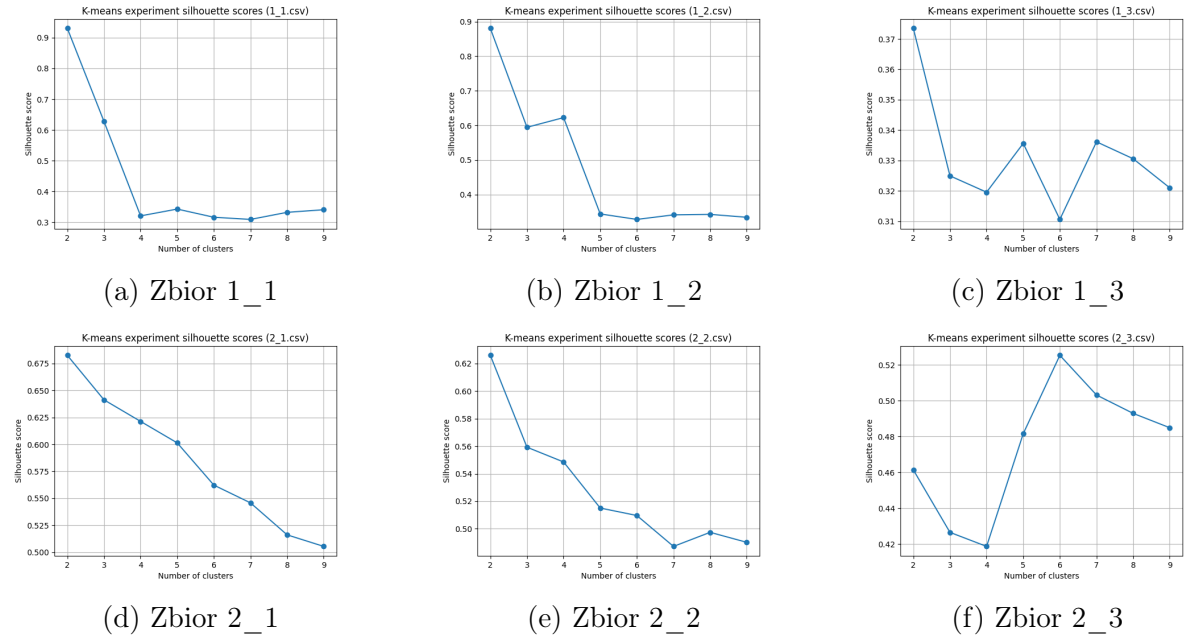
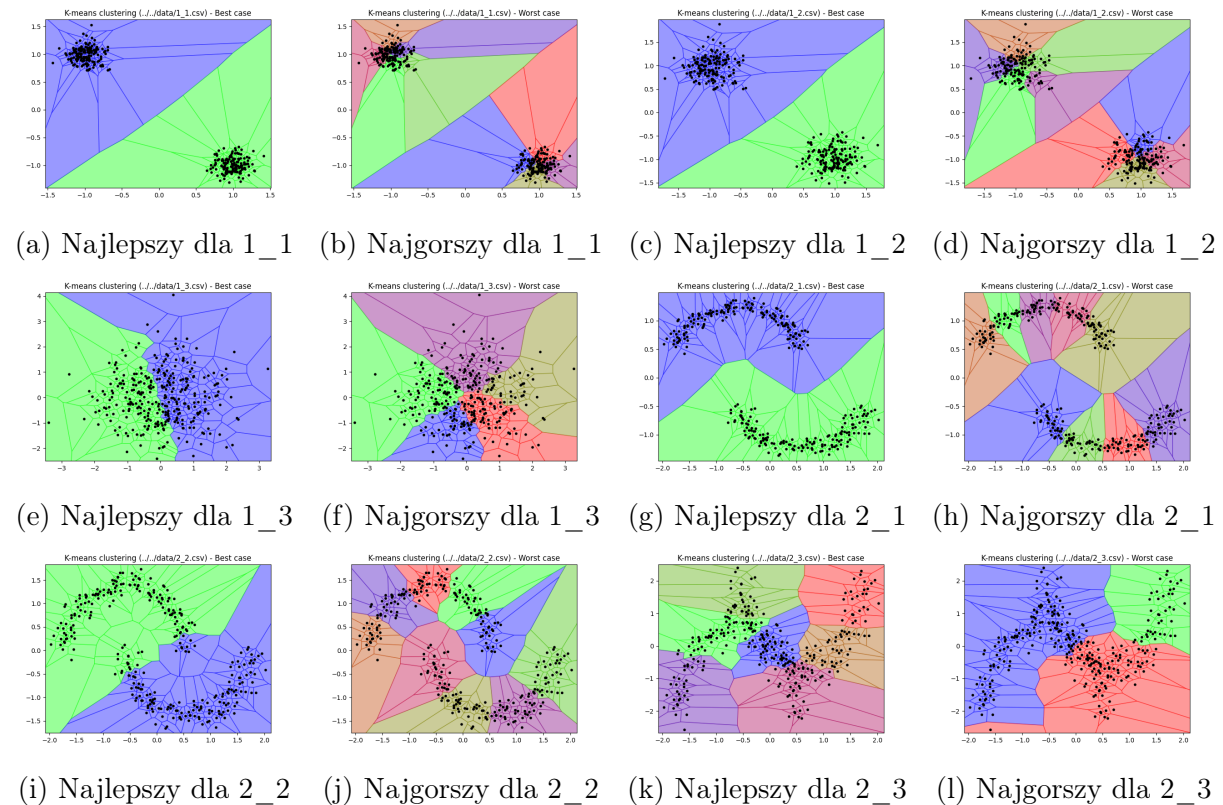


Obliczenia Inteligentne		Projekt 1	
Grupa: Grupa 1	Dzień i czas: Czwartek, 10:00		Rok akademicki: 2023/24
Imię i nazwisko: JAKUB PAWLAK		Imię i nazwisko: MAGDALENA PAKUŁA	

## 1 Eksperyment 1: Metoda K-Means

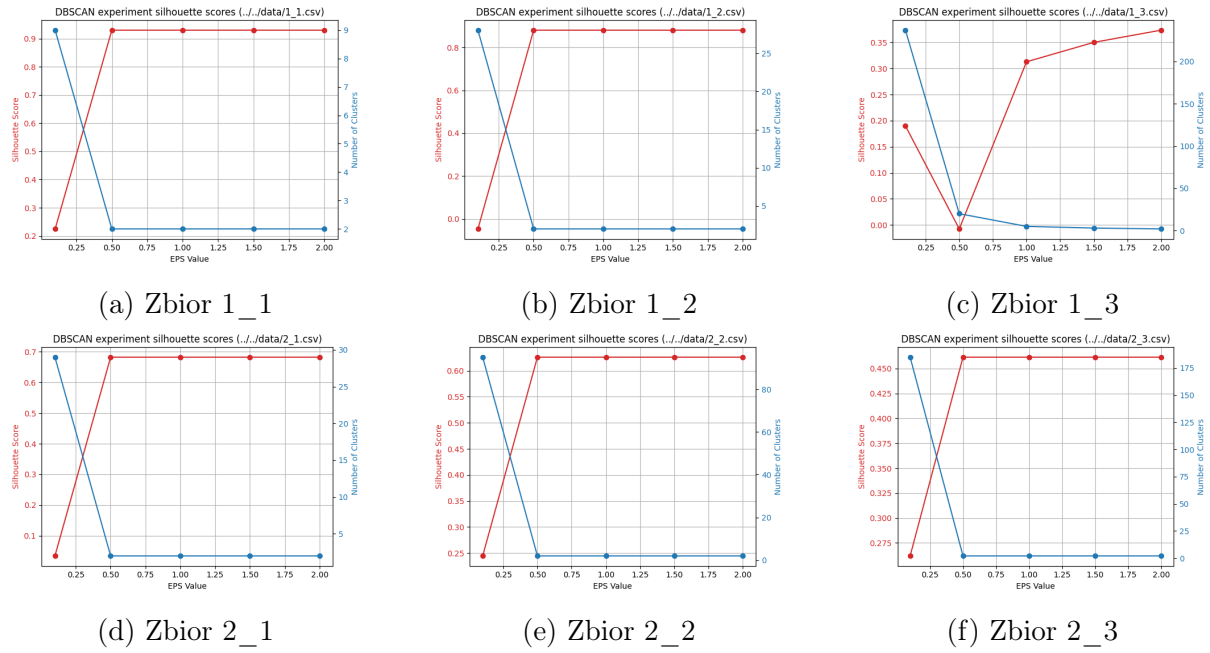


Rysunek 1: Zmiana wartości silhouette score dla wszystkich zbiorów w zależności od parametru `n_clusters` w metodzie K-means

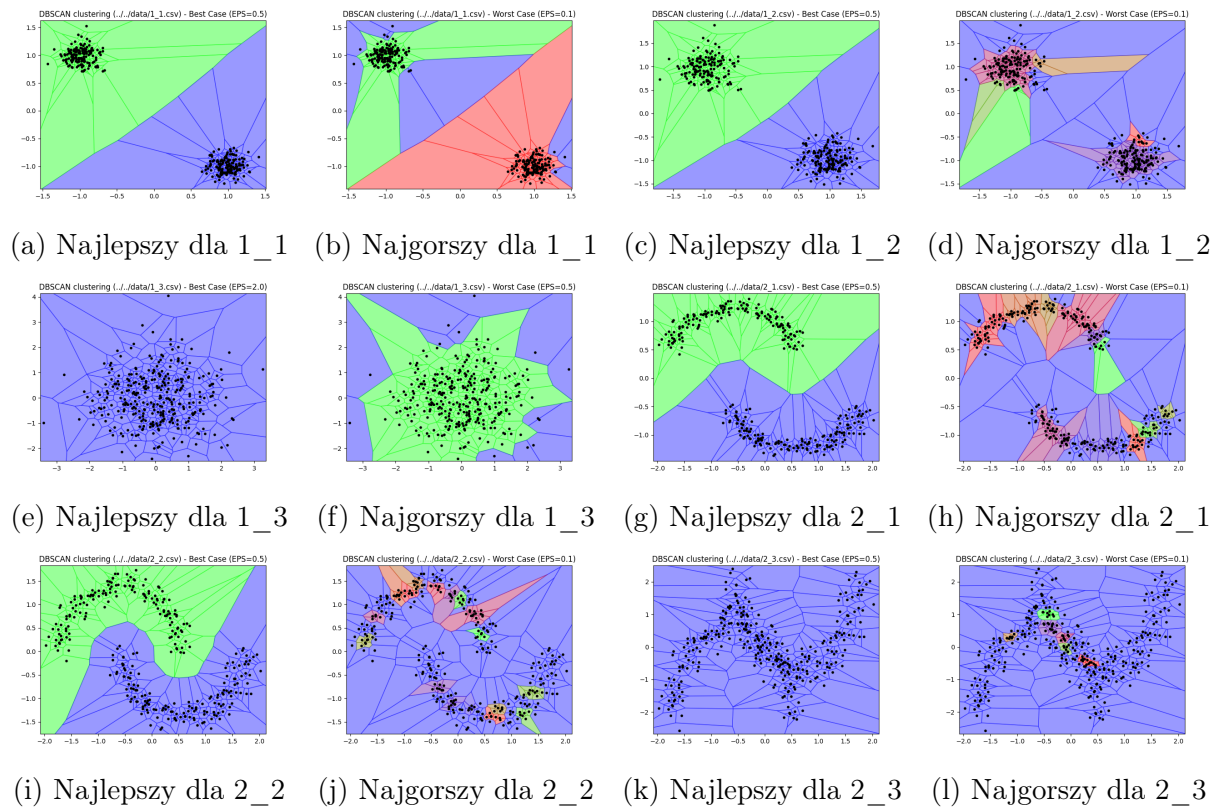


Rysunek 2: Wizualizacja klastrów dla wszystkich zbiorów na diagramie Voronoia dla najlepszego i najgorszego przypadku w metodzie K-means

## 2 Eksperyment 1: Metoda DBSCAN



Rysunek 3: Zmiana wartości silhouette score oraz n\_clusters dla wszystkich zbiorów w zależności od zmieniającego się parametru eps w metodzie DBSCAN



Rysunek 4: Wizualizacja klastrów dla wszystkich zbiorów na diagramie Voronoia dla najlepszego i najgorszego przypadku w metodzie DBSCAN

### 3 Eksperyment 2: Metoda K-Means z etykietami

Wyniki drugiego eksperymentu dla sześciu sztucznie wygenerowanych zbiorów danych i metody K-Means. Dla każdego zbioru należy pokazać wykres obrazujący zmianę wartości miar adjusted rand score, homogeneity score, completeness score oraz V-measure score przy zmieniającym się parametrze n-clusters oraz wizualizację klastrów (diagram Woronoja z pokazanymi prawdziwymi etykietami obiektów) dla najlepszego i najgorszego przypadku (wskazując, który to był przypadek i dlaczego).

## 4 Eksperyment 2: Metoda DBSCAN z etykietami

Wyniki drugiego eksperymentu dla sześciu sztucznie wygenerowanych zbiorów danych i metody DBSCAN. Dla każdego zbioru należy pokazać wykres obrazujący zmianę wartości miar adjusted rand score, homogeneity score, completeness score oraz V-measure score przy zmieniającym się parametrze eps oraz wizualizację klastrów (diagram Woronoja z pokazanymi prawdziwymi etykietami obiektów) dla najlepszego i najgorszego przypadku (wskazując, który to był przypadek i dlaczego). W przypadku metody DBSCAN warto również wskazać na wykresie jaką liczbę klastrów uzyskano dla różnych wartości parametru eps. Przykładowy wykres:

## 5 Wnioski

Opis wniosków z eksperymentów przeprowadzonych na sześciu sztucznie wygenerowanych zbiorach. W przypadku pierwszego eksperymentu należy stwierdzić, jak, obserwując silhouette score, można dobrać właściwe parametry metod klasteryzacji, aby klastry dobrze odkrywały strukturę danych. Dla eksperymentu drugiego należy wskazać jakie informacje o separowalności klas można wyciągnąć obserwując wartości miar: adjusted rand score, homogeneity score, completeness score oraz V-measure score. Wnioski powinny mieć charakter ogólny, pozwalający przenieść je na przypadek, w którym nie ma możliwości zwizualizowania danych. Każdy wniosek powinien być poparty odniesieniami do wyników przedstawionych na pierwszych czterech stronach raportu.

### Eksperyment 1

### Eksperyment 2

.

### Analiza

Analiza pozostałych, rzeczywistych zbiorów danych powinna uwzględniać wnioski z poprzednich eksperymentów. Można je również uzasadnić poprzez odwołanie do wartości miar uzyskanych na tych zbiorach.

## 6 Analiza pozostałych zbiorów danych

Opis analizy pozostałych, rzeczywistych zbiorów danych, w której to zastosowane zostaną wnioski z wcześniejszych eksperymentów. Wyniki tej analizy należy również uzasadnić poprzez odwołanie do wartości miar uzyskiwanych na tych zbiorach (warto wykorzystać tabele i/lub wykresy i się do nich odwołać). W przypadku zbioru Iris da się swoje wnioski poprzeć wizualizacjami rzutów cech obiektów na dwuwymiarowe przestrzenie wybranych kombinacji dwóch z nich.

### Wnioski

Na podstawie przeprowadzonych eksperymentów możemy wnioskować o strukturze rzeczywistych zbiorów danych i skuteczności wybranych metod klasteryzacji.

### Zbiór Iris

### Zbiór Wine

### Zbiór Breast Cancer Wisconsin