

Algorithm CART-SPLITREGION ($R, \mathcal{D}$)

Inputs R a region, $\mathcal{D}$ the data points contained in R , $|\mathcal{D}| = N$
for $j = 1 : n$ *coordinate axes*

1. sort $\mathcal{D}$ by coordinate j , denote the sorted data by $(x^{[1]:[N]}, y^{[1]:[N]})$
2. for $i = 1 : N - 1$
 - (a) calculate split points $\xi_{ji} = \frac{x^{[i]} + x^{[i+1]}}{2}$
 - (b) consider split into $D_{ji}^{\pm} = \{(x^l, y^l), x_j^l \lessgtr \xi_{ij}\}$
 - (c) compute $impurity_{ij}^{\pm}$

find the most pure split $i^*, j^* = \underset{i,j,\pm}{\operatorname{argmin}} impurity_{ij}^{\pm}$

Output $j^*, \xi_{i^*j^*}$ and $\mathcal{D}^{\pm} = \mathcal{D} \cap D_{j^*i^*}^{\pm}$

A split is good if at least one region it creates is pure.