Exercise SL.1. k-Nearest Neighbors algorithm (*knn*) and overfitting

In this exercise you will work with the *iris* dataset. The *iris* dataset is available in R. You should type *data(iris)* in order to get it.

This database consists of 150 cases and 5 features:

- Sepal length
- Sepal width
- Petal length
- Petal width
- Species (qualitative variable with 3 classes): setosa, virginica and versicolor.

You must perform *knn* on this dataset. You should proceed as follows:

- Create an object called *iris.training* which contains a sample of 110 cases from the *iris* dataset.
- Create an object called *iris.test* which contains the cases not included in *iris.training*.
- Perform *knn* taking as training sample the cases in *iris.training* and as test sample the cases in *iris.test*. Decide on a initial value for k.
- Compute the error rate for the test sample and repeat the previous step until you find an appropriate value for k. Which is the error rate you get?