2.285
youden	0.5216	0.4377	0.5985
pv.pos	0.6503	0.6031	0.6955
pv.neg	0.8453	0.8112	0.8754
lr.pos	3.041	2.621	3.529
lr.neg	0.2993	0.2453	0.3652
p.rout	0.5465	0.5142	0.5786
p.rin	0.4535	0.4214	0.4858
p.tpdn	0.2555	0.2207	0.2928
p.tndp	0.2228	0.1808	0.2695
p.dntp	0.3497	0.3045	0.3969
p.dptn	0.1547	0.1246	0.1888

- **tab:**

	<b>Outcome +</b>	<b>Outcome -</b>	<b>Total</b>
<b>Test +</b>	279	150	429
<b>Test -</b>	80	437	517
<b>Total</b>	359	587	946

- **method:** exact

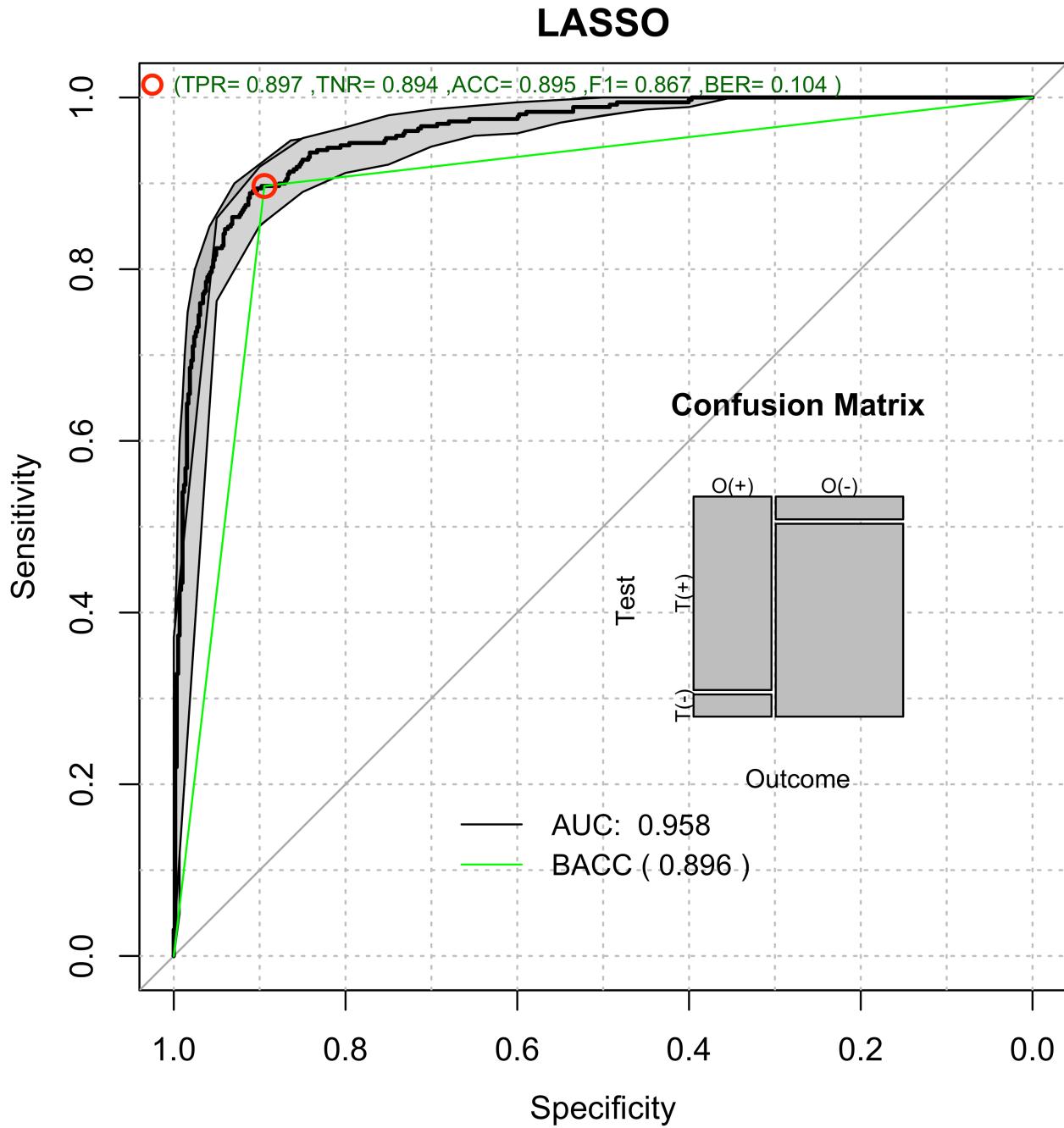
- **digits:** 2
- **conf.level:** 0.95

## 1.10 Learning Melanoma with LASSO

```
cvLASSO <- randomCV(MelanomaSeborrheicNevus,"Class",
                      LASSO_MIN,
                      trainSampleSets= cvKNN$trainSamplesSets,
                      family="binomial")
#> .....10  Tested: 946 Avg. Selected: 75.8 Min Tests: 1 Max Tests: 10 Mean Test
s: 4.630021 . MAD: 0.1503
#>
```

## 1.11 LASSO performance

```
performace <- predictionStats_binary(cvLASSO$medianTest,"LASSO")
```



```
par(op)
pander::pander(performace$aucs)
```

est	lower	upper
0.9577	0.9456	0.9697

```
pander::pander(performace$accc)
```

est	lower	upper

<b>est</b>	<b>lower</b>	<b>upper</b>
0.8953	0.8741	0.9141

```
pander::pander(performace$CM.analysis)
```

- **detail:**

<b>statistic</b>	<b>est</b>	<b>lower</b>	<b>upper</b>
ap	0.4059	0.3744	0.438
tp	0.3795	0.3485	0.4113
se	0.8969	0.8607	0.9264
sp	0.8944	0.8666	0.9181
diag.ac	0.8953	0.8741	0.9141
diag.or	73.69	47.93	113.3
nndx	1.264	1.184	1.375
youden	0.7913	0.7274	0.8444
pv.pos	0.8385	0.7979	0.8739
pv.neg	0.9342	0.9104	0.9532
lr.pos	8.492	6.693	10.77
lr.neg	0.1152	0.08482	0.1566
p.rout	0.5941	0.562	0.6256
p.rin	0.4059	0.3744	0.438
p.tpdn	0.1056	0.08194	0.1334
p.tndp	0.1031	0.07361	0.1393
p.dntp	0.1615	0.1261	0.2021
p.dptn	0.06584	0.04677	0.08961

- **tab:**

	<b>Outcome +</b>	<b>Outcome -</b>	<b>Total</b>
<b>Test +</b>	322	62	384
<b>Test -</b>	37	525	562
<b>Total</b>	359	587	946

- **method:** exact

- **digits:** 2
- **conf.level:** 0.95

```
barplot(cvLASSO$featureFrequency, las=2)
```

