

## Statistical Learning, Exercise 6 : Regression with k-NN

### Question 1

Create a function to compute the L2-norm (Euclidean distance) :

$$\|x\|_2 := \sqrt{x_1^2 + \cdots + x_n^2}$$

### Question 2

Create a function to compute the L1-norm (Manhattan distance) :

$$\|x\|_1 := \sum_{i=1}^n |x_i|$$

### Question 3

Load the Computer data set from the file ComputerData.txt and the Cars data set from the file Cars2Data.txt. You have now to predict the response variables (PRP and mpg) with the k-NN strategy. Apply the k-NN method to the Computer and Cars2 data set with  $K = 2$ . Is your model useful?

Be sure to compute the distance using the similar measurements across all the predictors. If you don't, the variable with the biggest variance will dominate the result. You have to normalize these values first (subtract the mean and divide the samples by the standard deviation).

### Question 4

You're now free to select the value of the parameter  $k$ . Build several models with different values of  $k$  and choose a model. Is your model useful? Why did you choose this value/model?