

## Zadanie 1

Dana jest próbka  $U = (u_1, \dots, u_N)$  mająca atrybuty  $a_1, \dots, a_k, a_i : U \rightarrow V_i$ , oraz decyzję  $d : U \rightarrow V_d$ . Błąd empiryczny klasyfikatora  $f$  na próbce  $U$  wynosi:

$$\mathcal{E}_U(f) = \frac{1}{N} \sum_{i=1}^N L(f(a_1(u_i), \dots, a_k(u_i)), d(u_i)),$$

gdzie

$$L(x, y) = \begin{cases} 0 & \text{gdy } x = y \\ 1 & \text{gdy } x \neq y \end{cases}$$

Podaj wzór na wartość błędu empirycznego dla drzewa decyzyjnego składającego się jedynie z liścia, które minimalizuje błąd empiryczny. Podaj wzór na wartość błędu empirycznego dla drzewa decyzyjnego mającego jeden węzeł wewnętrzny, które minimalizuje błąd empiryczny. Wywnioskuj na tej podstawie kryterium wyboru selektora w zachłanym algorytmie budowy drzewa decyzyjnego.

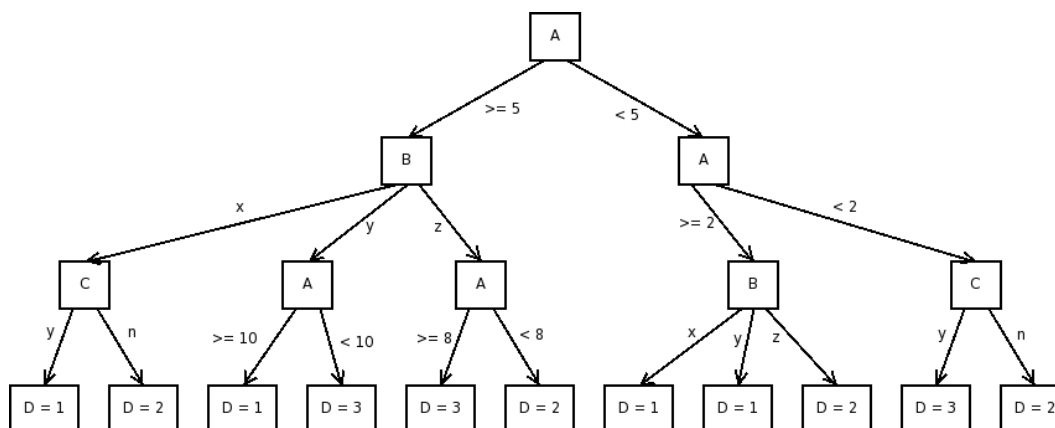
## Zadanie 2

Korzystając z kryterium wyboru selektora uzyskanego w poprzednim zadaniu, utwórz drzewo decyzyjne na podstawie następującej tabeli:

Example	Attributes										Goal
	Alt	Bar	Fri	Hun	Pat	Price	Rain	Res	Type	Est	Will Wait?
$X_1$	Yes	No	No	Yes	Some	\$\$\$	No	Yes	French	0-10	Yes
$X_2$	Yes	No	No	Yes	Full	\$	No	No	Thai	30-60	No
$X_3$	No	Yes	No	No	Some	\$	No	No	Burger	0-10	Yes
$X_4$	Yes	No	Yes	Yes	Full	\$	No	No	Thai	10-30	Yes
$X_5$	Yes	No	Yes	No	Full	\$\$\$	No	Yes	French	> 60	No
$X_6$	No	Yes	No	Yes	Some	\$\$	Yes	Yes	Italian	0-10	Yes
$X_7$	No	Yes	No	No	None	\$	Yes	No	Burger	0-10	No
$X_8$	No	No	No	Yes	Some	\$\$	Yes	Yes	Thai	0-10	Yes
$X_9$	No	Yes	Yes	No	Full	\$	Yes	No	Burger	> 60	No
$X_{10}$	Yes	Yes	Yes	Yes	Full	\$\$\$	No	Yes	Italian	10-30	No
$X_{11}$	No	No	No	No	None	\$	No	No	Thai	0-10	No
$X_{12}$	Yes	Yes	Yes	Yes	Full	\$	No	No	Burger	30-60	Yes

### Zadanie 3

Dane jest drzewo decyzyjne



oraz próbka walidująca

A	B	C	D
3	y	y	3
10	x	n	1
11	y	y	3
2	x	y	1
14	y	y	1
4	x	n	1
7	z	n	1
9	z	n	3
8	y	y	1
10	y	n	1
4	z	n	2
12	y	y	1

Przytnij drzewo zgodnie z algorytmem podanym na wykładzie. Zaznacz kolejność w jakiej algorytm odwiedza kolejne węzły. Dla każdego odwiedzonego węzła podaj jego postać po przycięciu, jego accuracy przed i po przycięciu oraz zaznacz, czy zostanie przycięty.