

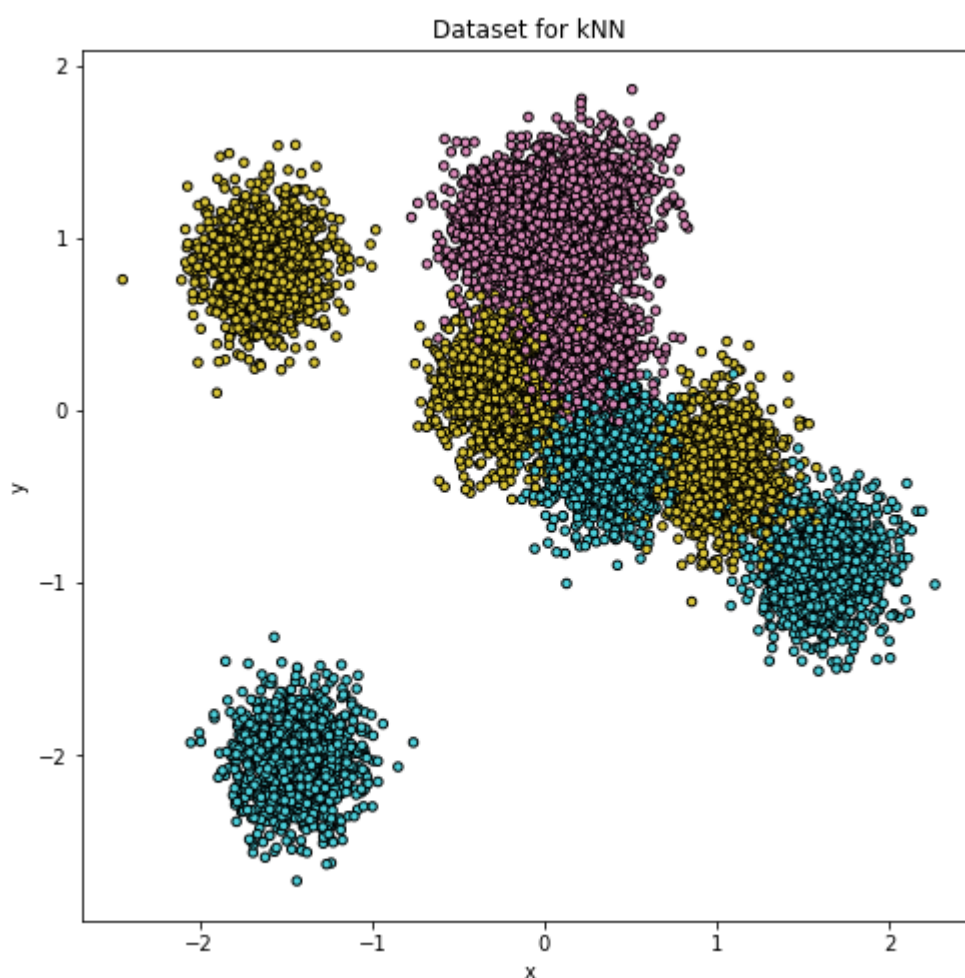
Podstawy Uczenia Maszynowego

Laboratorium 2

Maria Polak

1. Przygotowanie zbioru

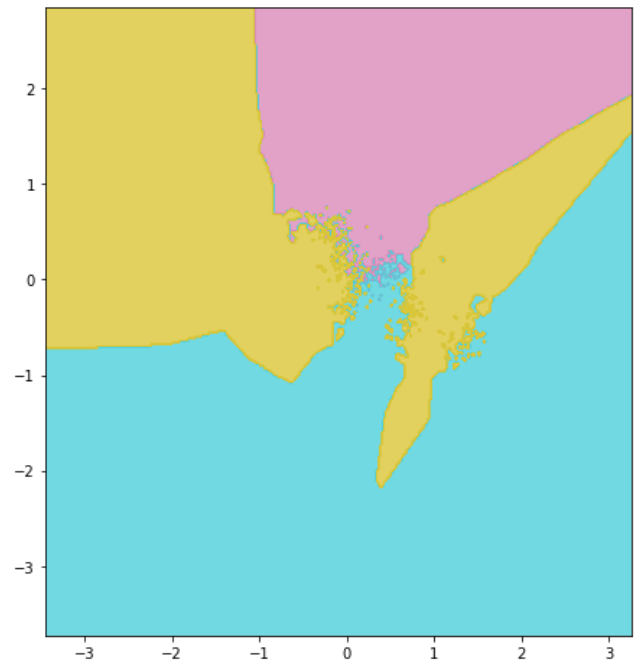
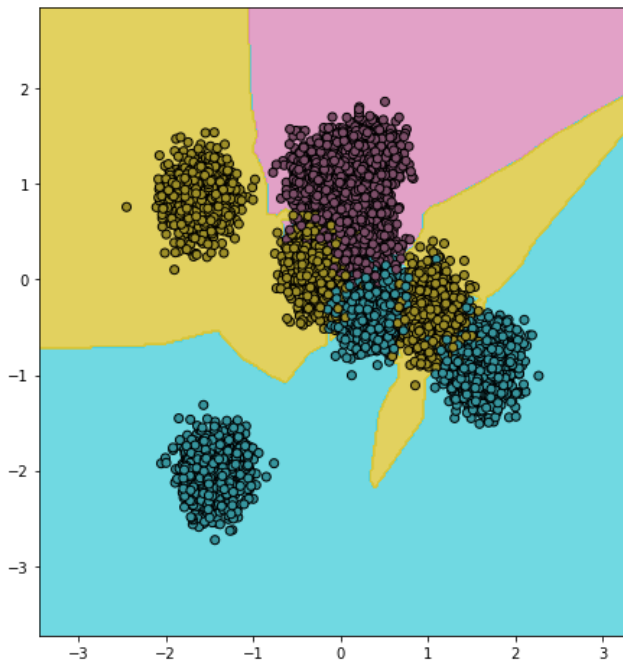
Zbiór danych został wygenerowany przy pomocy funkcji *make_blobs* z biblioteki *sklearn*. Wygenerowanych zostało 9 blobów, które następnie zostały podzielone na 3 klasy, dając zbiór widoczny na poniższym rysunku:



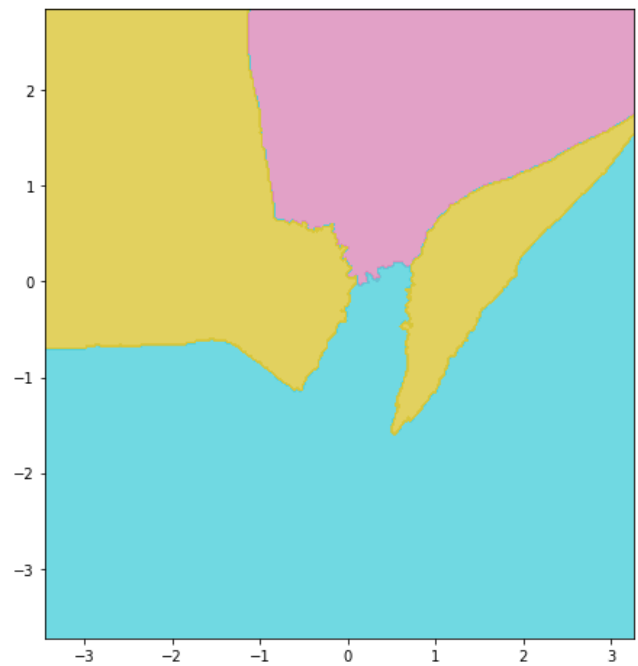
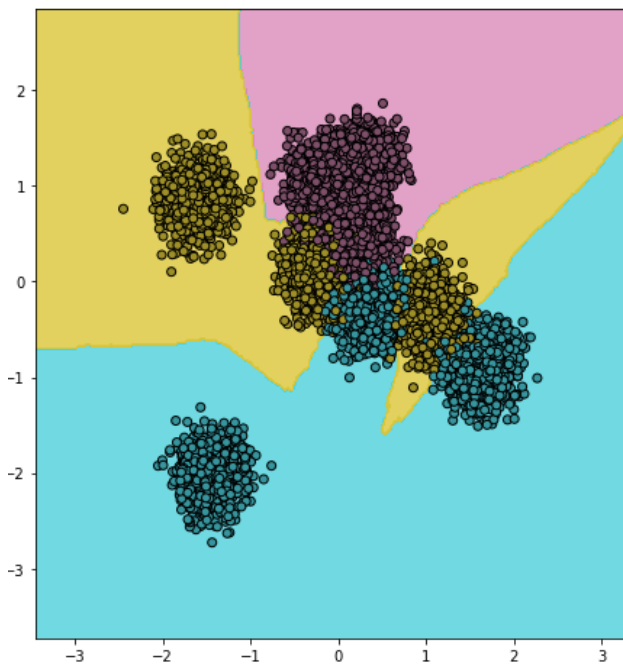
2. Porównanie metryk 'na oko'

Każdy z wariantów został wytrenowany. Wyniki były następujące:

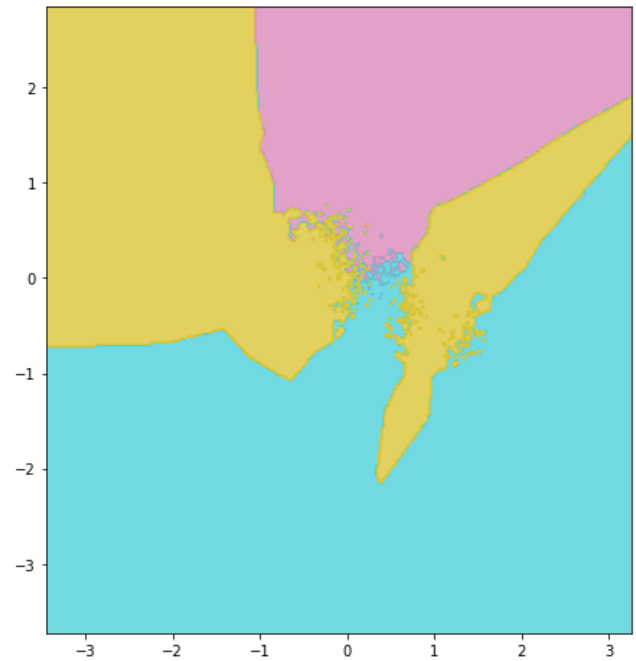
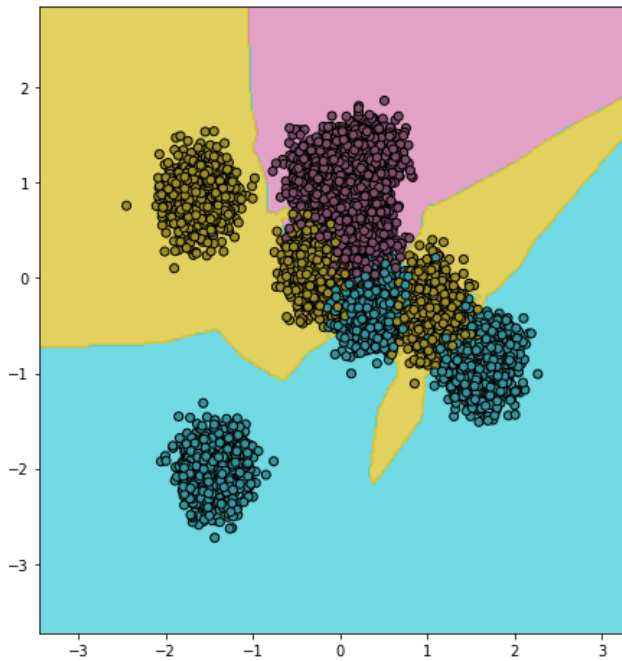
3-Class classification (k = 1, weights = 'uniform', metric = euclidean)



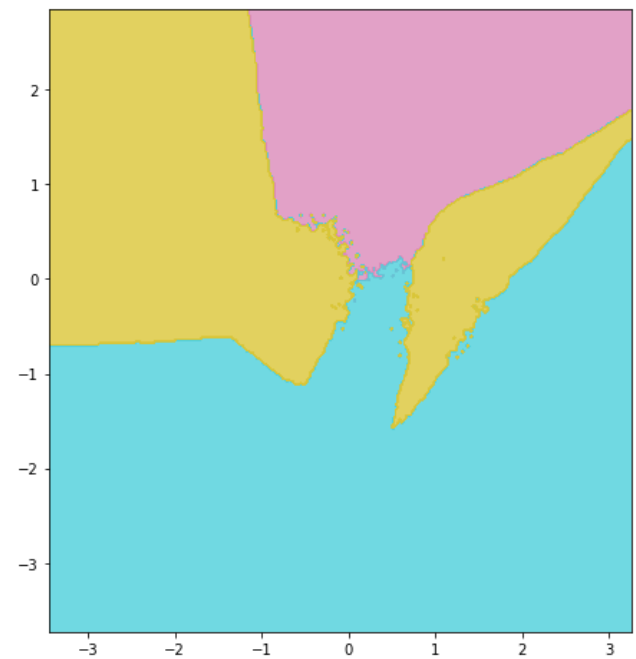
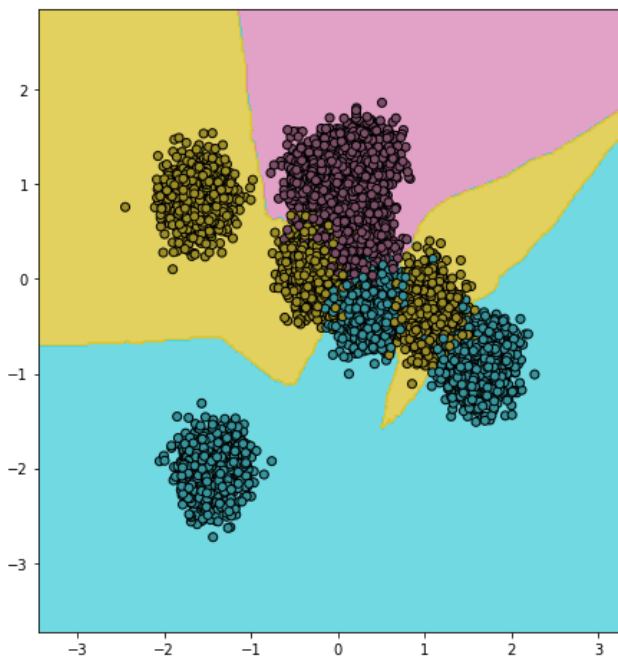
3-Class classification (k = 13, weights = 'uniform', metric = euclidean)



3-Class classification ($k = 1$, weights = 'uniform', metric = mahalanobis)



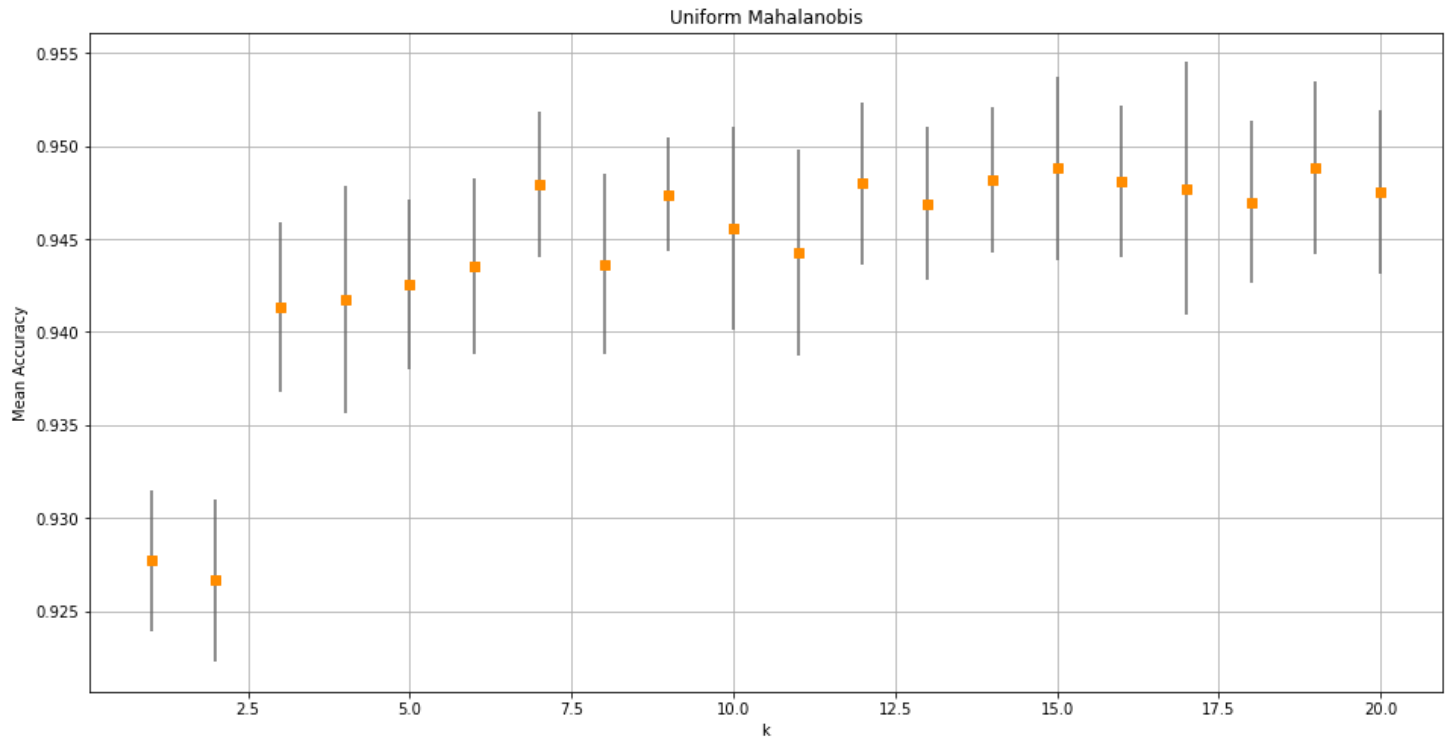
3-Class classification ($k = 9$, weights = 'distance', metric = euclidean)



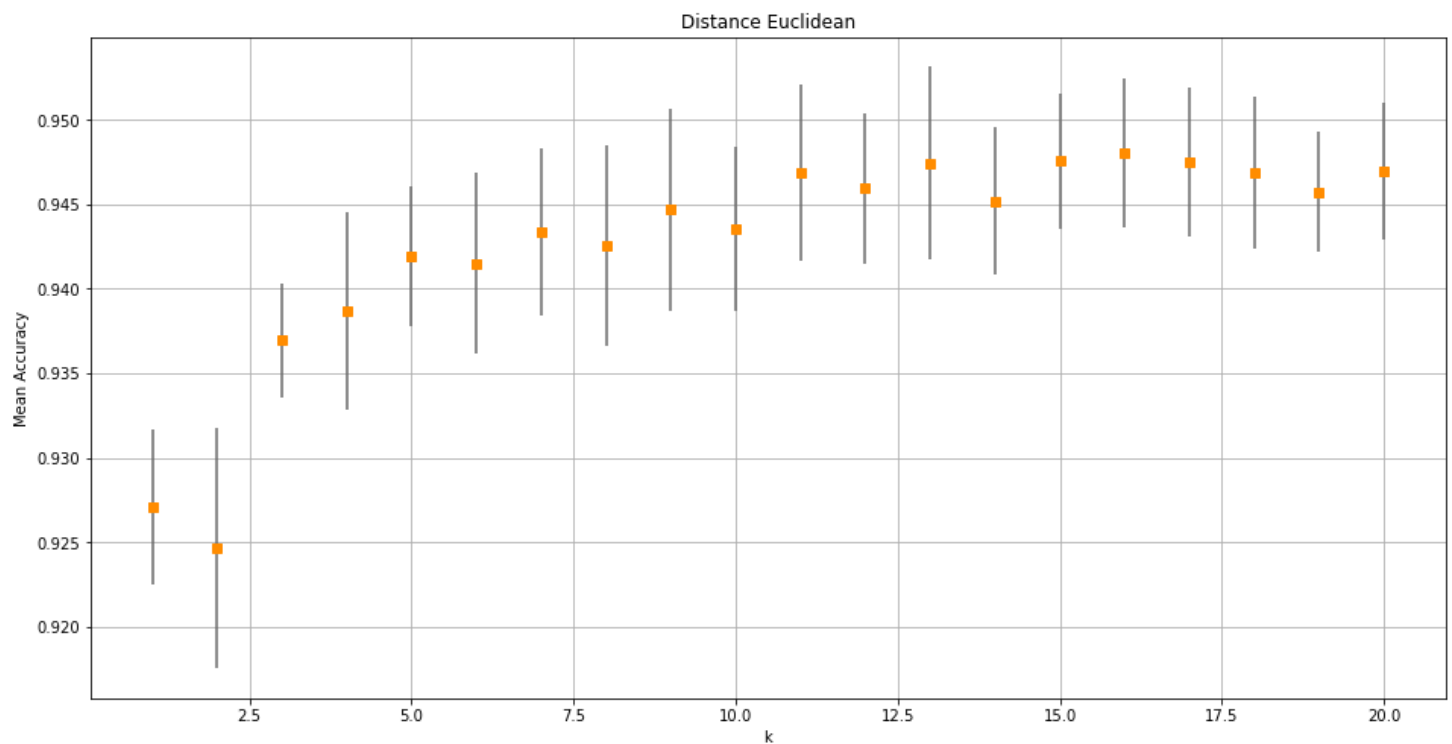
Jako najgorszy został wybrany wariant 3 (k-NN z $k=1$, głosowaniem większościowym i metryką Mahalanobisa), a jako najlepszy wariant 4 (k-NN z $k=9$, głosowaniem ważonym odległością i metryką Euklidesa)

3. Porównanie “najgorszej” i “najlepszej” metryki

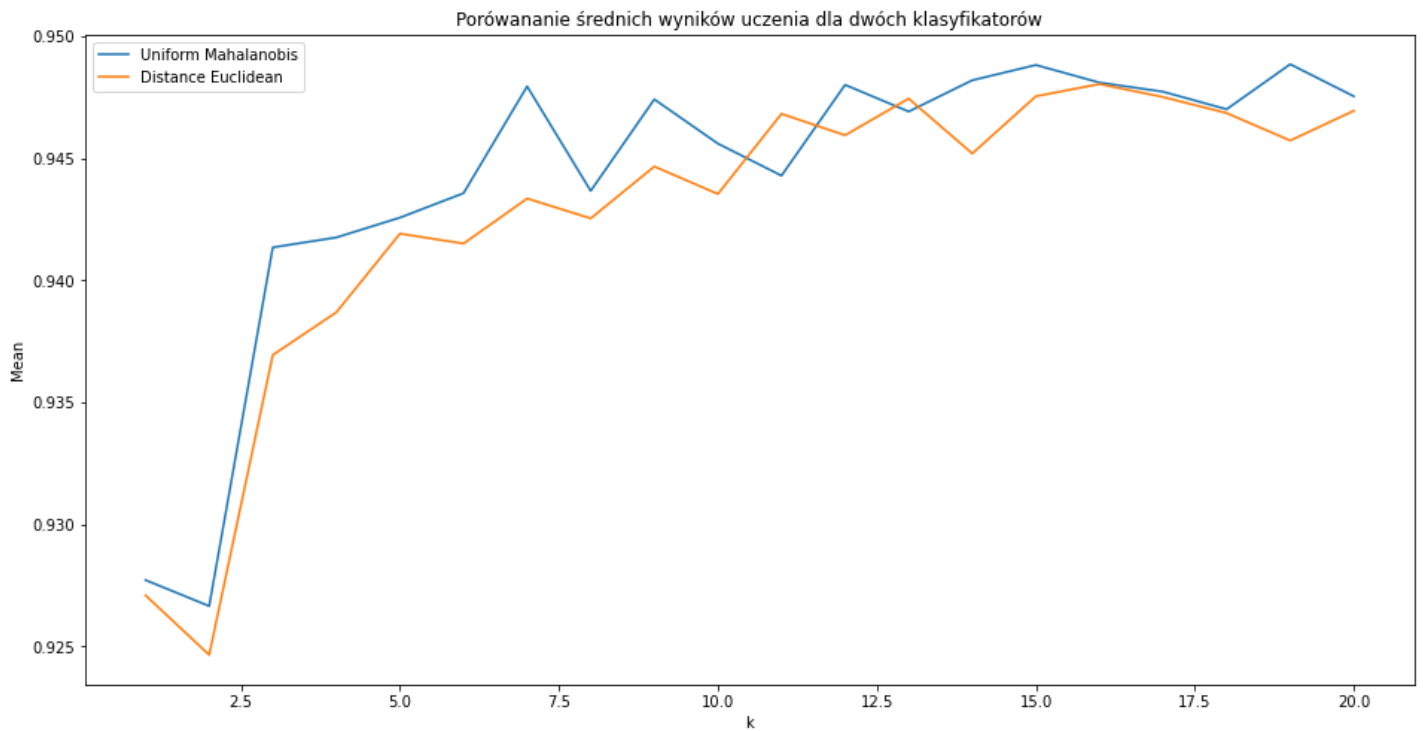
Dla obu wariantów klasyfikator został wytrenowany 20 razy (dla każdego k).



K with highest value = 19

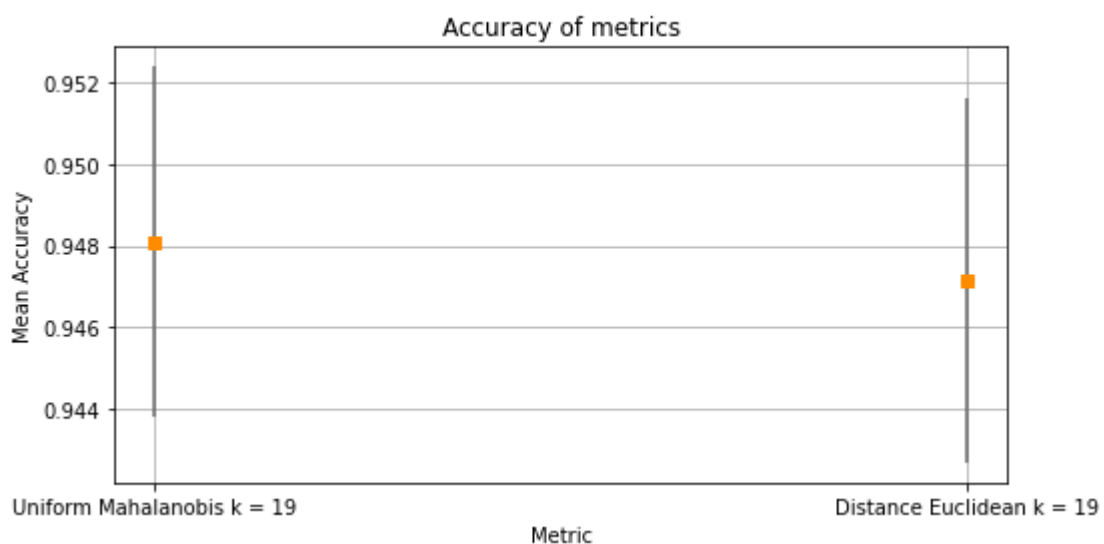


K with highest value = 16



Na powyższym wykresie możemy zauważyć, że wariant, który “na oko” był uznany jako gorszy przeważnie zwracał dokładniejsze wyniki, niż wariant uznany jako lepszy.

Następnie klasyfikator został wytrenowany dla najwyższych wartości k (19 dla Mahalanobis z głosowaniem większościowym i 16 dla Euclidean z głosowaniem ważonym odległością) i przetestowany 10-krotnie.



Wyniki prezentują się następująco:

Wariant	Średnia skuteczność	Odchylenie standardowe
Uniform Mahalanobis k = 19	94.81 %	0.43 %
Distance Euclidean k = 19	94.71 %	0.45 %

W praktyce wariant Mahalanobisa okazał się lepszy niż Euklidesa. Skuteczność Mahalanobisa jest o 0.1% większa i jednocześnie odchylenie standardowe wyników jest nieznacznie mniejsze (0.02%).