

# Metody rozpoznawania obrazów i podstawy uczenia maszynowego

Zadanie: Condensed Nearest Neighbours

Wykonanie: Kamil Kurp

Zbiór wejściowy - bitmapa:

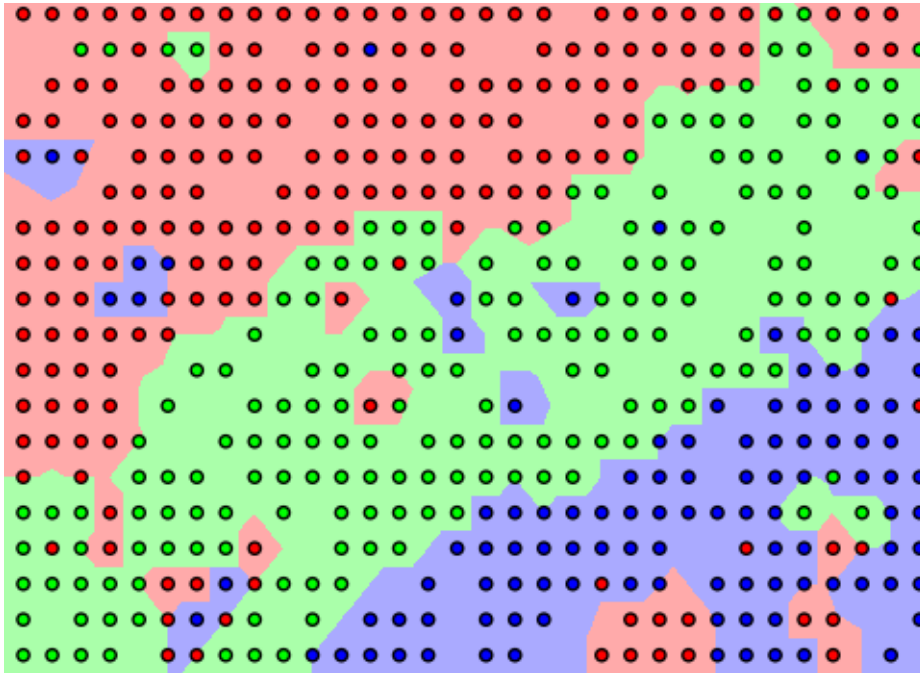


Klasa A – kolor czerwony

Klasa B – kolor zielony

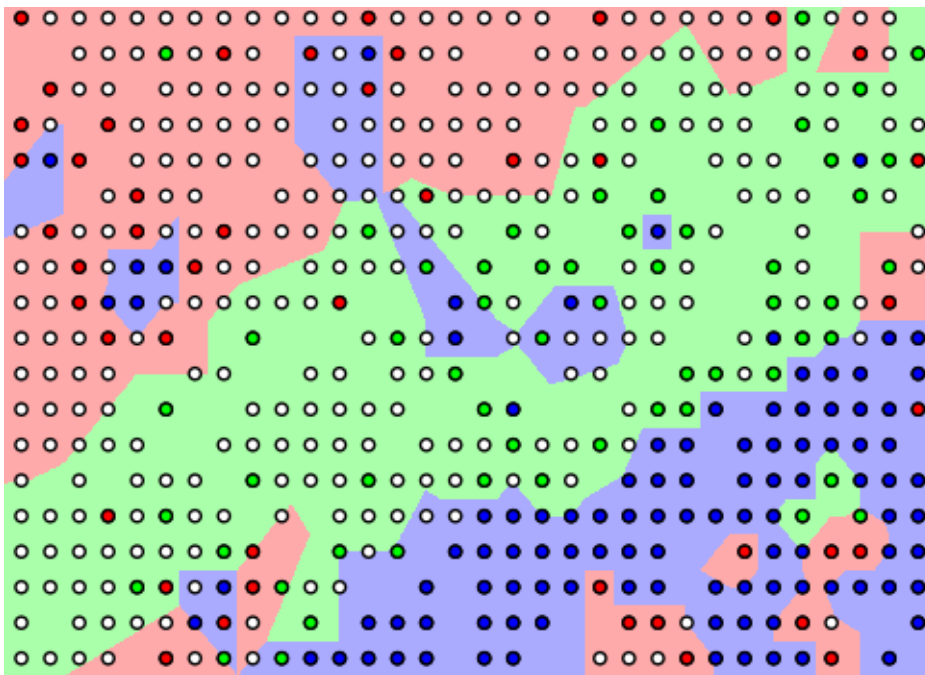
Klasa C – kolor niebieski

1. Zwykły k-NN z  $k=1$  i metryką Euklidesa



Błąd: 26.08%

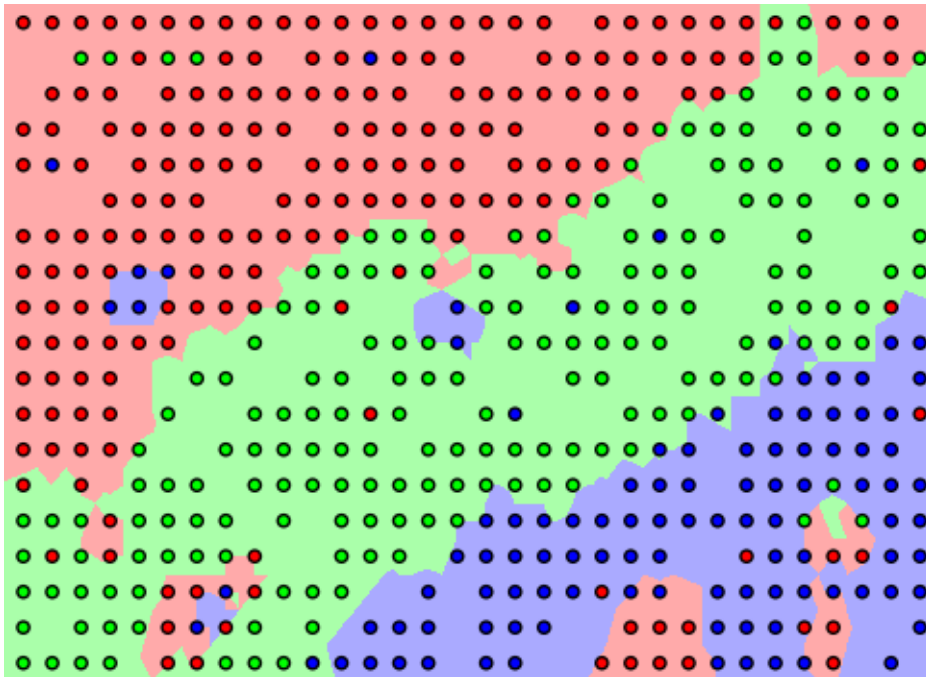
2. CNN z  $k=1$  i metryką Euklidesa (usunięte elementy na biało)



Błąd: 39.13%

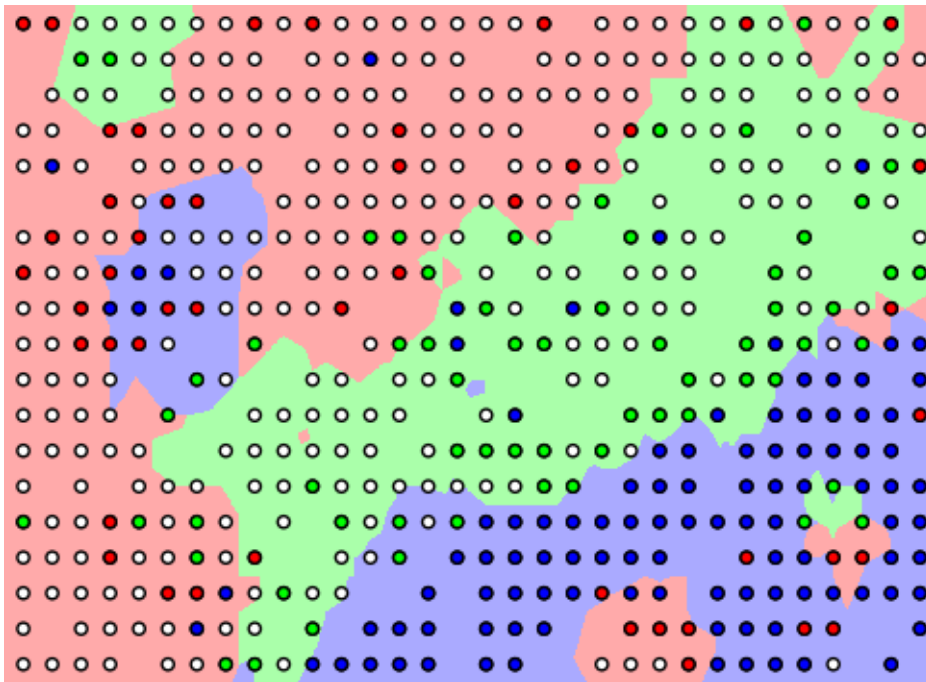
Punktów pozostałych po redukcji: 42.88%

3. Zwykły k-NN z  $k=3$  i metryką Euklidesa



Błąd: 19.87%

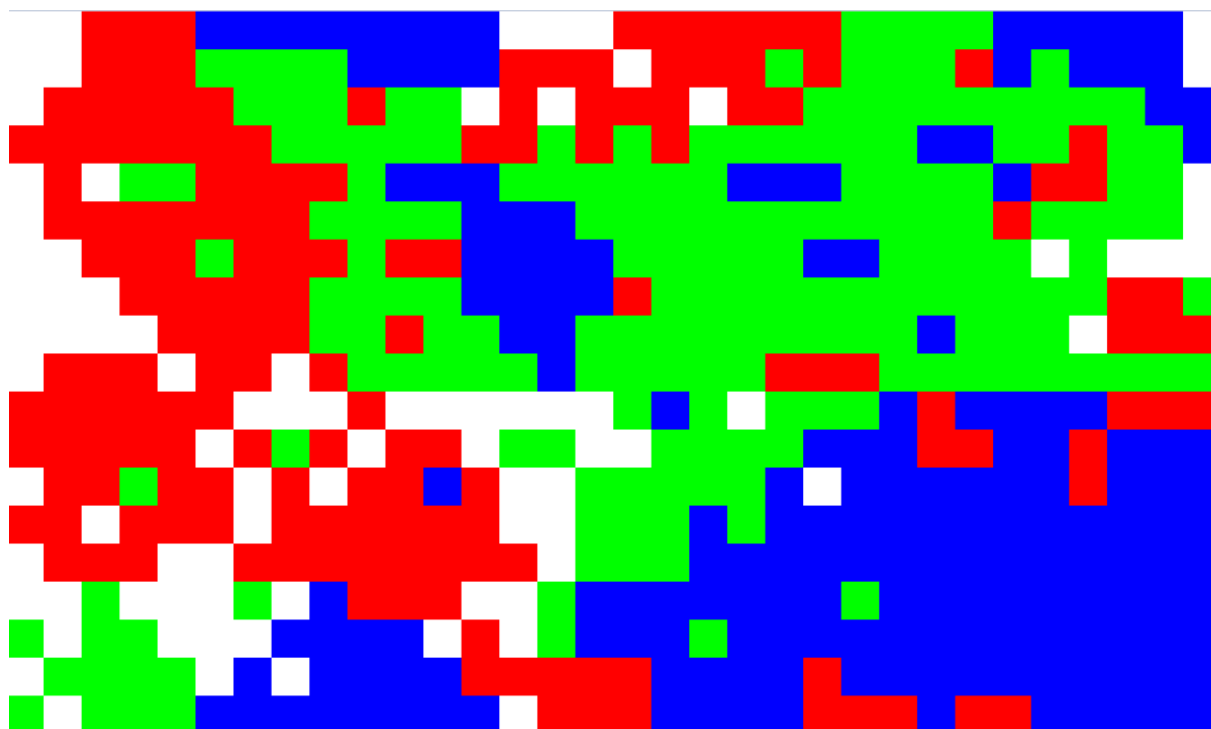
4. CNN z  $k=3$  i metryką Euklidesa (usunięte elementy na białą)



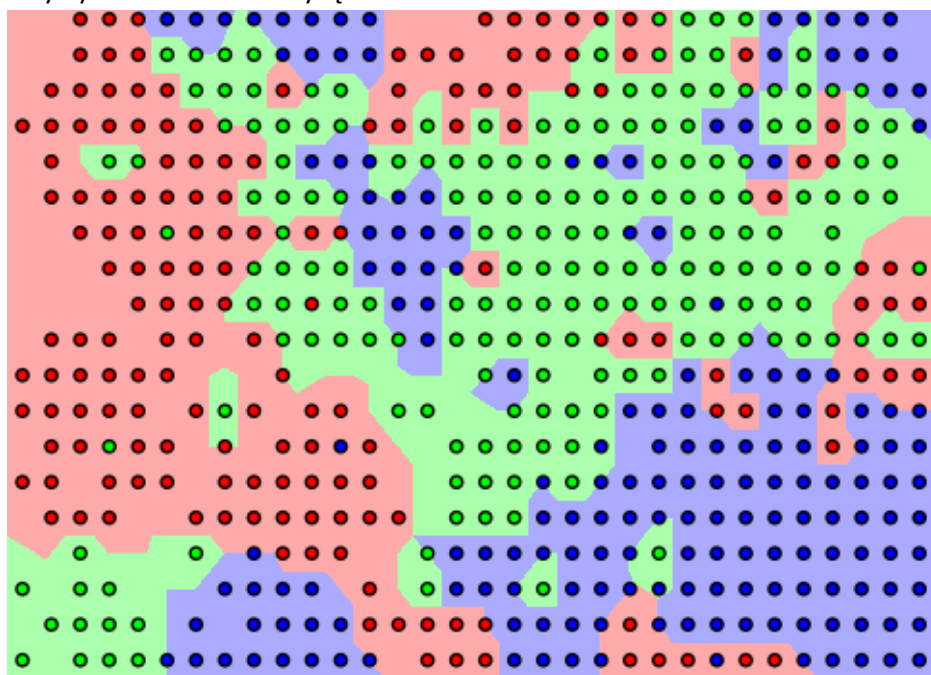
Błąd: 45.07%

Punktów pozostałych po redukcji: 44.32%

Drugi zbiór wejściowy:

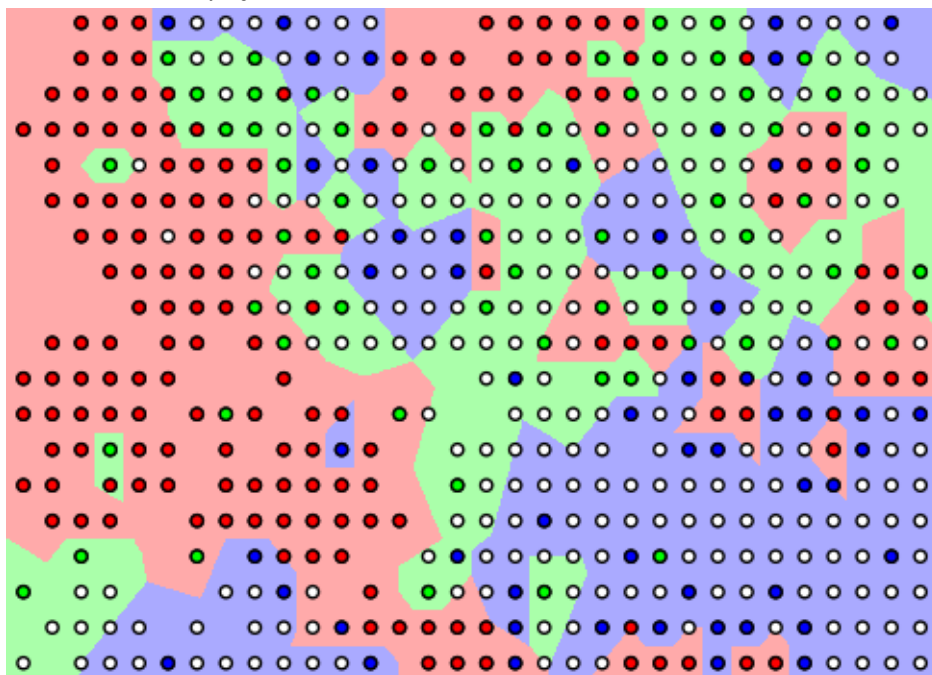


1. Zwykły k-NN z  $k=1$  i metryką Euklidesa



Błąd: 21.96%

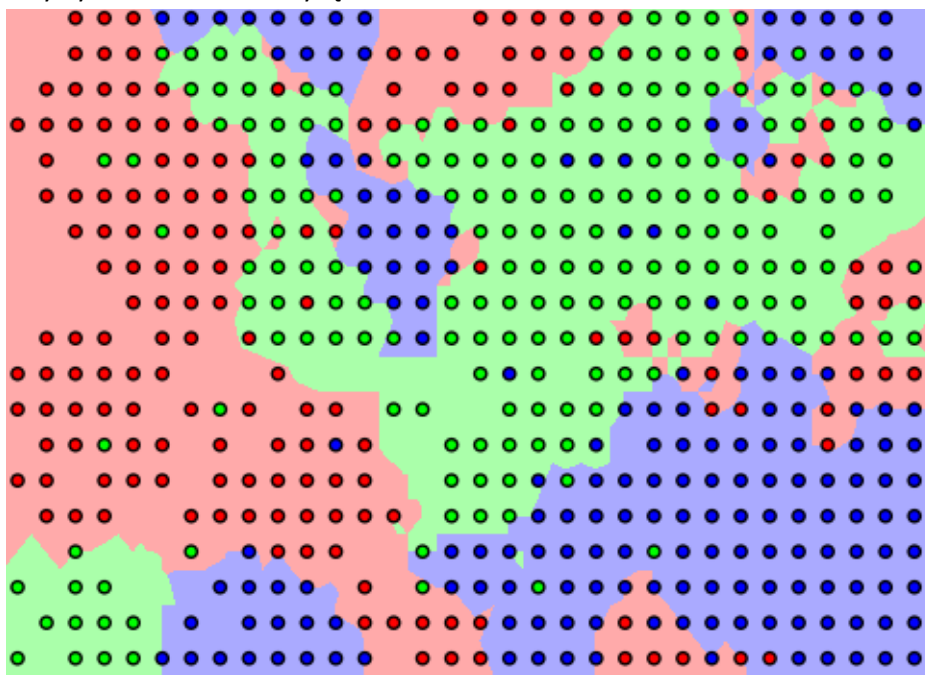
2. CNN z  $k=1$  i metryką Euklidesa



Błąd: 41.93%

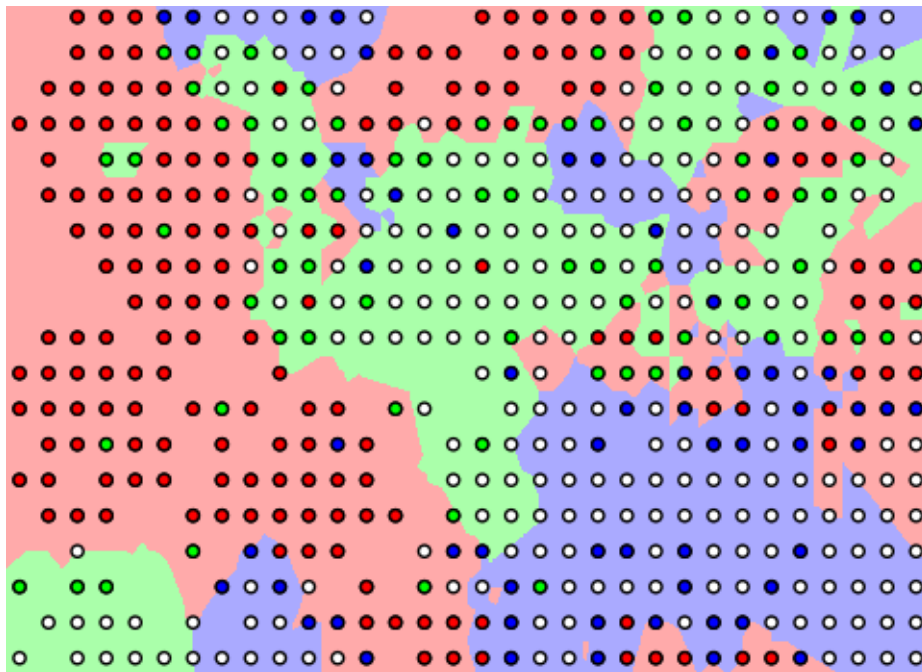
Punktów pozostałych po redukcji: 53.43%

3. Zwykły k-NN z  $k=3$  i metryką Euklidesa



Błąd: 21.96%

#### 4. CNN z $k=3$ i metryką Euklidesa



Błąd: 42.42%

Punktów pozostałych po redukcji: 56.87%

#### Podsumowanie:

Usunięcie elementów za pomocą CNN powoduje, że zdarzają się przypadki, gdzie jedna samotna obserwacja, która nie jest usunięta, może spowodować, że zostanie zaklasyfikowany do jej klasy duży obszar wokół niej, jeśli wiele punktów dookoła było usuniętych. Granica decyzyjna dla CNN z  $k = 1$  różni się dosyć znacząco od granicy dla CNN i  $k = 3$ . Dla CNN błąd rośnie prawie dwukrotnie, ale wzamian redukujemy ilość elementów prawie o połowę. Podając algorytmowi dane z trzema różnymi klasami, z jednej z tych klas elementy nie wydają się być wcale usuwane. Można zauważyć, że algorytm CNN faworyzuje elementy na granicach obszarów, a elementy wewnątrz obszarów są umieszczane w równych odległościach od siebie.