

# DRZEWA DECYZYJNE

## LAB 3 – Uczenie bez nadzoru

---

### Zadanie 1. (1 pkt.)

Zbiór danych: <http://archive.ics.uci.edu/ml/machine-learning-databases/00396/>

Dane dotyczące transakcji różnych produktów na przestrzeni 51-tygodni zawierają zarówno wartości oryginalne jak i znormalizowane.

### Zadanie 2. (3 pkt.)

Wykonaj analizę skupień danych dla następujących algorytmów:

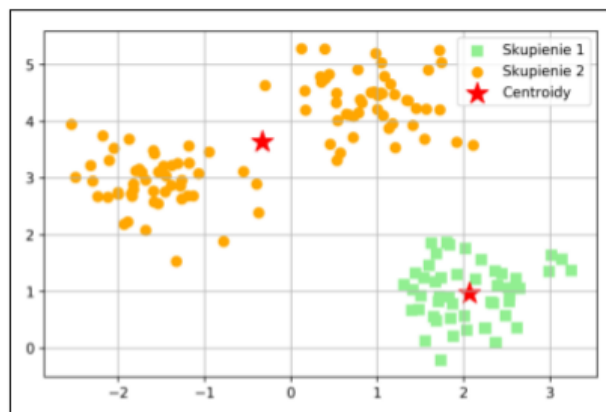
- K-MEANS
- K-MEANS++
- \*K-MEDIOD

W wyniku analizy skupień dodaj do każdej próbki kategorię odpowiadającą przypisanemu klastrowi (zakładamy kategorie całkowite: 1,2,3,...).

W zaimplementowanych algorytmach jako argument użytkownik może wprowadzić liczbę klastrów (domyślnie 5).

Przedstaw graficznie (na wspólnym wykresie) wyniki analizy skupień dla w/w metod:

- Każdy klastrow powinien być zaprezentowany unikatowym kolorem oraz stylem punktu (np. czerwone kółko, niebieski kwadrat, itp).
- Dla każdego klastra wyznacz centrum klastra.



Wynikiem działania algorytmu powinien być zbiór danych z etykietami klas przypisanymi na podstawie klasteryzacji.

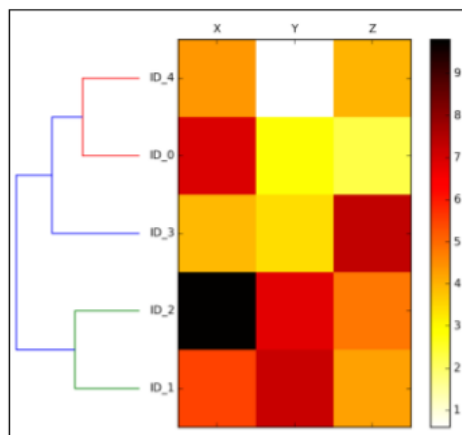
### Zadanie 3. (3 pkt.)

Wykonaj klasteryzację dla tego samego zbioru danych na podstawie algorytmów:

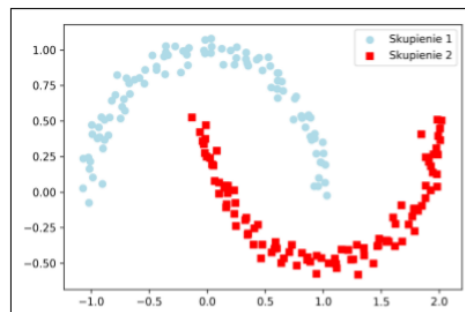
- AgglomerativeClustering z dowolnym typem klasteryzacji, typem wiązania oraz metryką odległości
- DBSCAN z domyślnymi wartościami argumentów

Zaprezentuj graficznie (na jednym wykresie) wyniki klasteryzacji:

- Klasteryzacja aglomeracyjna – dendrogram + mapa cieplna



- DBSCAN – wykres z oznaczeniem klastrow (kolor + symbol)



W wyniku analizy skupień dodaj do każdej próbki kategorię odpowiadającą przypisanemu klastrowi (zakładamy kategorie całkowite: 1,2,3,...).

**Wynikiem działania algorytmu powinien być zbiór danych z etykietami klas przypisanymi na podstawie klasteryzacji.**

#### Zadanie 4. (3 pkt.)

Korzystając z miary silhouette, a dokładnie średniej miary silhouette dla wszystkich klastrow oblicz i zaprezentuj graficznie optymalne liczby klastrow dla każdego z analizowanych algorytmow w zakresie od 2 do 15 klastrow.

Przyklad zobrazowania:

