Aprendizagem 2023

Lab 9: Clustering

Prof. Rui Henriques

Practical exercises

1. Consider the following training data without and the cluster centres

	y_1	y_2
$\mathbf{x}_1$	0	0
$\mathbf{x}_2$	1	0
$\mathbf{x}_3$	0	2
$\mathbf{x}_4$	2	2

$$\boldsymbol{u}_1 = \begin{pmatrix} 1/2 \\ 0 \end{pmatrix}, \qquad \boldsymbol{u}_2 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$$

a) Compute the silhouette of observation $\mathbf{x}_1$, cluster c_1 and overall solution

b) Knowing $\mathbf{x}_1$, $\mathbf{x}_2$ and $\mathbf{x}_4$ to be annotated as positive and $\mathbf{x}_3$ as negative (ground truth). Compute the purity of *k*-means against the given ground truth.