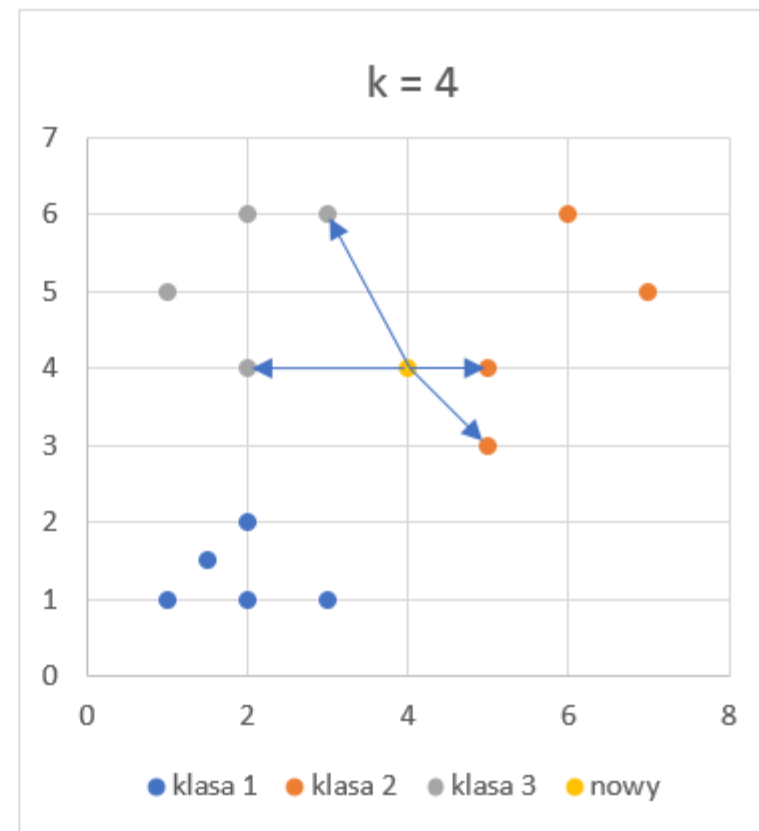
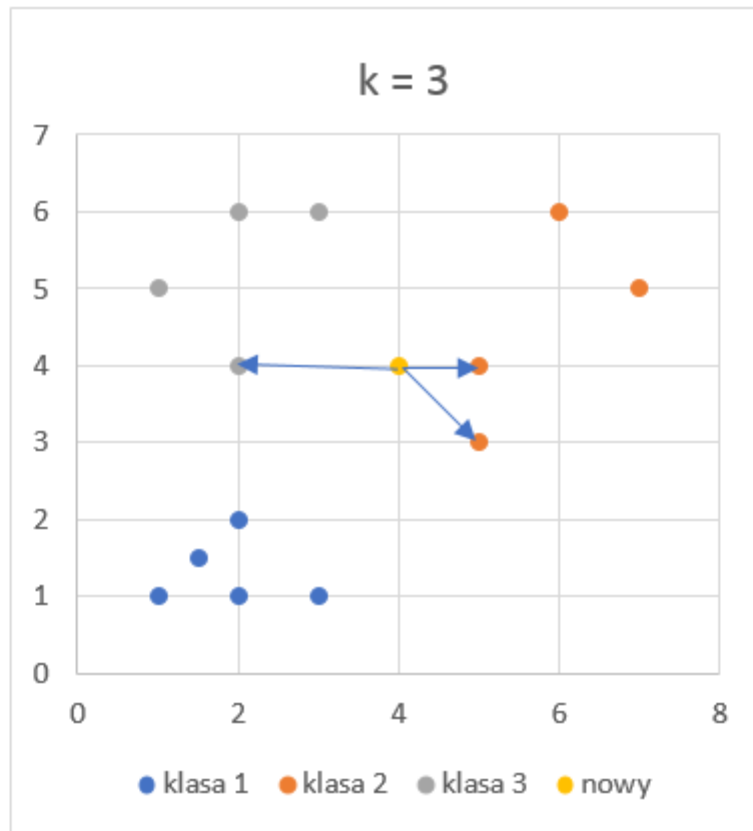
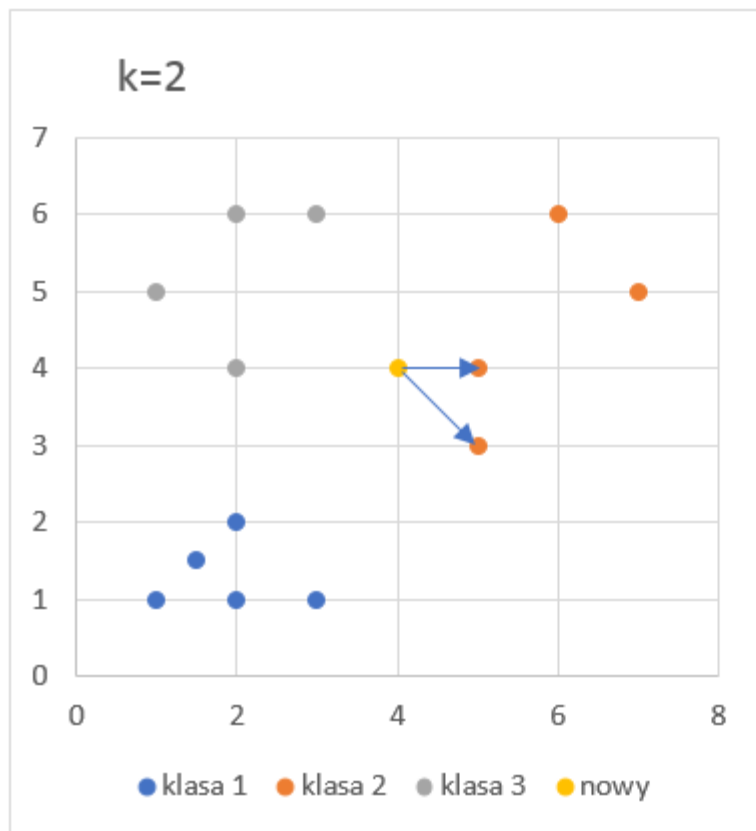


Metoda klasyfikacji k-nn

# Algorytm k-najbliższych sąsiadów (k-NN - k – nearest neighbors) głosowanie proste

1. Wyznacz wartość  $k$ , oznaczającą liczbę sąsiadów, z którymi nowy przykład będzie porównywany w celu wyznaczenia klasy
2. Porównać nowy przykład do  $k$  najbliższych sąsiadów, czyli tych które mają najmniejszą odległość w przyjętej metryce
3. Wyznacz klasę przez określenie największej liczby wzorców wyznaczających daną klasę

# Przykład

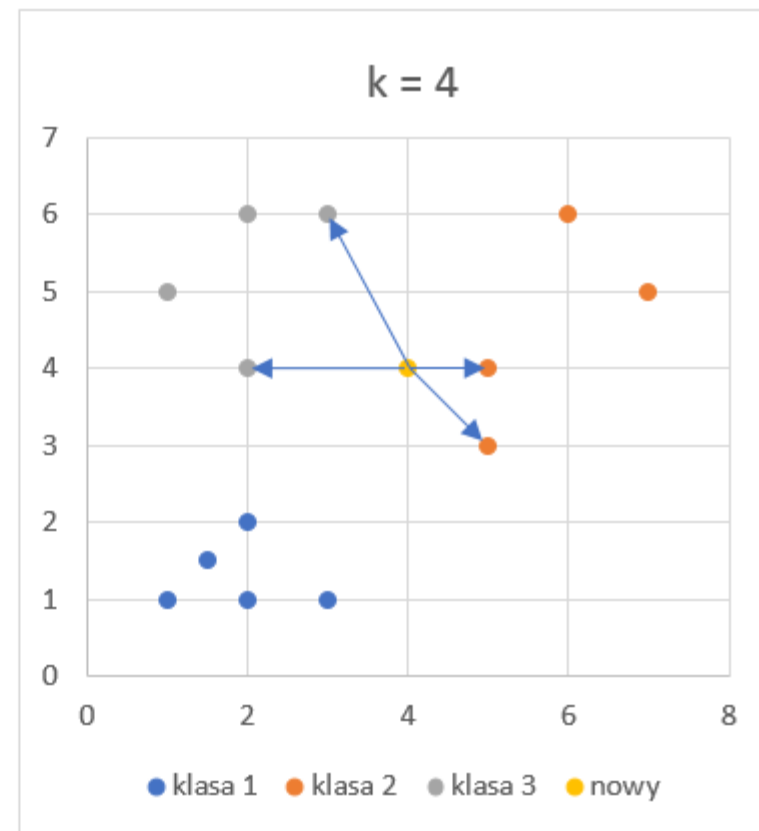
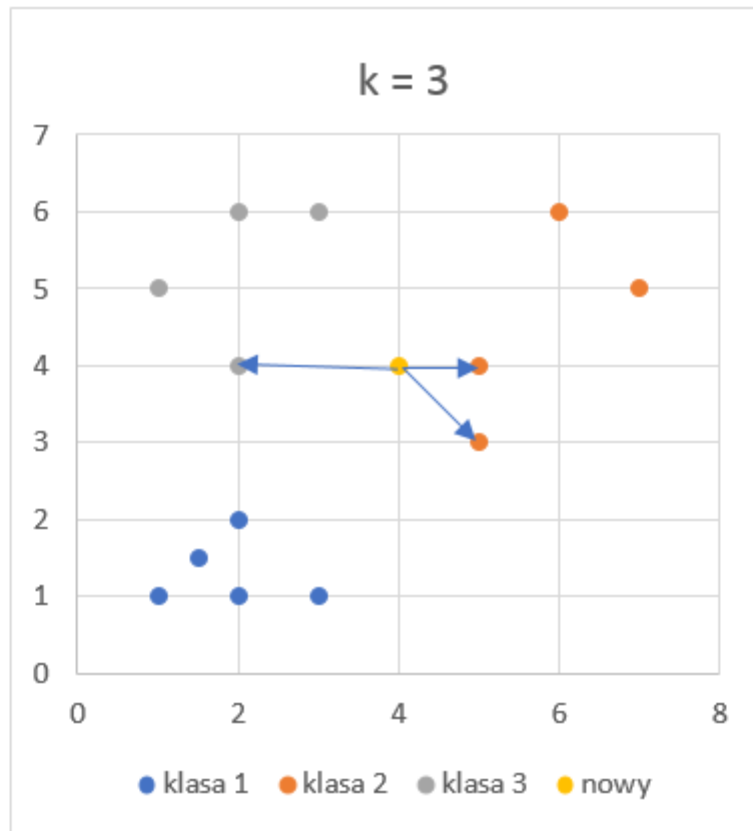
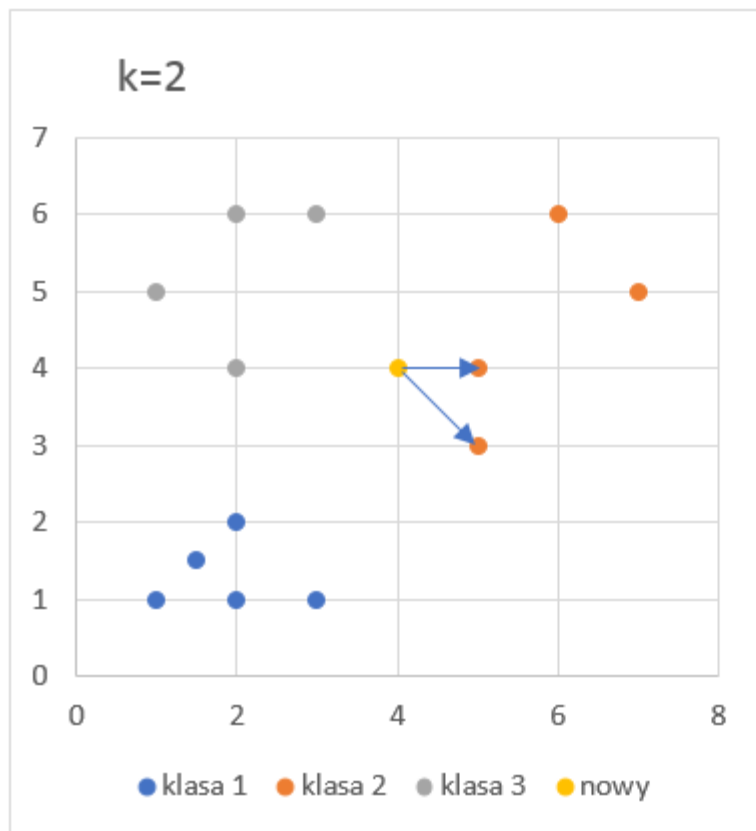


# Algorytm k-najbliższych sąsiadów (k-NN - k – nearest neighbors) głosowanie ważone

1. Wyznaczyć wartość k, oznaczającą liczbę sąsiadów, z którymi nowy przykład n będzie porównywany w celu wyznaczenia klasy
2. Porównać nowy przykład do k najbliższych sąsiadów  $x_i$ , czyli tych które mają najmniejszą odległość w przyjętej metryce
3. Wyznaczyć klasę przez określenie funkcji decyzyjnej bazującej na wyznaczonej odległości d od poszczególnych wzorców, np.:

$$p_i(x_i, n) = \frac{1}{1 + d(x_i, n)^2}$$

# Przykład



# Algorytm k-najbliższych sąsiadów (k-NN - k – nearest neighbors) głosowanie ważone

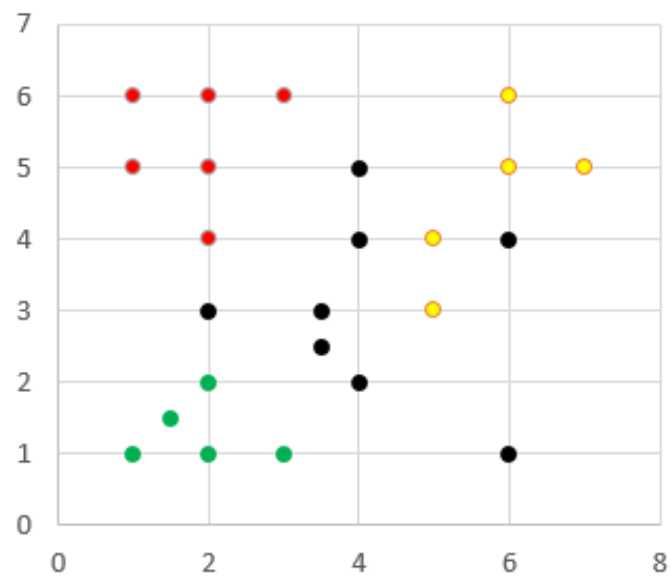
1. Wyznaczyć wartość k, oznaczającą liczbę sąsiadów, z którymi nowy przykład n będzie porównywany w celu wyznaczenia klasy
2. Porównać nowy przykład do k najbliższych sąsiadów  $x_i$ , czyli tych które mają najmniejszą odległość w przyjętej metryce
3. Wyznaczyć klasę przez określenie funkcji decyzyjnej bazującej na wyznaczonej odległości d od poszczególnych wzorców, np.:

$$p_i(x_i, n) = \frac{1}{1 + |d(x_i, n)|}$$

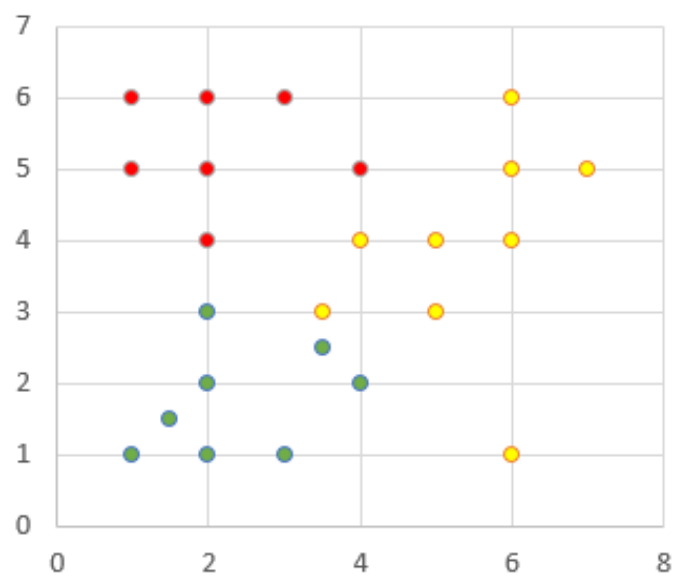
# Wybór wartości $k$

- Wybór małej wartości  $k$  :
  - Może uwidocznić się wpływ wzorców oddalonych od środka danej klasy (może być to szum lub błąd w danych)
  - Dla  $k = 1$  algorytm zwróci klasę najbliższego punktu (wzorca)
- Wybór bardzo dużej wartości  $k$ :
  - Utrata informacji o lokalnym wpływie poszczególnych wzorców, zbyt duże uogólnienie danych zawartych w przykładach

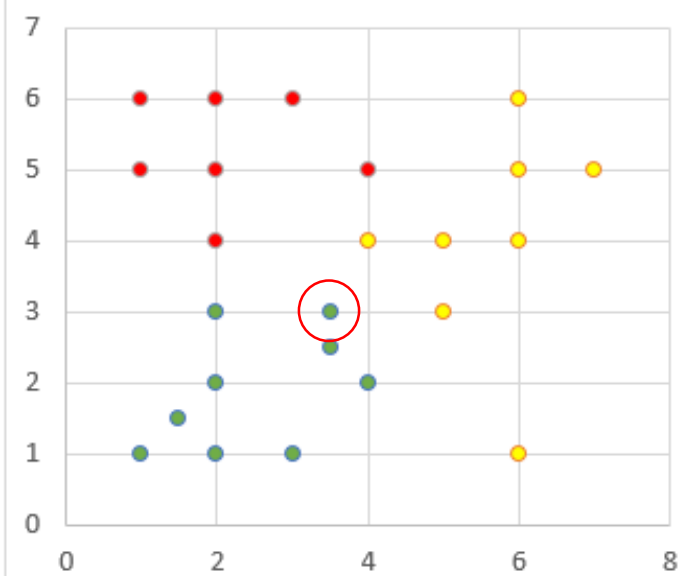
knn



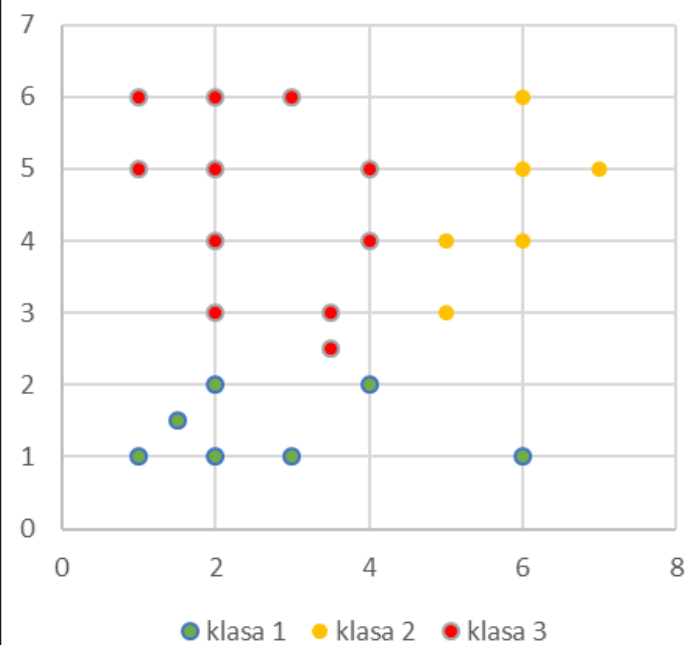
knn k = 3 w2



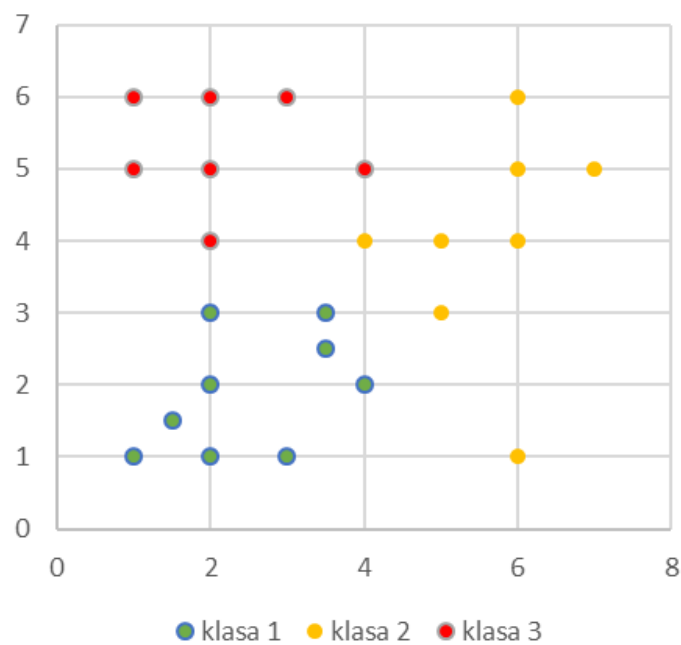
knn k = 7 w2



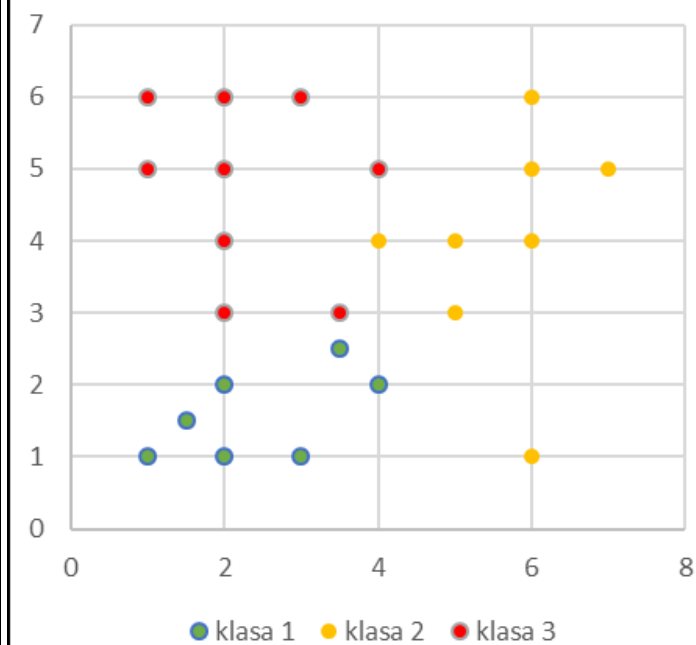
knn k = 14 w1



knn k = 14 w2



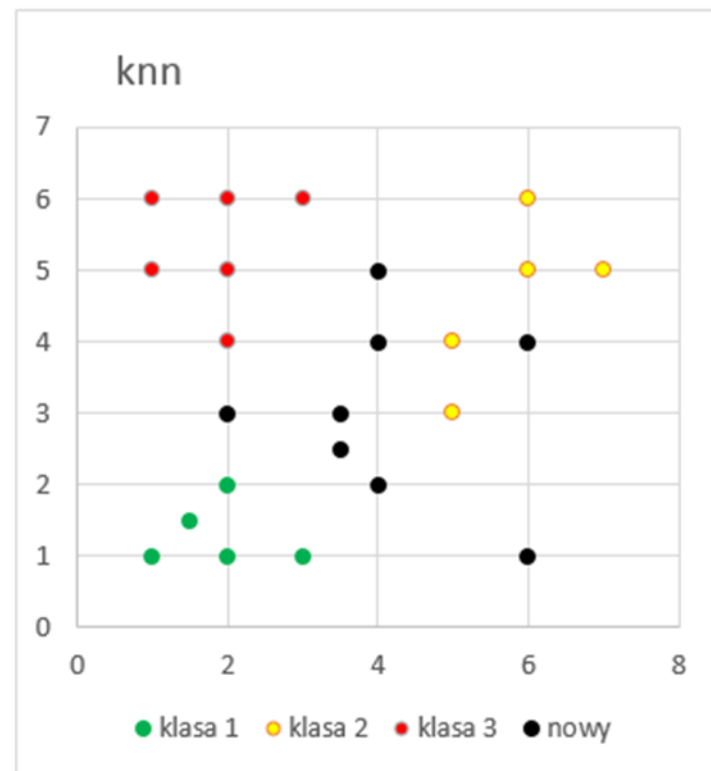
knn k = 14 w3





# Przykład

x1	x2	Klasa
1	1	1
1	5	3
1	6	3
1,5	1,5	1
2	1	1
2	2	1
2	4	3
2	5	3
2	6	3
3	1	1
3	6	3
5	3	2
5	4	2
6	5	2
6	6	2
7	5	2



# Pseudokod programu

```
For j = 1 To liczba_nowych_przykładów
    For i = 1 To liczba_wzorców
        todl (i, 1) = odleglosc(i, j)
        todl(i, 2) = i
        todl(i, 3) = tdane(i, 3)
    Next i
    sortuj(todl) // posortuj dane wg odleglosci
    klasa = zlicz_klasy(k, wariant)
    znowe(j, 3) = klasa
Next j
```

## Funkcja **zlicz klasy**

```
maxk = 0, klasa_max = 0
For i = 1 To k
    klasa = todl(i, 3)
    If wariant = 1 Then
        tklas(klasa) = tklas(klasa) + 1
    Elseif wariant = 2 Then
        tklas(klasa) += 1 / (1+ todl(i, 1)2)
    Elseif wariant = 3 Then
        tklas(klasa) += 1 / (1+ |todl(i, 1)|)
    End If

    If maxk < tklas(klasa) Then
        klasa_max = klasa
        maxk = tklas(klasa)
    End If
Next i
return klasa_max
```