

**OPENCOURSEWARE
ADVANCED PROGRAMMING
STATISTICS FOR DATA SCIENCE**

Ricardo Aler



**ADVANCED PROGRAMMING
MASTER IN STATISTICS FOR DATA SCIENCE.
RCPP ASSIGNMENT: PROGRAMMING 1-NEAREST NEIGHBOUR IN C++**

1.5 POINTS

Introduction

The aim of this assignment is to program in C++ a simple but useful machine learning method: KNN or k-nearest neighbor. In order to save time, I have already programmed KNN with K=1 in R, which you will find in Aula Global. That is the code that you have to translate into C++.

There are two ways of programming and compiling functions in Rcpp: *cppFunction* and *sourceCpp*. In order to practice what you have been taught in class, you will use both of them.

What to do

1. Translate the KNN R code available in Aula Global into Rcpp, using *cppFunction*
2. Compile the C++ code you wrote using also *sourceCpp*
3. Use the library *microbenchmark* in order to determine whether the code you wrote is faster or slower than the standard R *knn* function, that belongs to the CRAN library *class* (*library(class)*)
4. Write a short report (1 page) about the code you wrote and the results of the *microbenchmark*.
5. Hand in your code and report

