

#### Ćwiczenie 4. (5pt) Klasyfikacja – algorytm k-najbliższych sąsiadów

##### Zadania:

Przeprowadzić walidację krzyżową modelu dostarczonego algorytmem k-najbliższych sąsiadów dla danych z pliku "seeds.csv":

1. Walidacja krzyżowa z dwoma zbiorami: podziel zbiór danych na zbiór trenujący i zbiór testowy o rozmiarze podanym w pliku "przydziały.docs". Porównaj wyniki klasyfikatora z prawdziwymi wartościami zmiennej celu. Pokaż macierz pomyłek i wskaż źle sklasyfikowane rekordy. Zapisz poziom błędu.
2. Walidacja krzyżowa na  $k$  podzbiorach: ustaw liczbę podzbiorów  $k$  tak, jak wskazano w pliku "przydziały.docs". Zapisz średnią wartość poziomu błędu, jak również wartości poziomów błędów dla wszystkich podzbiorów.

Liczba sąsiadów  $k$  dla algorytmu k-najbliższych sąsiadów powinna zostać wybrana z przedziału od 1 do 6.