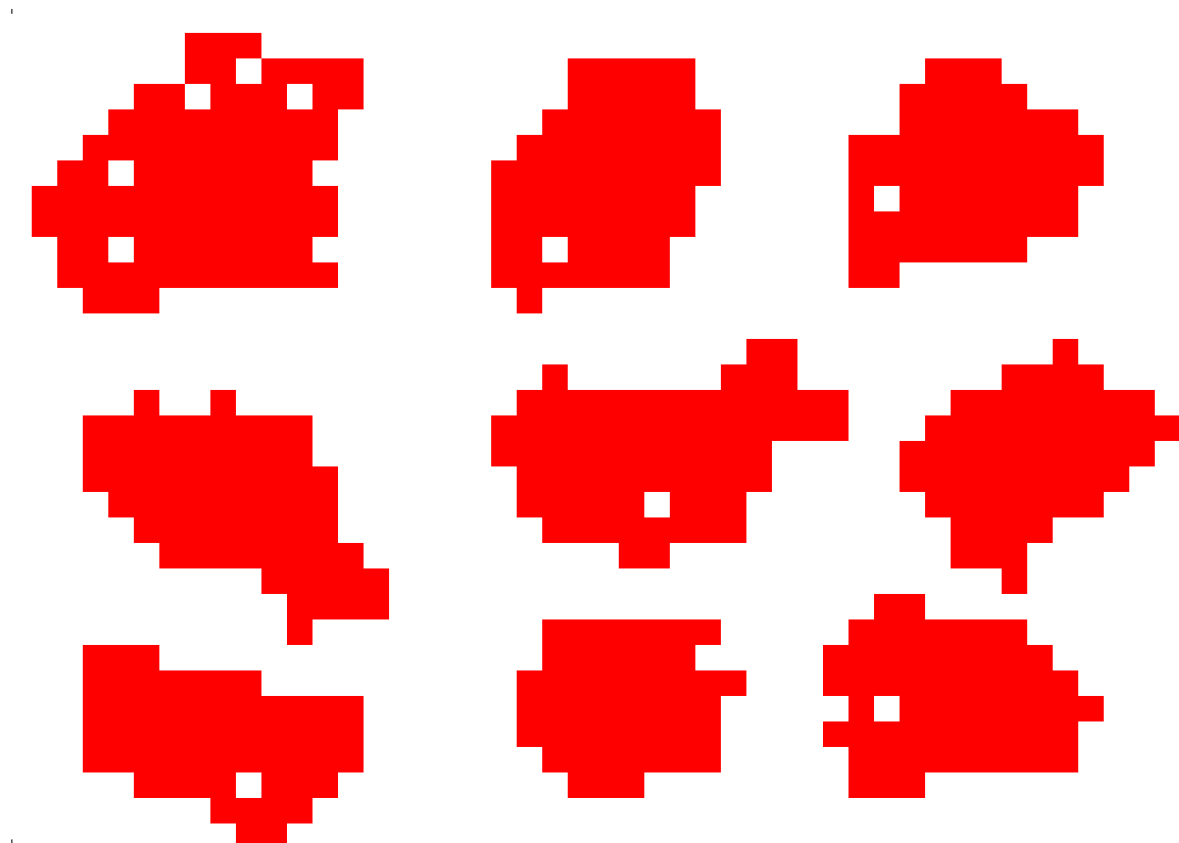


Metody rozpoznawania obrazów i podstawy uczenia maszynowego

Zadanie: Metody inicjalizacji k-means

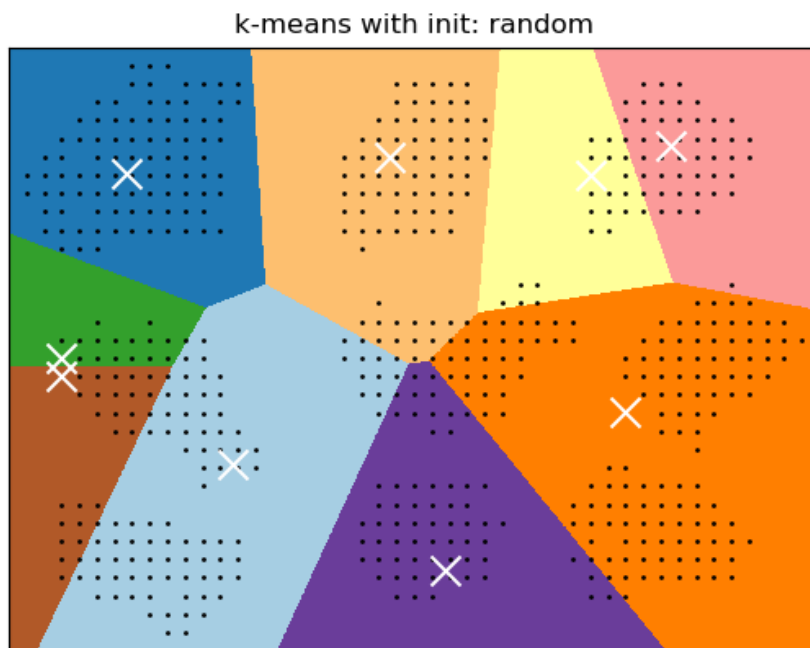
Wykonanie: Kamil Kurp

Zbiór wejściowy - bitmapa:

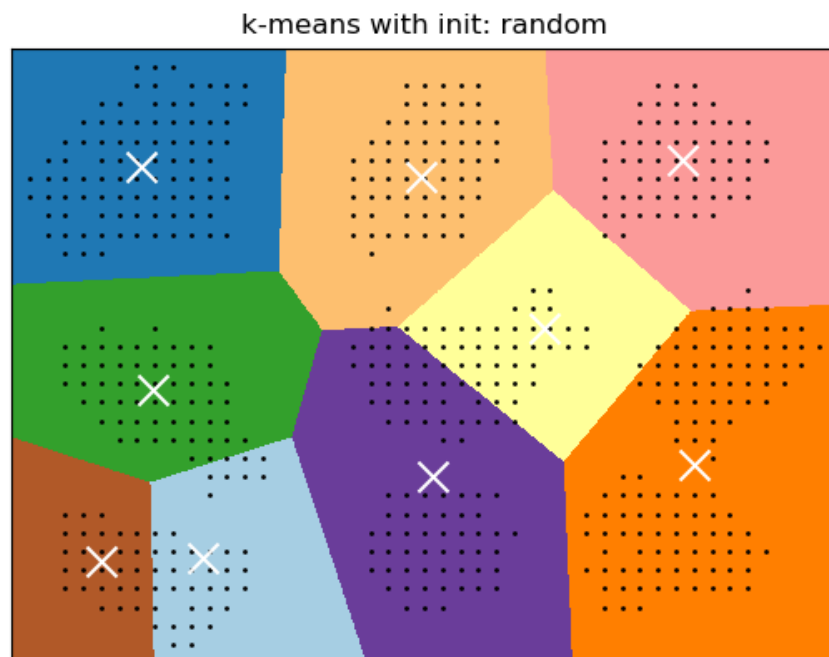


1. Inicjalizacja random

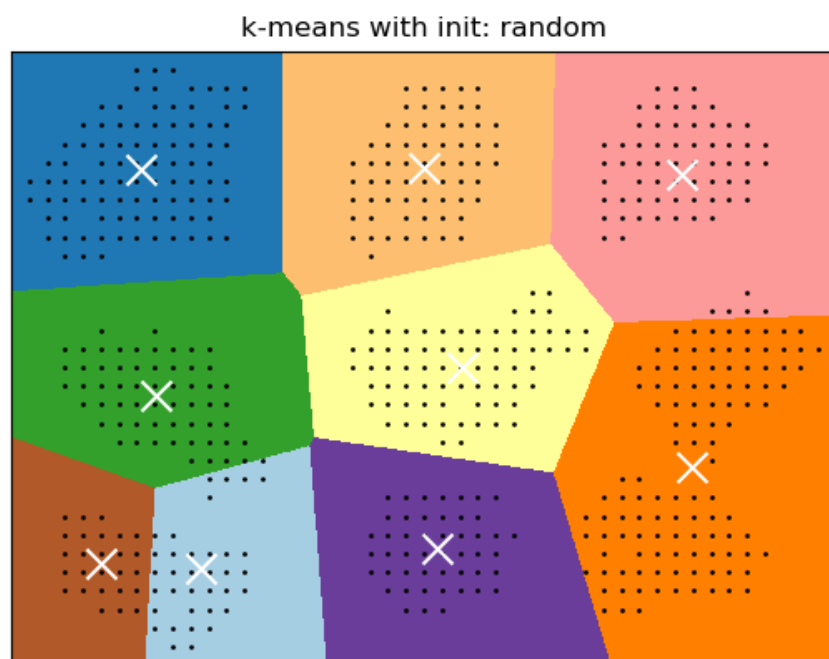
Pierwsza iteracja:



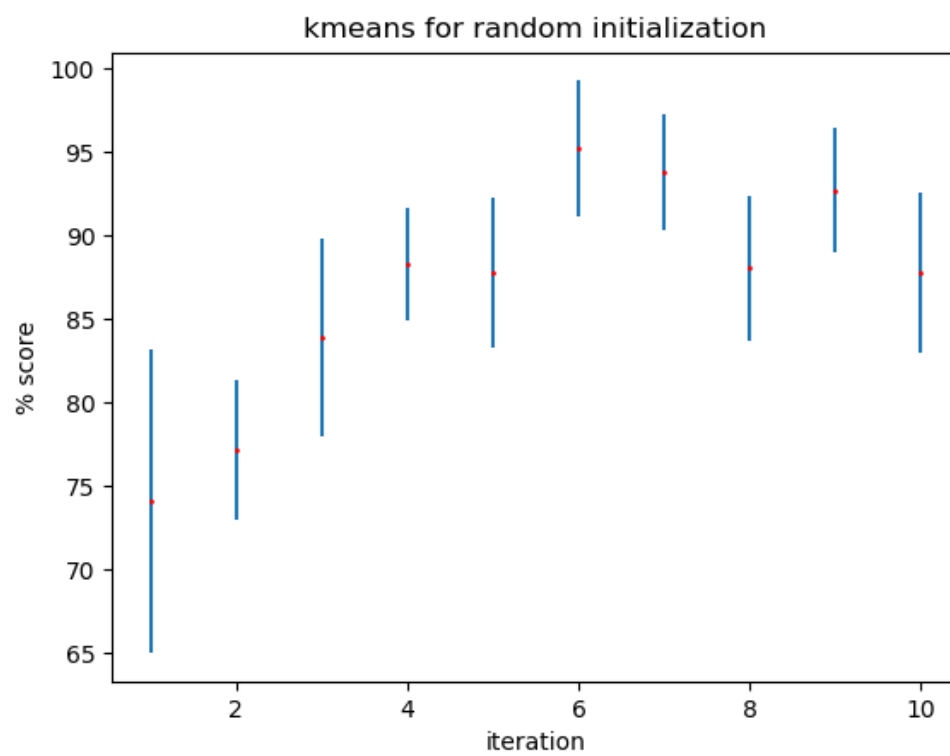
Piąta iteracja:



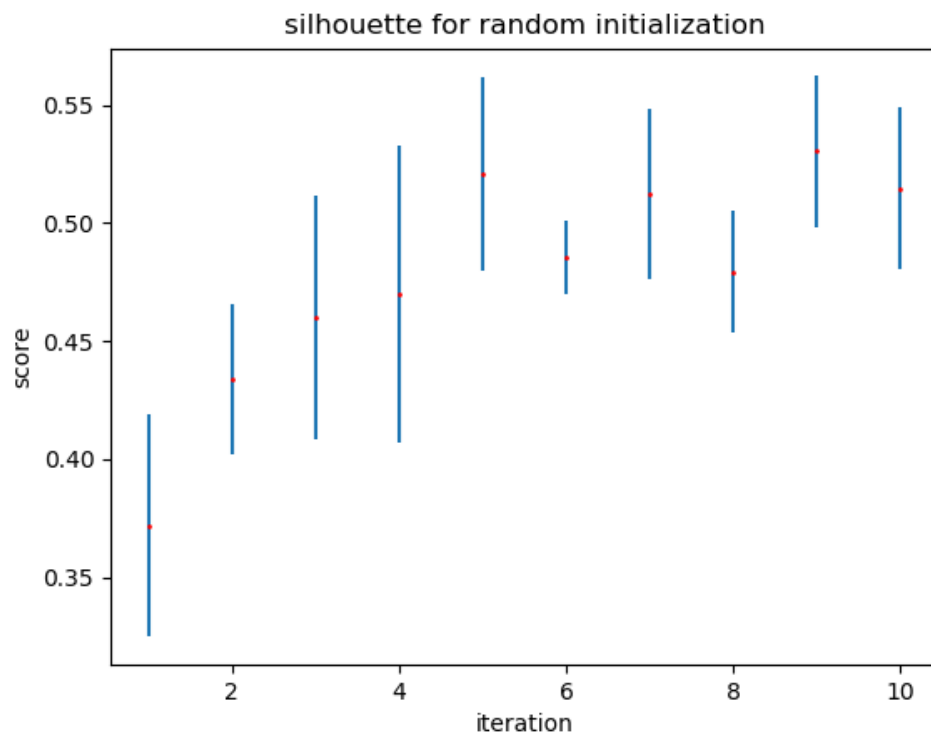
Dziesiąta iteracja:



Metryka jakości Adjusted Mutual Information (AMI):

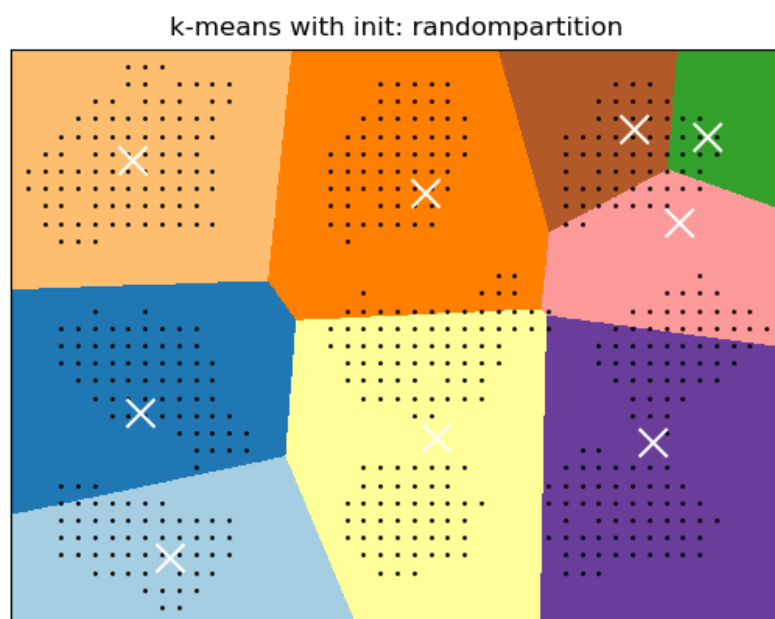


Miara Silhouette:

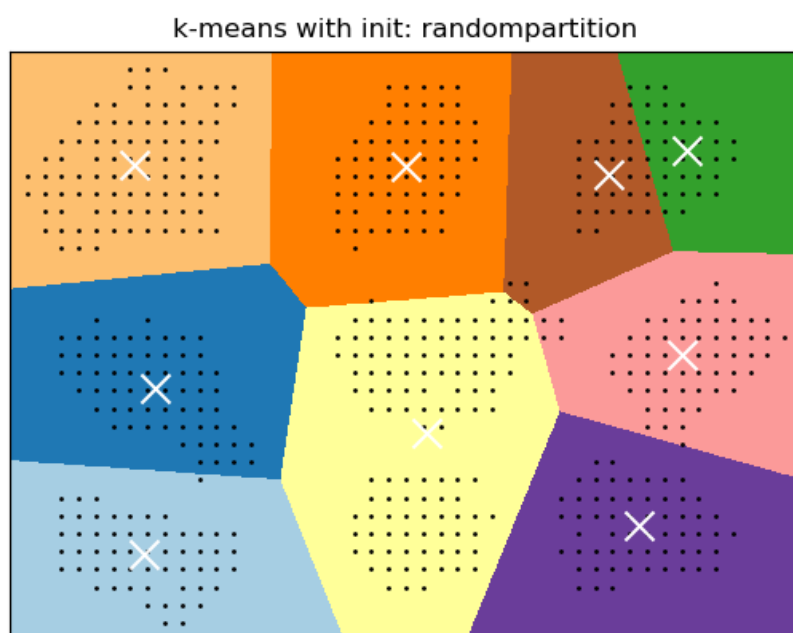


2. Inicjalizacja random partition

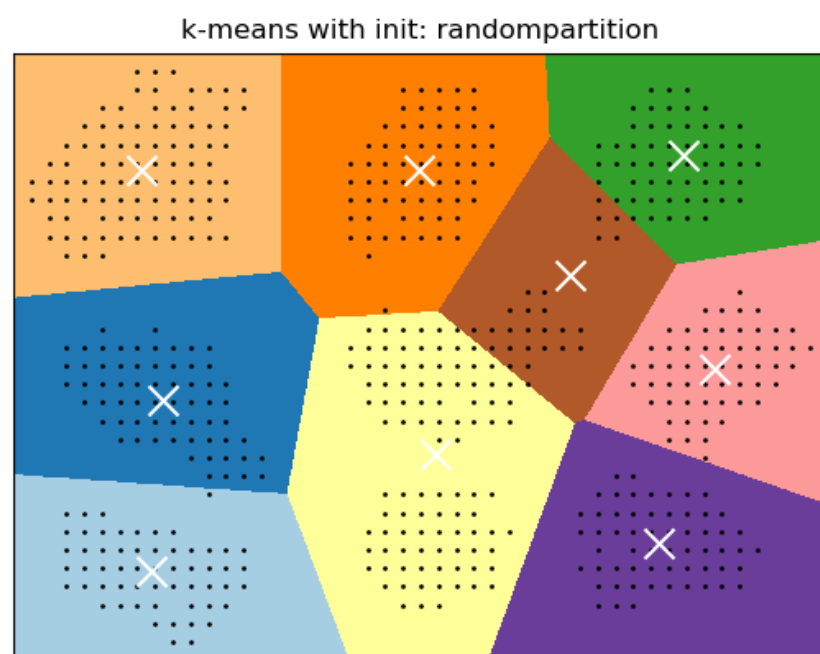
Pierwsza iteracja:



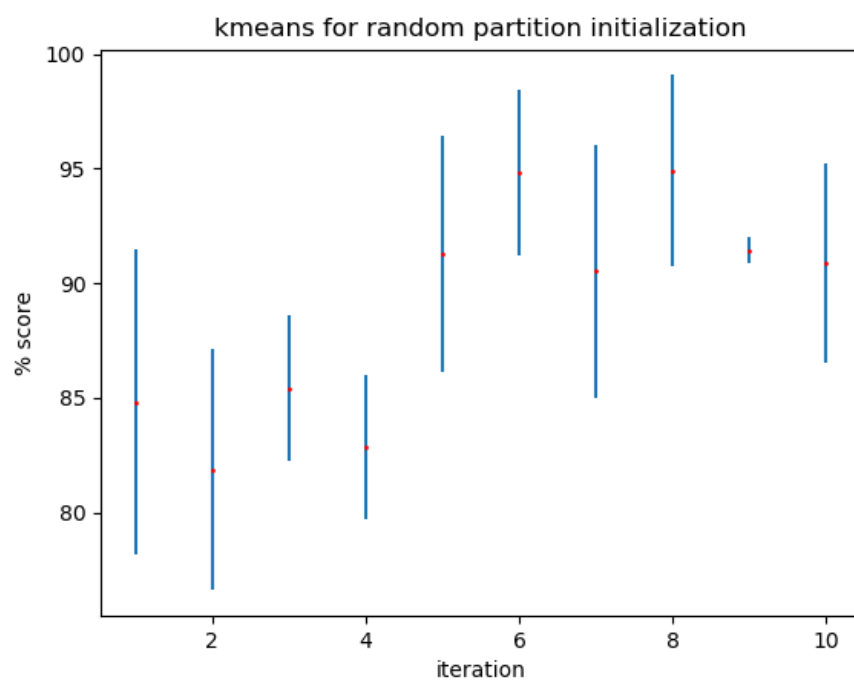
Piąta iteracja:



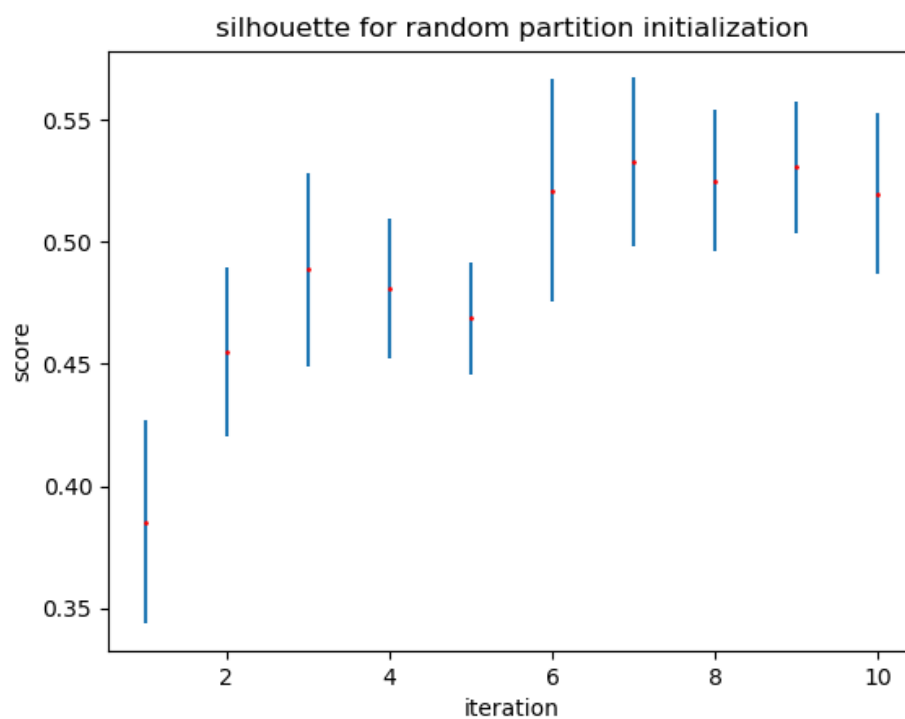
Dziesiąta iteracja:



Metryka jakości Adjusted Mutual Information (AMI):

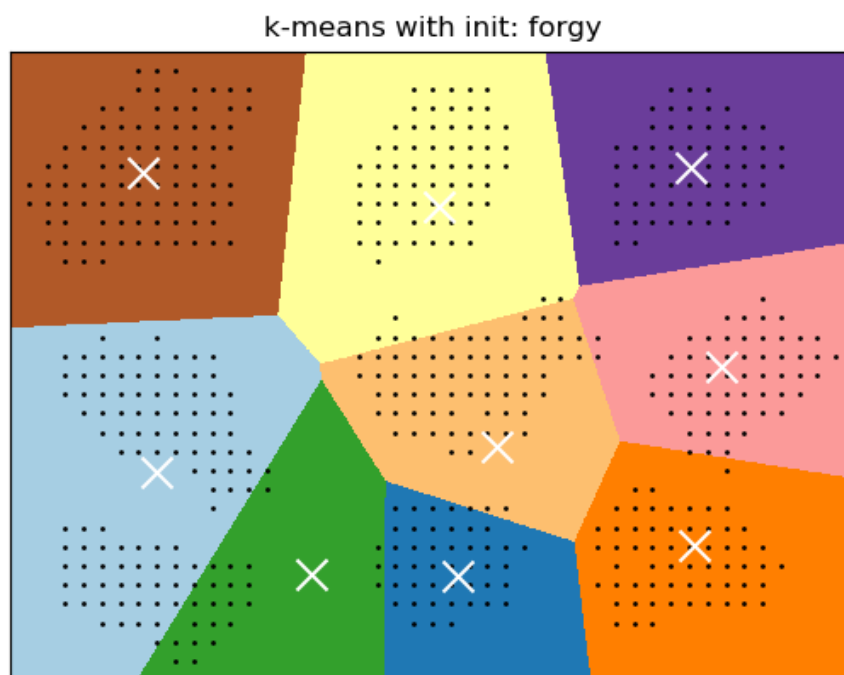


Miara Silhouette:

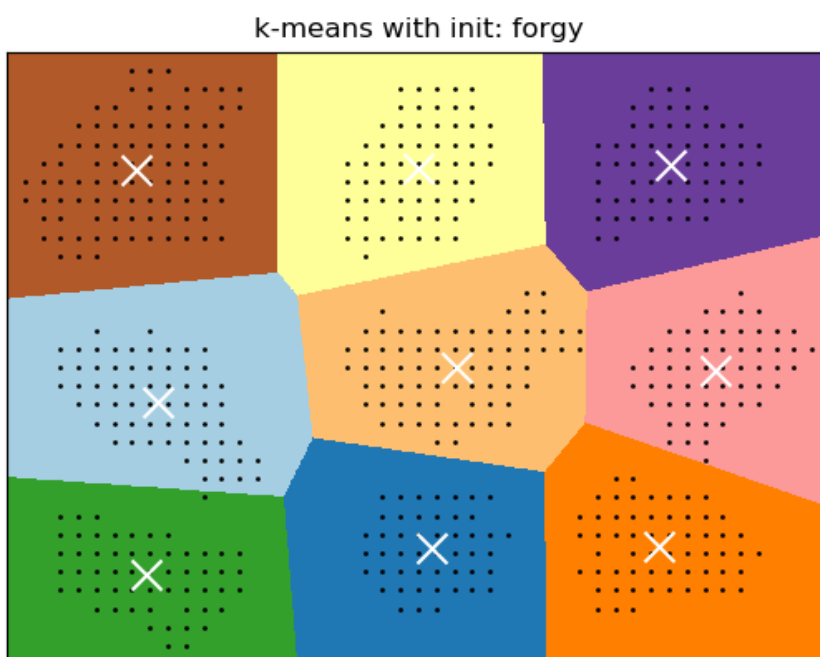


3. Inicjalizacja Forgy

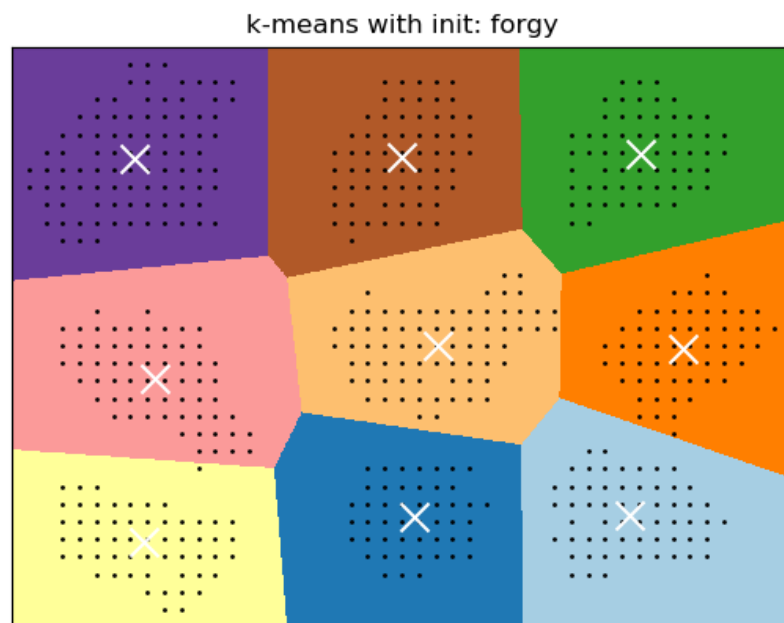
Pierwsza iteracja:



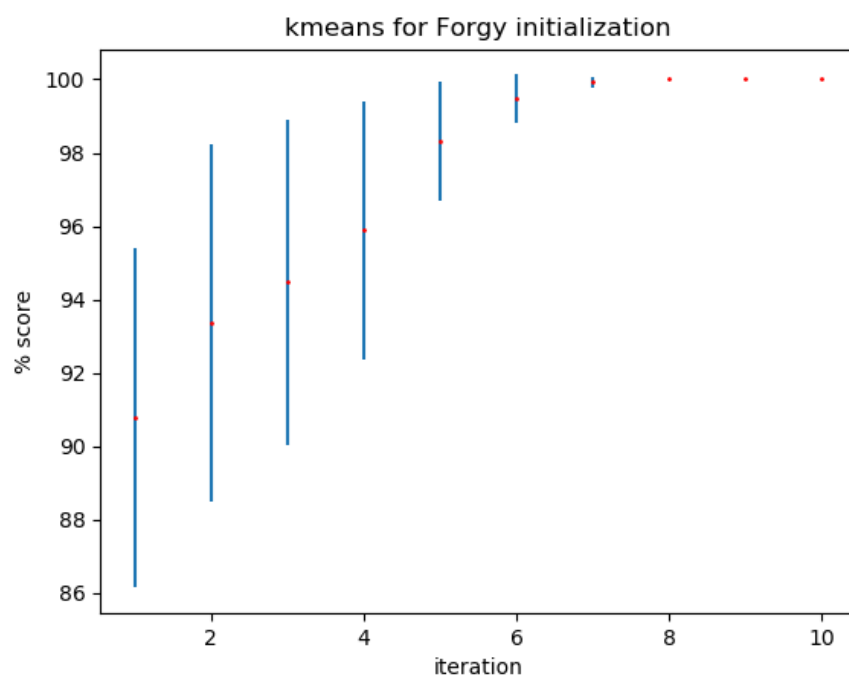
Piąta iteracja:



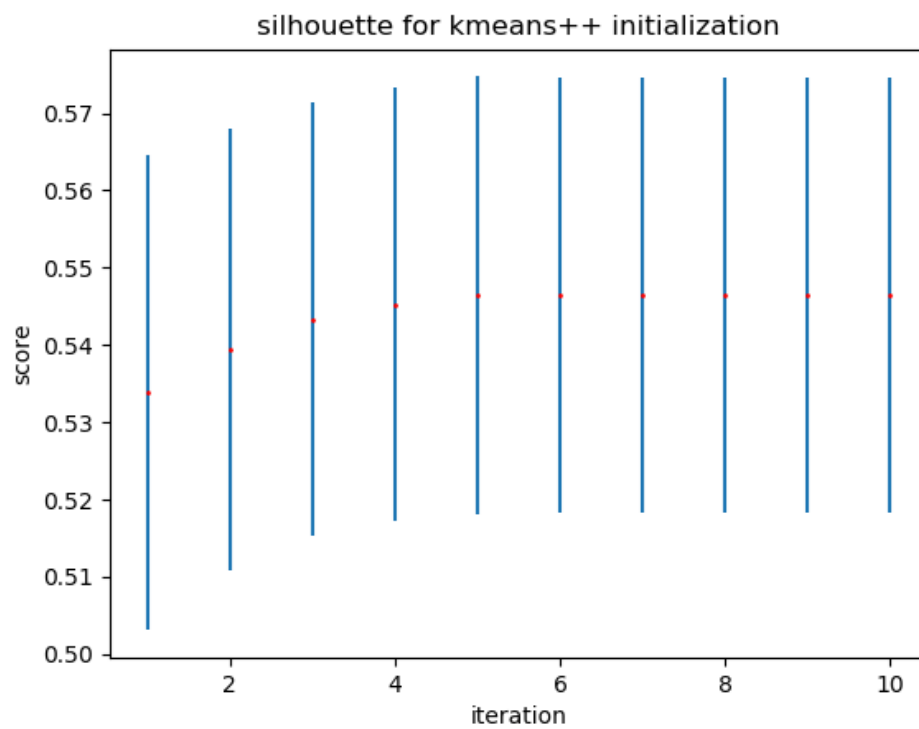
Dziesiąta iteracja:



Metryka jakości Adjusted Mutual Information (AMI):

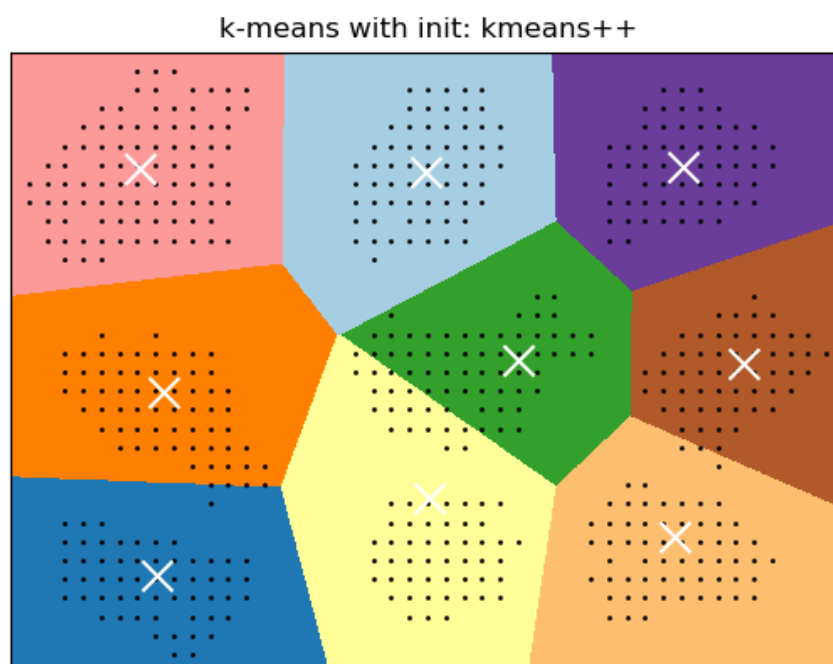


Miara Silhouette:

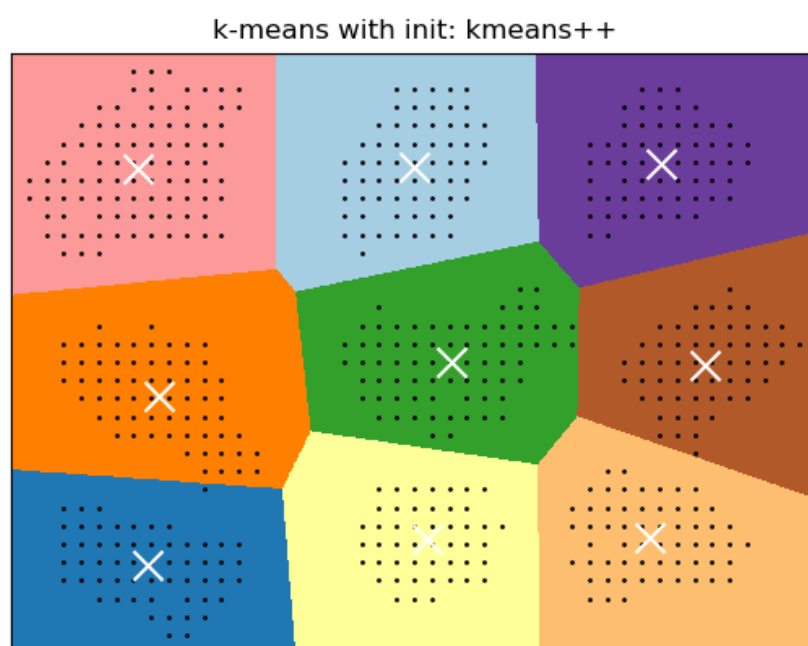


4. Inicjalizacja k-means++

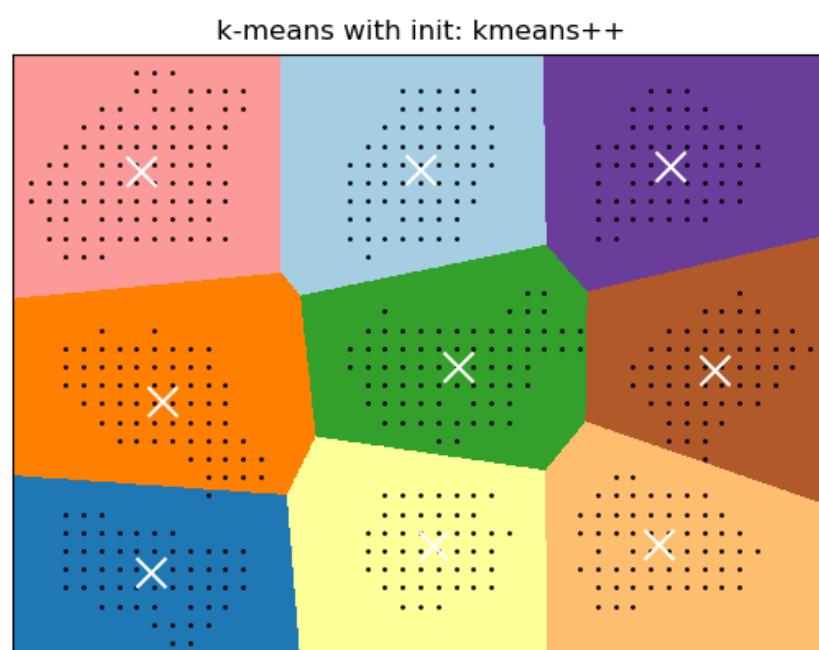
Pierwsza iteracja:



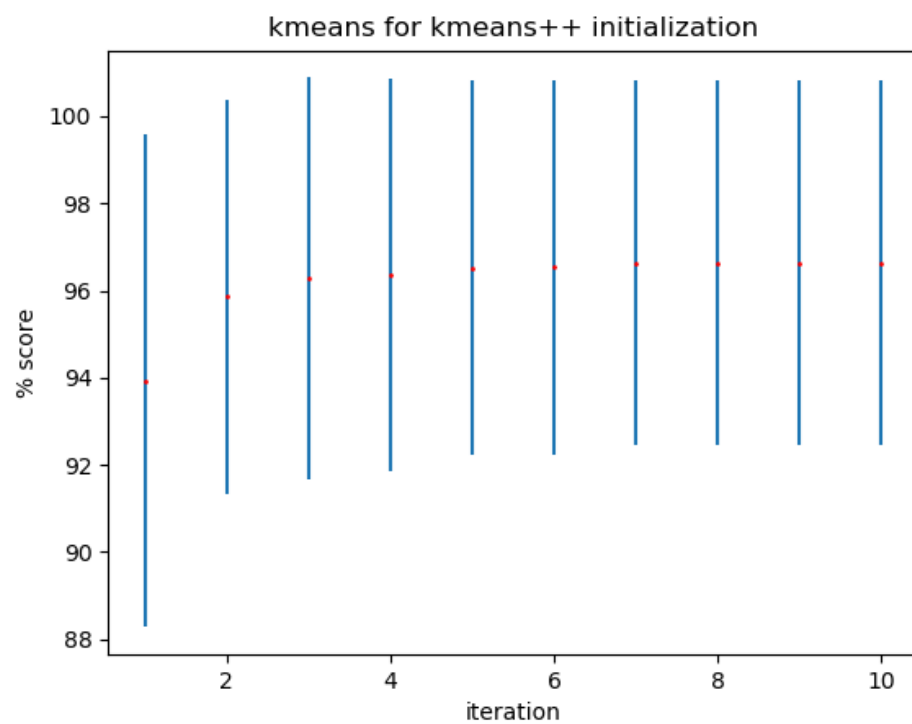
Piąta iteracja:



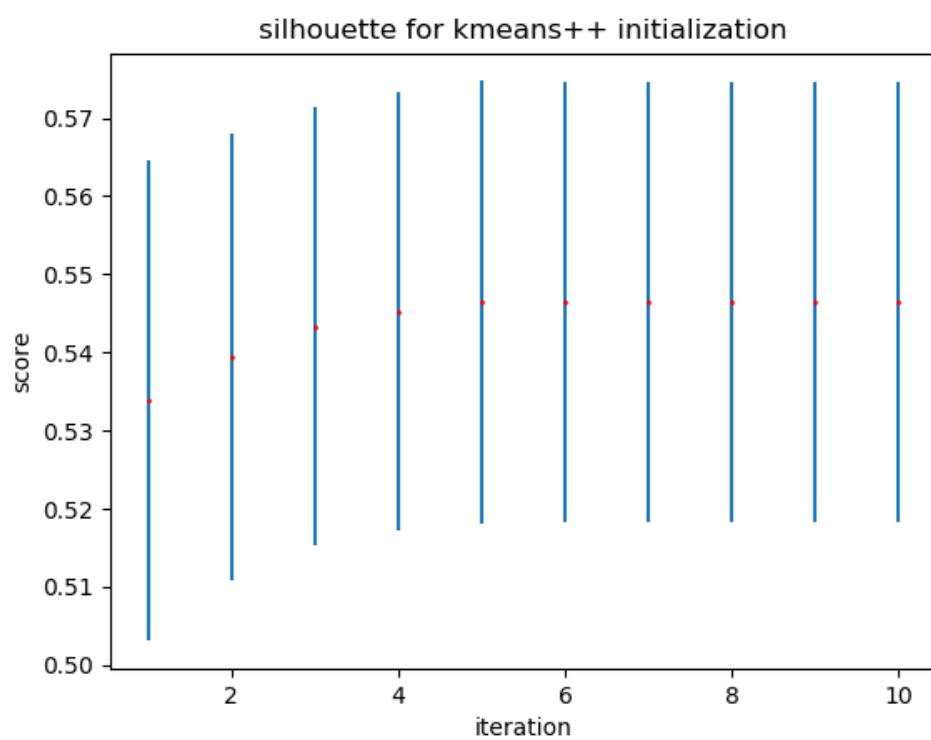
Dziesiąta iteracja:



Metryka jakości Adjusted Mutual Information (AMI):



Miara Silhouette:



5. Wnioski

Inicjalizacja random nie jest w stanie rozróżnić dwóch klastrów, które są bardzo blisko (algorytm wpada w lokalne minimum) ze względu na to, że początkowe punkty centralne klastrów zostały niefortunnie wylosowane. W random partition występuje podobny problem. Skutkuje to tym, że te metody inicjalizacji nie są w stanie osiągnąć 100% w punktacji metryki jakości. Inicjalizacja Forgy osiąga dla takich danych wejściowych maksymalną punktację w najmniejszej ilości iteracji ze wszystkich zastosowanych metod. K-means++ osiąga podobnie dobry wynik, jednak nie dla wszystkich prób inicjalizacji jest osiągnięte 100%.