

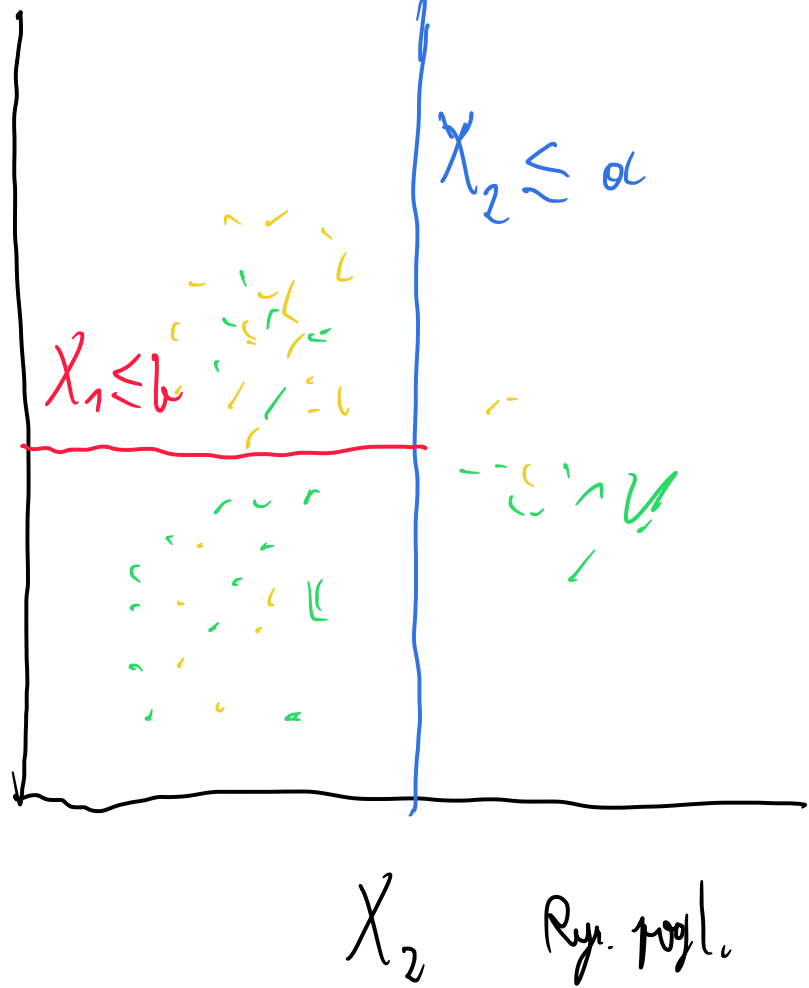
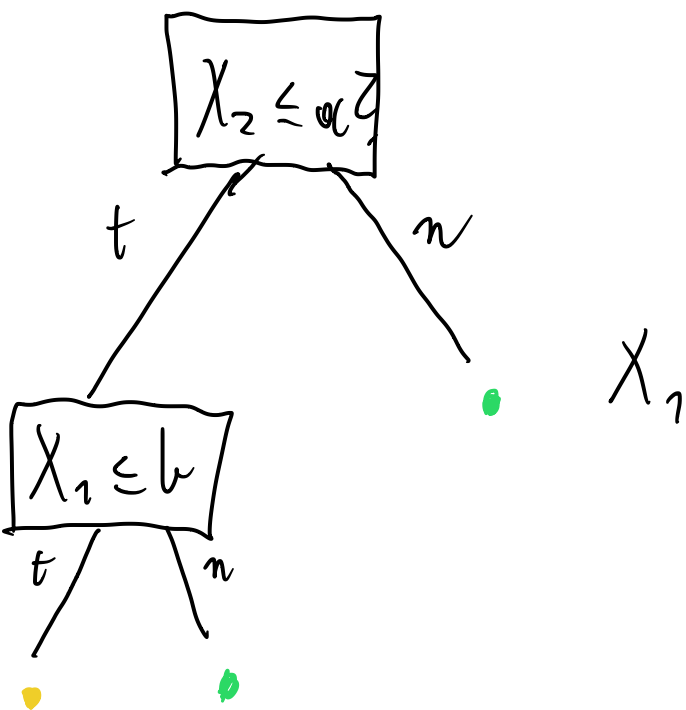
31/ Drzewa decyzyjne: konstrukcja (dla liniowej)

Dane:  $X_1, X_2, \dots, X_n$  - cechy

$y_1, y_2, \dots, y_n$  - klasy

$(x^1, y^1), (x^2, y^2), \dots, (x^n, y^n)$  - zbiór uczący  
 $x_i$

1. Wybieramy pierwszą cechę według ustalonego porz.
2. Dzielimy zbiór uczący tak, aby zminimalizować funkcję straty. (Np. % popr. klasyfikacji, zakładając, że przewidywana wartość albo wezta to ta o najw. %). W ten sposób otrzymujemy węzeł wewnętrzny.
3. Powtarzamy albo podzbiórów i bez cechy  $x_i$
4. Kończymy gdy zostanie spełniony określony warunek. Np. rozmiar drzewa albo niskie tempo poprawy.



Najpierw dzielimy

$X_1 \times X_2$  po  $X_n$ . Prawy  
węzeł ma słab. wartości błęd.

• -  $y_1$   
+ -  $y_2$

Słab., dzielimy  $X_2 \leq \alpha$  po  $X_1$

Nowo otrzymane węzły mają ok. błęd.

Ponieważ  $q$  nie jest uprzed. liniowo, nie ma ograniczonego kryterium podziału, po którym można podzielić wartości  $q$ .  
 Wszystkie możliwości obrotu przedziału jest bardzo dużo:  
 $n > 2$ :

$$\left\{ \begin{smallmatrix} n \\ 2 \end{smallmatrix} \right\} = 2 \cdot \left\{ \begin{smallmatrix} n-1 \\ 2 \end{smallmatrix} \right\} + \left\{ \begin{smallmatrix} n-1 \\ 1 \end{smallmatrix} \right\} \quad \leftarrow \text{1. Stw. 2 roz.}$$

$$= 2 \left( 2 \left\{ \begin{smallmatrix} n-2 \\ 2 \end{smallmatrix} \right\} + 1 \right) + 1 = 4 \left\{ \begin{smallmatrix} n-2 \\ 2 \end{smallmatrix} \right\} + 2 + 1 = \dots =$$

$$2^{n-2} + \sum_{i=1}^{n-2} 2^{n-2-i} = \sum_{i=0}^{n-2} 2^{n-2-i} 1^i =$$

$$\left\{ \begin{smallmatrix} n \\ k \end{smallmatrix} \right\} = k \left\{ \begin{smallmatrix} n-1 \\ k \end{smallmatrix} \right\} + \left\{ \begin{smallmatrix} n-1 \\ k-1 \end{smallmatrix} \right\}$$

$$\left\{ \begin{smallmatrix} n \\ 1 \end{smallmatrix} \right\} = 1, \left\{ \begin{smallmatrix} n \\ n \end{smallmatrix} \right\} = 1$$

$$= 2^{n-1} - 1$$

Możliwe rozw.:

- Przekształcenie każdej wartości na przedział binarny (ostrożnie! wtedy trzeba rozważyć sl. l. cedh)
- Rozważenie lotowej próbki partyjnej