

Nearest neighbour classification

W. Grossmann

```
#####  
# Nearest neighbour classification demo  
#####  
library(class)  
  
## kNN for the iris data  
data(iris)  
variable.names(iris)  
  
## [1] "Sepal.Length" "Sepal.Width" "Petal.Length" "Petal.Width"  
## [5] "Species"  
  
## Splitting in training and tests set (70% to 30%)  
train<- sample(1:nrow(iris),as.integer(0.7*nrow(iris)))  
trainIris <- iris[train,]  
testIris <- iris[-train,]  
train.def <- trainIris$Species  
  
## A 3-nearest neighbours model  
nn3 <- knn(trainIris[,1:4],testIris[,1:4],train.def, k=3, pro=TRUE)  
  
summary(nn3)  
  
##      setosa versicolor  virginica  
##      15          17          13  
  
attributes(nn3)  
  
## $levels  
## [1] "setosa"      "versicolor" "virginica"  
##  
## $class  
## [1] "factor"  
##  
## $prob  
## [1] 1.0000000 1.0000000 1.0000000 1.0000000 1.0000000 1.0000000 1.0000000  
## [8] 1.0000000 1.0000000 1.0000000 1.0000000 1.0000000 1.0000000 1.0000000  
## [15] 1.0000000 1.0000000 1.0000000 1.0000000 0.6666667 1.0000000 1.0000000  
## [22] 0.7500000 1.0000000 1.0000000 1.0000000 1.0000000 1.0000000 1.0000000  
## [29] 1.0000000 1.0000000 1.0000000 1.0000000 1.0000000 1.0000000 1.0000000  
## [36] 1.0000000 1.0000000 1.0000000 1.0000000 0.7500000 1.0000000 1.0000000  
## [43] 1.0000000 0.6666667 1.0000000
```

```

#Prediction for test case 5
nn3[5]

## [1] setosa
## Levels: setosa versicolor virginica

#Probability of assignment for test case 5
attributes(nn3)$prob[5]

## [1] 1

attributes(nn3)$levels

## [1] "setosa"      "versicolor" "virginica"

## Confusion matrix

table(testIris[, 'Species'], nn3)

##           nn3
##           setosa versicolor virginica
## setosa           15          0          0
## versicolor        0          16          0
## virginica         0           1          13

## classification with 5-nearest neighbours
nn5 <- knn(trainIris[,1:4], testIris[,1:4], train.def, k=5, pro=TRUE)

## The resulting confusion matrix
table(testIris[, 'Species'], nn5)

##           nn5
##           setosa versicolor virginica
## setosa           15          0          0
## versicolor        0          16          0
## virginica         0           0          14

```