

Lab 05 - Classificação usando algoritmos k-NN

Machine Learning com o R - Análise Macro

Thalles Quinaglia Liduares

2022-08-11

Upload pacotes

```
library(caret)
library(class)
```

Upload dados

```
setwd("C:\\Program Files\\R\\Dados")

database<-read.csv(file="wisc_bc_data.csv", stringsAsFactors = F)

attach(database)

data<-database[,-1] # Remover coluna ID
```

Normalização dos dados

```
data[,2:31]<-scale(data[,2:31])
```

Divisão dos dados entre treino e teste

```
set.seed(1608)

part_data<-floor(0.75*nrow(data))

treino_data <-sample(seq_len(nrow(data)), size = part_data)

treino<-data[treino_data, ]

teste<-data[-treino_data,]
```

Modelo

```
prev<-knn(treino[, -1], teste[, -1], cl=treino[,1],k=5)
```

Acurácia

```
matrix_conf<-table(teste[,1],prev)
```

```
confusionMatrix(matrix_conf)
```

```
## Confusion Matrix and Statistics
##
##      prev
##      B  M
## B 95  0
## M  0 48
##
##              Accuracy : 1
##              95% CI : (0.9745, 1)
##      No Information Rate : 0.6643
##      P-Value [Acc > NIR] : < 2.2e-16
##
##              Kappa : 1
##
##  Mcnemar's Test P-Value : NA
##
##      Sensitivity : 1.0000
##      Specificity : 1.0000
##      Pos Pred Value : 1.0000
##      Neg Pred Value : 1.0000
##      Prevalence : 0.6643
##      Detection Rate : 0.6643
##      Detection Prevalence : 0.6643
##      Balanced Accuracy : 1.0000
##
##      'Positive' Class : B
##
```