Machine Learning I

KNN y Naives Bayes

Dante Conti, Sergi Ramirez, (c) IDEAI

2025-10-11

Table of contents


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

```
set.seed(42)
Default$student = as.numeric(Default$student) - 1
default_idx = sample(nrow(Default), 5000)
default_trn = Default[default_idx, ]
default_tst = Default[-default_idx, ]
```

```
# training data
X_default_trn = default_trn[, -1]
y_default_trn = default_trn$default

# testing data
```

```
X_default_tst = default_tst[, -1]
y_default_tst = default_tst$default
```

```
prediccion <- knn(train = X_default_trn, test  = X_default_tst,
                  cl = y_default_trn, k = 3)
head(prediccion)
```

```
[1] No No No No No No
Levels: No Yes
```

```
calc_class_err = function(actual, predicted) {
  mean(actual != predicted)
}
```

```
calc_class_err(actual    = y_default_tst,
               predicted = knn(train = X_default_trn,
                               test  = X_default_tst,
                               cl    = y_default_trn,
                               k     = 5))
```

```
[1] 0.0312
```

```
calc_class_err(actual    = y_default_tst,
               predicted = knn(train = scale(X_default_trn),
                               test  = scale(X_default_tst),
                               cl    = y_default_trn,
                               k     = 5))
```

```
[1] 0.0284
```

```
set.seed(42)
k_to_try = 1:100
err_k = rep(x = 0, times = length(k_to_try))

for (i in seq_along(k_to_try)) {
  pred = knn(train = scale(X_default_trn),
             test  = scale(X_default_tst),
             cl    = y_default_trn,
             k     = k_to_try[i],
```