

MPP 1

Termin oddania: 1 tydzień.

Dane wejściowe:

Dane treningowe – plik iris_training.txt

Dane testowe – plik iris_test.txt

Program musi wczytać dane z podanego pliku tekstowego. Zakładamy, że:

- Atrybut decyzyjny znajduje się w ostatniej kolumnie.
- Wszystkie atrybuty poza decyzyjnym są numeryczne.

Program musi akceptować dowolną liczbę atrybutów warunkowych, tzn. nie może zakładać, że ich jest ustalona liczba.

Następnie program wczyta wartość parametru k od użytkownika, a potem zaklasyfikuje wszystkie przykłady wczytane z pliku tekstowego zawierającego zbiór testowy algorytmem k -NN.

Jako wynik ma wypisać liczbę prawidłowo zaklasyfikowanych przykładów oraz dokładność eksperymentu wyrażoną w procentach.

Program musi umożliwiać wielokrotne ręczne wpisanie wektora atrybutów i wypisać dla takiego wektora jego wynik klasyfikacji k -NN.

Opcjonalnie można dodać wykres (np. w Excelu) zależności uzyskanej dokładności od k oraz krótka dyskusję.

Nie można używać żadnych bibliotek ML, wszystko ma być zaimplementowane od zera w pętlach, if-ach, odległość trzeba liczyć za pomocą działań arytmetycznych, etc.

Zadanie do wykonania na ćwiczeniach:

- ~~1. Wczytać dane ze zbioru treningowego.~~
 - ~~2. Wczytać dane ze zbioru testowego.~~
 - ~~3. Dla pierwszej obserwacji ze zbioru testowego policzyć jej odległości euklidesowe do wszystkich obserwacji w zbiorze treningowym.~~
 - ~~4. Ze znalezionych w pkt. 3 odległości znaleźć 3 najmniejsze. Wydrukować wartości decyzyjne obserwacji, których one dotyczą.~~
1. Wczytać z klawiatury dwa wektory 4-wymiarowe.
 2. Wydrukować odległość euklidesową/miejską między nimi.