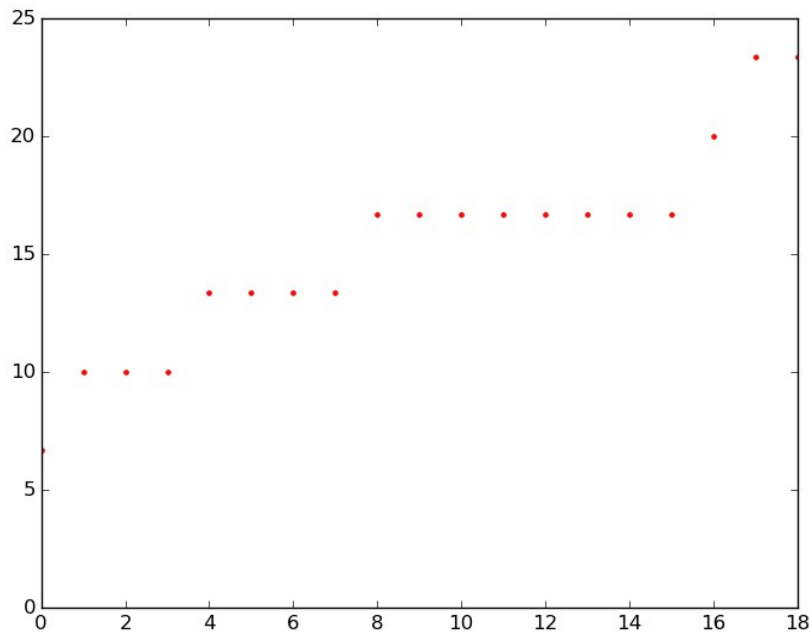


MRO

Paulina Żak
Laboratorium 1

4.1-NN

1. Funkcja realizująca klasyfikację 1-NN została napisana z użyciem metody `scipy.spatial.distance.euclidean`
2. Błąd klasyfikacyjny na zbiorze `facesYale.mat` wynosi 6.67%
3. Wykres błędu klasyfikacyjnego przy skalowanej cesze 10 przedstawia się następująco

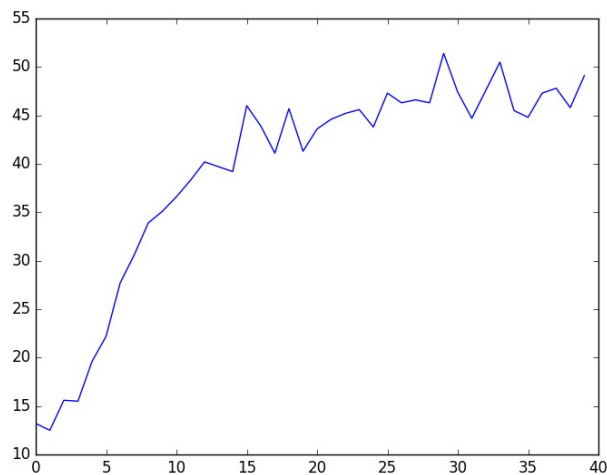


4. Błąd klasyfikacyjny na zbiorze `spambase.mat` wynosi: 17.905021586%
5. Błąd klasyfikacyjny na zbiorze `spambase.mat` przy wykorzystaniu walidacji krzyżowej wynosi:

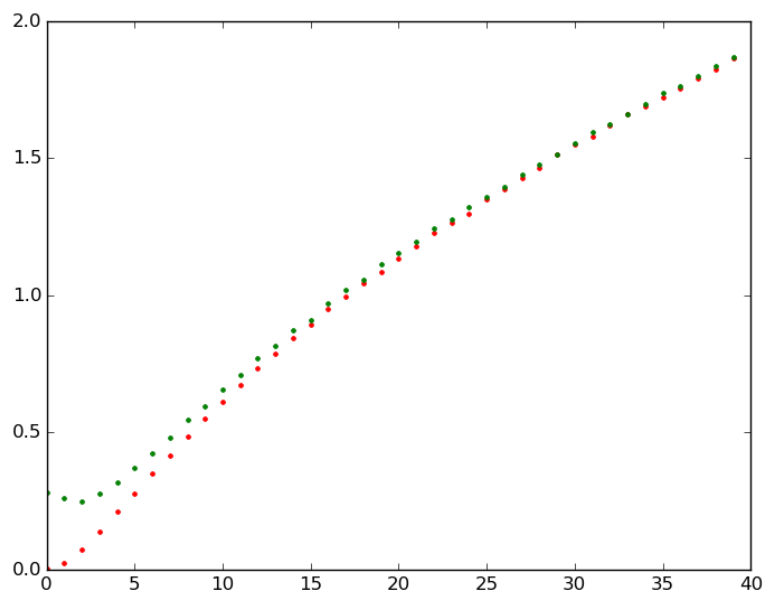
6.40608034745%
9.88056460369%
9.01194353963%
9.33767643865%
9.66340933768%
9.01194353963%
10.8577633008%
9.01194353963%
8.5776330076%
9.9891422367%

5.1-NN w wielu wymiarach

1. Błąd klasyfikacji w zależności od wymiarowości przestrzeni przedstawia się następująco



2. Średnie odległości od najbliższego sąsiada do klasy poprawnej oraz przeciwnej



Wraz z zwiększaniem wymiarowości, różnice między sąsiadami poprawnymi i niepoprawnymi się zacierają. Jest to przykład “przekleństwa wymiarowości”. Odległość euklidesowa traci użyteczność wraz z wzrostem wymiarowości, gdyż wektory do szukanego punktu stają się coraz bardziej podobne do siebie.

Poniżej znajduje się wykres stosunków obu średnich.

