

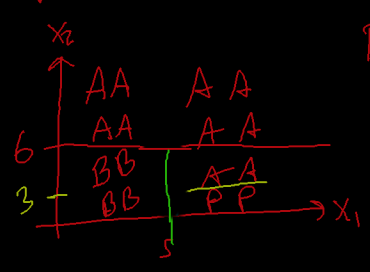
CART (1984)
 ↓
 classification tree
 regression tree

Categorical (nominal)
 $Y = \{C_1, C_2, \dots, C_K\}$ K levels

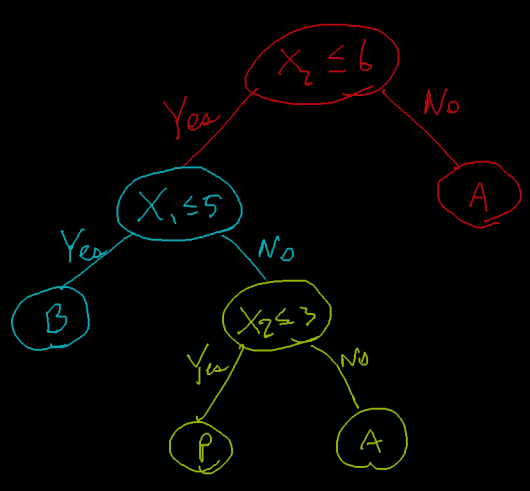
if $K=2 \Rightarrow$ binary classification

$C_1=0, C_2=1$

$p=2$



II



$$Gini_r = \frac{3}{4} \cdot \frac{1}{4} + \frac{1}{4} \cdot \frac{3}{4} = 6/16$$

$$Gini_L = \frac{1}{2} \cdot \frac{1}{2} + \frac{1}{2} \cdot \frac{1}{2} = 0 = 8/16$$

$$= 10/16$$

Default $N_0 = 1$.

Classification Algorithm

Same as the regression tree algorithm except

(I) The objective function which evaluates each split is not weighted $SS E_L, SS E_r$. Instead we use the weighted avg. of $Gini_L, Gini_r$

$$Gini_w = \frac{n_L Gini_L + n_r Gini_r}{n_L + n_r}$$

left

$$Gini_L = \sum_{k=1}^K \hat{p}_k (1 - \hat{p}_k) \quad \text{where} \quad \hat{p}_k = \frac{\sum_{i=1}^{n_L} \mathbb{1}_{y_i = C_k}}{n_L}$$

observations in node

This is called "Gini Impurity" thanks to Kurt!

(II) Leaf Assignment $\hat{y} = \text{Mode}[y's]$

