

## NAI – mały projekt programistyczny 1

### Klasyfikator $k$ -NN

Napisz program, który wykonuje algorytm  $k$ -NN dla danych w formacie **Iris** (dane są w pliku `iris.csv`, zaś opis formatu w pliku `iris_description.txt`). Pobierz plik z danymi i podziel go na część treningową (po 35 przykładów dla każdej klasy irysów) oraz testową (pozostałe 45 przykładów).

Program powinien pobierać następujące argumenty:

- `k`: dodatnia liczba naturalna będąca hiperparametrem  $k$ -NN
- `train-set`: nazwa pliku zawierającego zbiór treningowy (w formacie CSV jak wyżej)
- `test-set`: nazwa pliku zawierającego zbiór testowy (w formacie CSV jak wyżej)

Program ma dokonać klasyfikacji  $k$ -NN wszystkich obserwacji z pliku `test-set` na podstawie pliku `train-set` oraz podać dokładność (accuracy) tej klasyfikacji.

Program ma też dostarczać testowy interfejs (niekoniecznie graficzny), który umożliwia (zapętlone) podawanie przez użytkownika pojedynczych wektorów do klasyfikacji i podaje ich etykietę  $k$ -NN na podstawie `train-set`.

Opcjonalne rozszerzenie (przyznam za to punkty z aktywności):

Wykonaj wykres zależności accuracy od wartości  $k$  (dowolna technika - Excel, Python, R, itp.).

Uwaga: Program należy wykonać samodzielnie. **Plagiat** lub **niezrozumienie** rozwiązania skutkuje **brakiem zaliczenia projektu**.

Nie można korzystać z gotowych bibliotek do uczenia maszynowego ani operacji na wektorach. Wszystkie szczegóły algorytmu  $k$ -NN należy samodzielnie przećwiczyć kodując.