

Reporte - Análisis Melanoma vs. Ketosis Seborreica

Equipo CT

17/03/2023

Melanoma vs Seborreic_Keratosis

Loading data sets

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

MelanomaFeatures <- read.csv("/Users/juanluis/Downloads/LearningMelanoma/MelanomaLesionFeatures.csv", h
SeborrheicFeatures <- read.csv("/Users/juanluis/Downloads/LearningMelanoma/SeborrheicLesionFeatures.csv
MelanomaControlFeatures <- read.csv("/Users/juanluis/Downloads/LearningMelanoma/MelanomaControlFeatures
SeborrheicControlFeatures <- read.csv("/Users/juanluis/Downloads/LearningMelanoma/SeborrheicControlFeatu
LesionFeatureNames <- read_excel("/Users/juanluis/Downloads/LearningMelanoma/FeatureNames.xlsx")
ControlFeatureNames <- read_excel("/Users/juanluis/Downloads/LearningMelanoma/FeatureNames.xlsx",sheet = "PerC
FeatureNames <- read_excel("/Users/juanluis/Downloads/LearningMelanoma/FeatureNames.xlsx",sheet = "PerC

colnames(MelanomaFeatures) <- LesionFeatureNames$FinalName
colnames(SeborrheicFeatures) <- LesionFeatureNames$FinalName
colnames(MelanomaControlFeatures) <- ControlFeatureNames$FinalName
colnames(SeborrheicControlFeatures) <- ControlFeatureNames$FinalName

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

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",FeatureNames$Features,sep="")
colnames(SMelanomaFeatures) <- paste("S",FeatureNames$Features,sep="")

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

```

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",FeatureNames$Features,sep="")
colnames(SSeborrheicFeatures) <- paste("S",FeatureNames$Features,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_C",FeatureNames$Features,sep="")
colnames(SMelanomaControlFeatures) <- paste("S_C",FeatureNames$Features,sep="")

MelanomaControlFeatures <- cbind(MelanomaControlFeatures,SMelanomaControlFeatures/(0.001+abs(MMelanomaC

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_C",FeatureNames$Features,sep="")
colnames(SSeborrheicControlFeatures) <- paste("S_C",FeatureNames$Features,sep="")

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

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

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

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

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

```

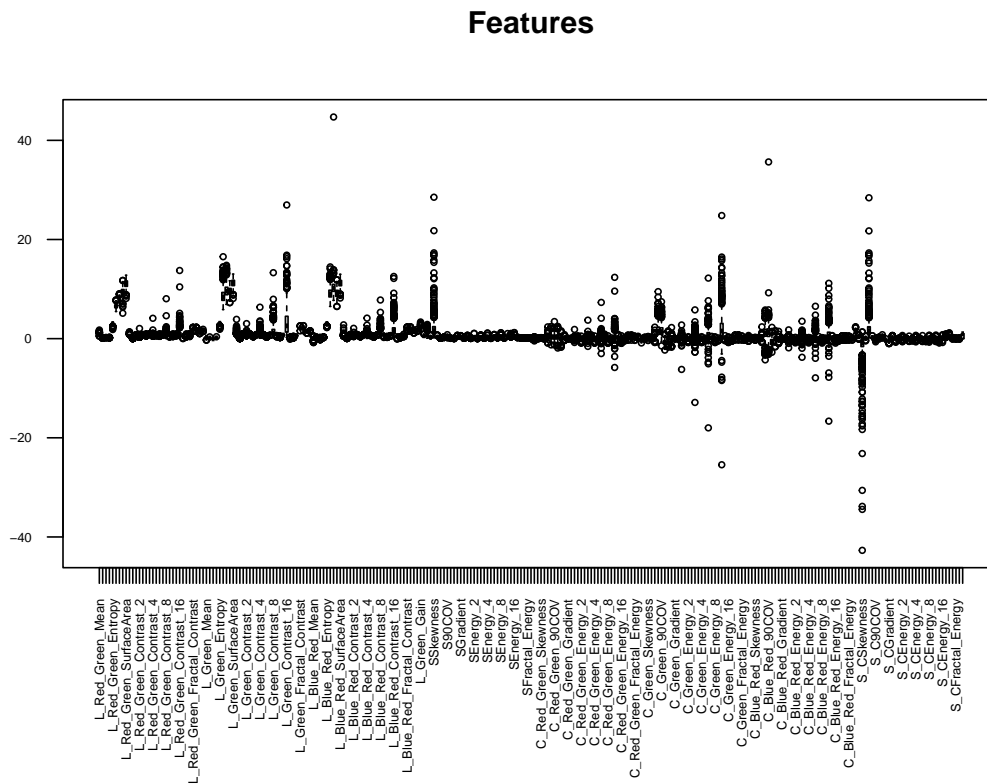
```
MelanomaSeborrheic <- rbind(MelanomaFeatures,SeborrheicFeatures)

table(MelanomaSeborrheic$Class)
```

The Heatmap

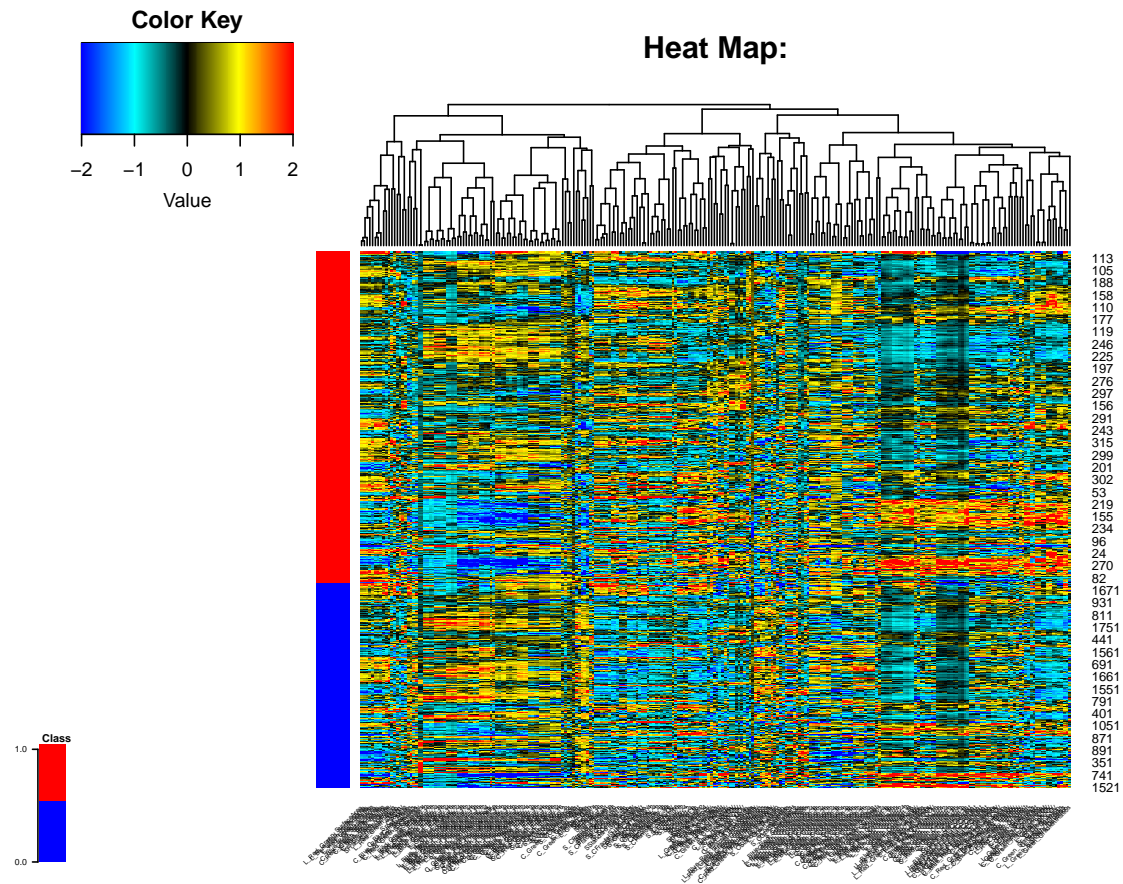
```
par(op)

par(mar=c(10,5,5,5))
boxplot(MelanomaSeborrheic,las=2,cex=0.5,cex.axis=0.5,main="Features")
```



```
par(op)

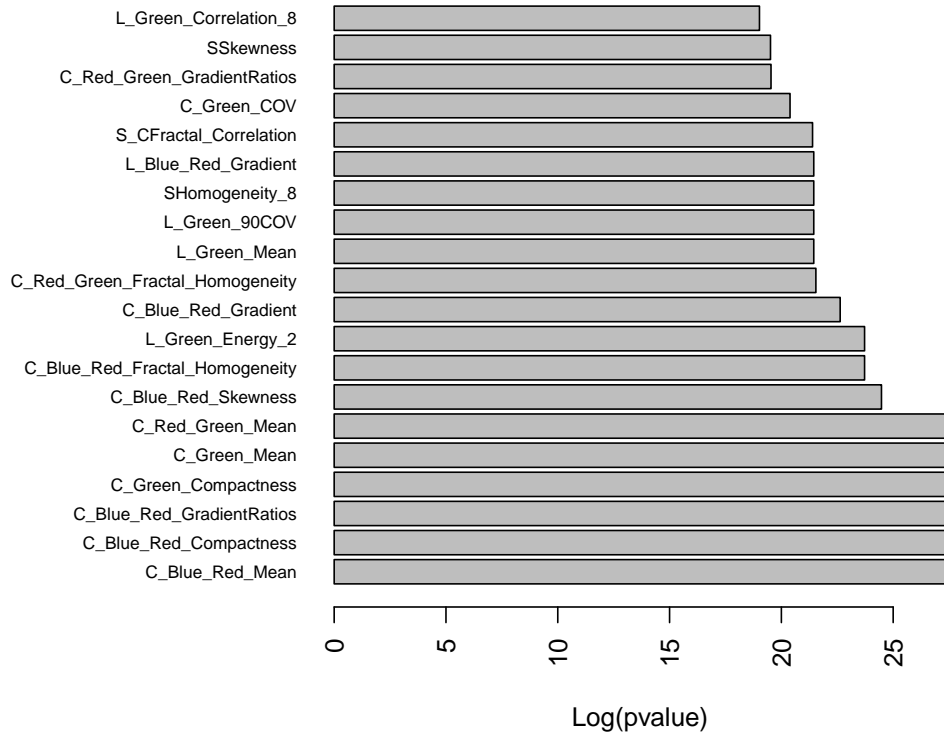
hm <- heatMaps(Outcome = "Class",
  data = MelanomaSeborrheic,
  title = "Heat Map:",Scale = TRUE,
  hCluster = "col",cexRow = 0.75,cexCol = 0.35,srtCol = 45)
```



```
uks <- univariate_KS(MelanomaSeborrheic,"Class")

par(mar=c(5,15,5,5))
barplot(-log(1.0e-12+uks[1:20]),las=2,xlab="Log(pvalue)",cex.names = 0.65,main="Top Features",horiz=TRUE)
```

Top Features



```
print(uks[1:20])
#>          C_Blue_Red_Mean          C_Blue_Red_Compactness
#>          0.000000e+00          0.000000e+00
#> C_Blue_Red_GradientRatios          C_Green_Compactness
#>          0.000000e+00          2.864375e-14
#>          C_Green_Mean          C_Red_Green_Mean
#>          9.166001e-14          1.145750e-13
#> C_Blue_Red_Skewness C_Blue_Red_Fractal_Homogeneity
#>          2.243215e-11          4.889171e-11
#>          L_Green_Energy_2          C_Blue_Red_Gradient
#>          4.889171e-11          1.475812e-10
#> C_Red_Green_Fractal_Homogeneity          L_Green_Mean
#>          4.390983e-10          4.829891e-10
#>          L_Green_90COV          SHomogeneity_8
#>          4.829891e-10          4.829891e-10
#>          L_Blue_Red_Gradient          S_CFractal_Correlation
#>          4.829891e-10          5.100486e-10
#>          C_Green_COV          C_Red_Green_GradientRatios
#>          1.396939e-09          3.274915e-09
#>          SSkewness          L_Green_Correlation_8
#>          3.351653e-09          5.488690e-09

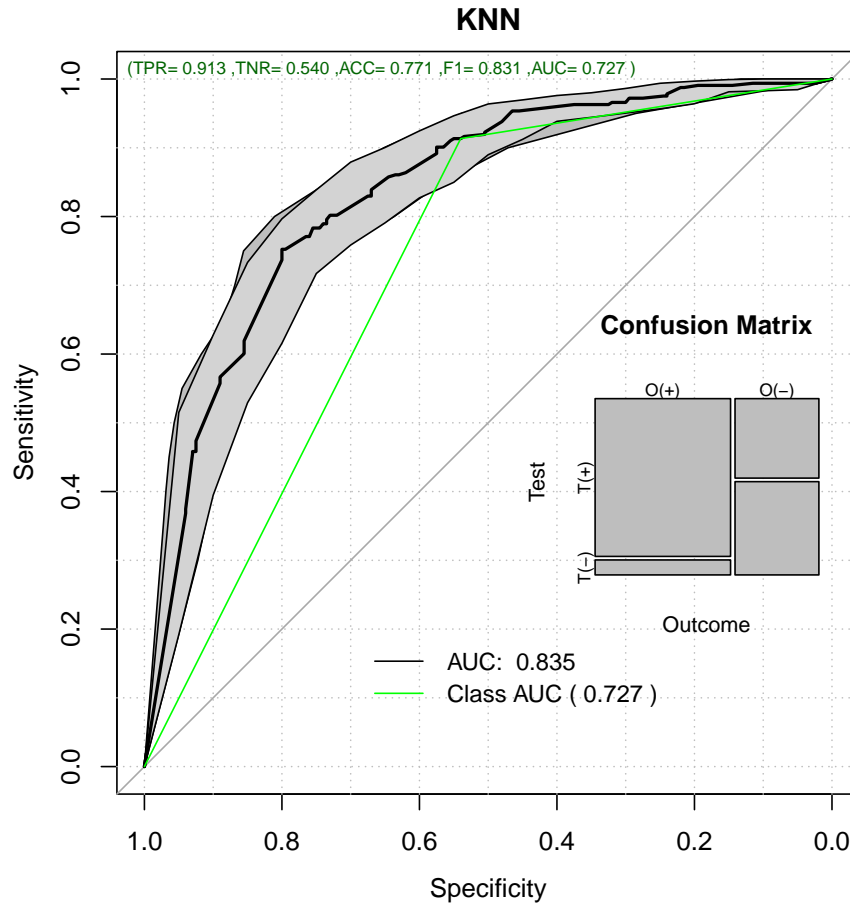
par(op)
```

Learning Melanoma with KNN

```
cvKNN <- randomCV(MelanomaSeborrheic,"Class",
  KNN_method,
  trainFraction = 0.90,
  repetitions = 100,
  classSamplingType = "Pro",
  featureSelectionFunction = univariate_KS,
  featureSelection.control = list(pvalue=0.05,limit= -1),
  kn=5
)
#> .....10  Tested: 352 Avg. Selected: 157.4 Min Tests: 1 Max Tests: 5 Mean Tests: 1.505682 . MAD:
#> .....20  Tested: 472 Avg. Selected: 159.55 Min Tests: 1 Max Tests: 6 Mean Tests: 2.245763 . MAD
#> .....30  Tested: 505 Avg. Selected: 159.4667 Min Tests: 1 Max Tests: 9 Mean Tests: 3.148515 . M
#> .....40  Tested: 512 Avg. Selected: 159.525 Min Tests: 1 Max Tests: 10 Mean Tests: 4.140625 . M
#> .....50  Tested: 520 Avg. Selected: 159.48 Min Tests: 1 Max Tests: 12 Mean Tests: 5.096154 . MA
#> .....60  Tested: 522 Avg. Selected: 158.75 Min Tests: 1 Max Tests: 14 Mean Tests: 6.091954 . MA
#> .....70  Tested: 522 Avg. Selected: 158.8286 Min Tests: 1 Max Tests: 15 Mean Tests: 7.10728 . M
#> .....80  Tested: 522 Avg. Selected: 159.25 Min Tests: 1 Max Tests: 17 Mean Tests: 8.122605 . MA
#> .....90  Tested: 523 Avg. Selected: 159.2 Min Tests: 1 Max Tests: 18 Mean Tests: 9.120459 . MAD
#> .....100 Tested: 523 Avg. Selected: 158.92 Min Tests: 1 Max Tests: 20 Mean Tests: 10.13384 . M
#>
```

Plot performance

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



```
par(op)
```

Learning Melanoma with KNN

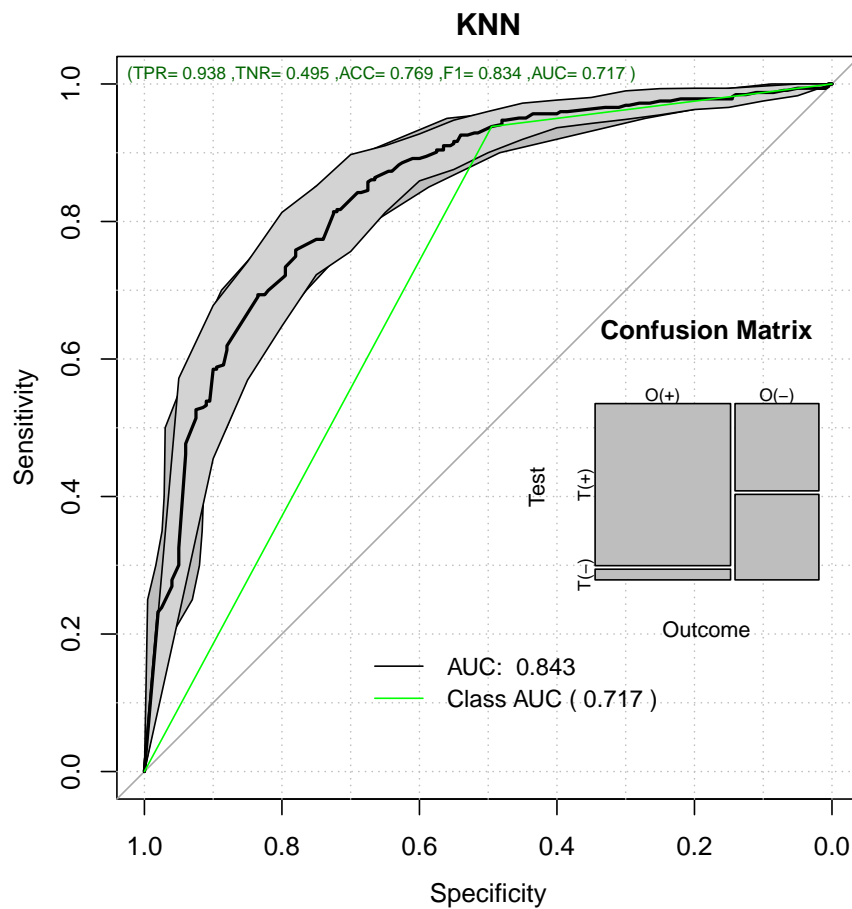
```
cvKNN <- randomCV(MelanomaSeborrheic,"Class",
  KNN_method,
  trainFraction = 0.90,
  repetitions = 100,
  classSamplingType = "Pro",
  featureSelectionFunction = univariate_KS,
  featureSelection.control = list(pvalue=0.05,limit= -1),
  kn=10
)
```

```
#> .....10 Tested: 345 Avg. Selected: 159.6 Min Tests: 1 Max Tests: 5 Mean Tests: 1.536232 . MAD:
#> .....20 Tested: 471 Avg. Selected: 159.35 Min Tests: 1 Max Tests: 7 Mean Tests: 2.250531 . MAD:
#> .....30 Tested: 503 Avg. Selected: 159.1333 Min Tests: 1 Max Tests: 7 Mean Tests: 3.161034 . MAD:
#> .....40 Tested: 518 Avg. Selected: 158.675 Min Tests: 1 Max Tests: 10 Mean Tests: 4.092664 . MAD:
#> .....50 Tested: 520 Avg. Selected: 158.94 Min Tests: 1 Max Tests: 11 Mean Tests: 5.096154 . MAD:
#> .....60 Tested: 522 Avg. Selected: 158.4833 Min Tests: 1 Max Tests: 13 Mean Tests: 6.091954 . MAD:
#> .....70 Tested: 523 Avg. Selected: 158.6286 Min Tests: 1 Max Tests: 15 Mean Tests: 7.09369 . MAD:
#> .....80 Tested: 523 Avg. Selected: 158.825 Min Tests: 1 Max Tests: 16 Mean Tests: 8.107075 . MAD:
```

```
#> .....90 Tested: 523 Avg. Selected: 159.0667 Min Tests: 1 Max Tests: 16 Mean Tests: 9.120459 . M
#> .....100 Tested: 523 Avg. Selected: 159.25 Min Tests: 2 Max Tests: 19 Mean Tests: 10.13384 . M
#>
```

Plot performance

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



```
par(op)
```