Aprendizagem 2023

Lab 4: Linear Regression and *k*NN

Prof. Rui Henriques

Practical exercises

I. Lazy learning

1. Consider the following data:

	input		output	
	y_1	y_2	y_3	y_4
$\mathbf{x}_1$	1	1	A	1.4
$\mathbf{x}_2$	2	1	B	0.5
$\mathbf{x}_3$	2	3	B	2
$\mathbf{x}_4$	3	3	B	2.2
$\mathbf{x}_5$	2	2	A	0.7
$\mathbf{x}_6$	1	2	A	1.2

Assuming a *k*-nearest neighbor with *k*=3 applied within a leave-one-out schema:

a) Let y_3 be the output variable (*categoric*). Considering an Euclidean (*l*2) distance, provide the classification estimates for x_1.

$\lVert\mathbf{x}_i - \mathbf{x}_j\rVert_2$	x_1	x_2	x_3	x_4	x_5	x_6
x_1	-	1	$\sqrt{5}$	$\sqrt{8}$	$\sqrt{2}$	1

$$\hat{z}_1 = mode(B, A, A) = A$$

b) Let y_4 be the output variable (*numeric*). Considering cosine similarity, provide the mean regression estimate for x_1.

cos	x_1	x_2	x_3	x_4	x_5	x_6
x_1	-	0.95	0.98	1	1	0.95

$$\hat{z}_2 = mean(2, 2.2, 0.7) = 1.6(3)$$

c) Consider a weighted-distance *k*-nearest neighbor with Manhattan (l_1) distance, identify the:

i. weighted mode estimate of x_1 for y_3 outcome

l_1	x_1	x_2	x_3	x_4	x_5	x_6
x_1	-	1	3	4	2	1

$$\hat{z}_1 = weighted\_mode\left(\frac{1}{1}B, \left(\frac{1}{2} + \frac{1}{1}\right)A\right) = weighted\_mode(1 \times B, 1.5 \times A) = A$$

ii. weighted mean estimate of x_1 for y_4 outcome

$$\hat{z}_1 = \frac{\frac{1}{1}0.5 + \frac{1}{2}0.7 + \frac{1}{1}1.2}{\frac{1}{1} + \frac{1}{2} + \frac{1}{1}} = 0.82$$

II. Linear regression

1. Considering the following data to learn a model $z = w_1 y_1 + w_2 y_2 + \varepsilon$, where $\varepsilon \sim N(0, 0.1)$

Compare:

	y_1	y_2	output
x_1	3	-1	2
x_2	4	2	1
x_3	2	2	1

a) $\boldsymbol{w} = [w_1\ w_2]^T$ using the maximum likelihood approach

Maximum likelihood can be given by (proof on the slides): $\mathbf{w} = (X^T X)^{-1} X^T Z$

Solve exercise similarly as previous ones.

b) $\boldsymbol{w}$ using the Bayesian approach, assuming $p(\boldsymbol{w}) = N\left(\boldsymbol{w} \mid \boldsymbol{u} = [0\ 0], \boldsymbol{\sigma} = \begin{bmatrix} 0.2 & 0 \\ 0 & 0.2 \end{bmatrix}\right)$

Maximum posterior is given by (proof on the slides): $\mathbf{w} = (X^T X + \lambda I)^{-1} X^T Z$

$\lambda = \frac{{\sigma_{posterior}}^2}{{\sigma_{prior}}^2} = \frac{0.1^2}{0.2^2}$. Solve exercise similarly as $1.f$).