

## Drzewa decyzyjne (II) + algorytm k-NN

---

1. Binarne drzewa decyzyjne różnią się od drzew decyzyjnych następującymi własnościami
  - a) składają się z węzłów mających dokładnie dwie krawędzie,
  - b) umożliwiają klasyfikację danych należących do jednej z dwóch klas,
  - c) mają w swojej budowie wyłącznie liście probabilistyczne,
  - d) zwykle są drzewami o większej wysokości.
  
2. Liść probabilistyczny to:
  - a) element decyzyjnego drzewa binarnego,
  - b) element drzewa decyzyjnego,
  - c) element, którego wykorzystanie jest możliwe tylko w przypadku dwuklasowego klasyfikatora,
  - d) żadna z powyższych odpowiedzi.
  
3. Przycinanie drzewa oznacza:
  - a) wybranie  $n$  węzłów drzewa,
  - b) usunięcie tych gałęzi drzewa, które reprezentują najmniejszą liczbę przykładów uczących,
  - c) usunięcie takich gałęzi, że błąd klasyfikacji nie wzrasta powyżej założonego poziomu,
  - d) żadna z powyższych odpowiedzi.
  
4. Algorytm  $k$ -NN to:
  - a) algorytm  $k$  nienazwanych przykładów,
  - b) algorytm, który może być dwukrotnie gorszy od klasyfikatora Bayesa,
  - c) inna wersja algorytmu  $k$  średnich
  - d) żadna z powyższych odpowiedzi.
  
5. Algorytm  $k$ -NN wiąże się z:
  - a) obliczaniem średniej ważonej punktów reprezentujących dane uczące,
  - b) obliczaniem odległości między punktami reprezentującymi dane treningowe,
  - c) obliczaniem mediany i macierzy kowariancji,
  - d) żadna z powyższych odpowiedzi.

## Ćwiczenie 1

Poniżej przedstawiono dwa wiersze danych, które należy zaklasyfikować wykorzystując poniższe drzewo decyzyjne. Jeśli klasyfikacja nie jest możliwa uzasadnij: dlaczego? Na jakiej podstawie możliwa była klasyfikacja i do jakich klas zostaną owe dwa wiersze zaklasyfikowane?

A1	A2	A3	A4	A5	A6	klasa
high	high	5more	?	?	med	?
low	med	?	?	?	med	?

