

1. Drzewa decyzyjne przedstawiane są w postaci:
  - a) grafów,
  - b) grafów skierowanych,
  - c) grafów, w których z każdego wierzchołka wychodzą maksymalnie dwie krawędzie,
  - d) grafów nie zawierających cykli.

2. Które z poniższych wyrażeń było wyrażeniem, które wykorzystano przy liczeniu entropii:

$$\frac{2}{7} \log_2 \frac{2}{7} + \log_2 \frac{5}{5}$$

$$\frac{2}{7} \log_2 \frac{2}{7} + \frac{3}{4} \log_2 \frac{3}{4}$$

$$\frac{0}{7} \log_2 \frac{0}{7} + \log_2 \frac{7}{7}$$

$$\frac{2}{7} \log_2 \frac{2}{7} + \frac{5}{7} \log_2 \frac{5}{7}$$

Dlaczego? .....

.....

.....

.....

3. W węzłach drzewa decyzyjnego dla zbioru uczącego przedstawionego na wykładzie 5 mogą znaleźć się warunki:
  - a)  $A1 == \text{high}$ ,
  - b)  $A2 == \text{low}$  and  $A6 == \text{med}$ ,
  - c)  $A5 < > \text{bign}$ ,
  - d) żadna z powyższych odpowiedzi.

**Ćwiczenie 1**

Biorąc pod uwagę dane dotyczące liczby wystąpień poszczególnych wartości atrybutu  $A9$  w zbiorze uczącym  $T$  wyznacz wartość  $E(T, A9)$  obliczaną podczas tworzenia drzewa decyzyjnego:

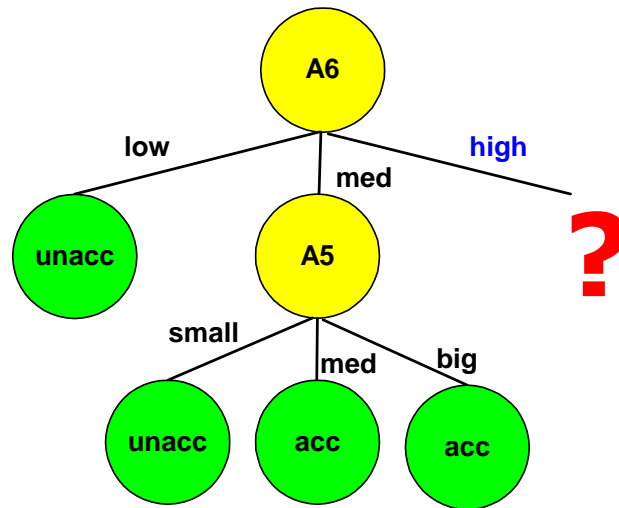
	klasa:	dobry	zły	niepewny
<b>A9:</b>	mała	0	1	2
	duża	2	4	3
	średnia	3	3	3
	wielka	3	5	1

.....

.....

## Ćwiczenie 2

Zbuduj do końca drzewo decyzyjne. Które z atrybutów ( $A_1, A_2, A_3, A_4, A_5$ ) zostaną wybrane do tworzenia kolejnych elementów drzewa?



A1	A2	A3	A4	A5	A6	klasa
low	vhigh	5more	4	big	high	acc
high	med	5more	4	big	high	acc
low	vhigh	3	4	med	high	acc
med	vhigh	5more	more	med	high	acc
med	vhigh	2	4	small	high	acc
high	low	3	2	small	high	unacc
high	low	3	2	med	high	unacc
high	med	4	2	med	high	unacc

## Ćwiczenie 3 (na samotny wieczór)

Określ jak będzie wyglądało drzewo decyzyjne dla zbioru uczącego przedstawionego poniżej.

A1	A2	A3	A4	A5	A6	klasa
med	low	2	2	med	high	unacc
low	vhigh	2	2	small	med	unacc
vhigh	vhigh	3	4	small	low	unacc
high	low	2	more	med	med	unacc
low	vhigh	5more	2	small	high	unacc
high	med	5more	more	med	med	acc
med	med	3	4	med	med	acc
high	med	4	more	small	high	acc
low	low	4	more	small	high	good
low	med	4	more	big	med	good
low	med	3	4	small	high	good
low	low	2	4	med	high	good
low	med	3	more	small	high	good