

Egzamin, termin I

zadanie algoryt- miczne

imię i nazwisko _____

liczba punktów _____ / 10

dokument klasy **B1**, archiwizować do **2019-01-01**

Dany jest plik tekstowy zawierający współrzędne n dwuwymiarowych punktów wraz z przyporządkowanymi etykietami:

x_0	y_0	e_0
x_1	y_1	e_1
x_{n-1}	y_{n-1}	e_{n-1}

Współrzędne x i y to liczby zmiennoprzecinkowe podwójnej precyzji, etykieta e to nieujemna liczba całkowita. Wiadomo, że $n \leq N$, oraz $e \in \{0, 1, \dots, E-1\}$, gdzie N i E to pewne znane w programie stałe.

Zadanie 1. [1 pkt]

Zdeklaruj strukturę `Obiekt` reprezentującą punkt wraz z przyporządkowaną mu etykietą.

Zadanie 2. [2 pkt]

Zdefiniuj funkcję `int wczytajObiekty(const string& wejscie, Obiekty obiekty[])`, która wczytuje obiekty z pliku `wejscie` i umieszcza je w tablicy `obiekty` (założyć, że tablica posiada odpowiedni rozmiar). Zwrócić z funkcji liczbę n wczytanych obiektów (-1 w przypadku braku istnienia pliku; 0 gdy plik jest pusty). Jeżeli plik istnieje, to ma poprawny format. Składniki oceny:

- otwarcie pliku, sprawdzenie czy istnieje: 1 pkt,
- wczytanie danych do tablicy, zwrócenie wartości: 1 pkt.

Zadanie 3. [2 pkt]

Zdefiniuj funkcję `odleglosc` wyznaczającą odległość euklidesową pomiędzy dwoma przekazanymi obiektami. Odległość dwóch punktów na płaszczyźnie dana jest wzorem $d = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$, funkcja biblioteki standardowej wyznaczająca pierwiastek to `sqr()`. Ustal nagłówek funkcji, by zminimalizować rozmiar kopiowanych do funkcji danych. Składniki oceny:

- poprawne wyznaczenie i zwrócenie odległości: 1 pkt,
- ograniczenie rozmiaru przekazywanych do funkcji danych: 1 pkt.

Zadanie 4. [5 pkt]

Niech zbiór obiektów wczytanych w Zadaniu 2 nazwany będzie zbiorem treningowym. Dany jest pewien obiekt Q , o znanych współrzędnych x i y , ale nieznaney etykiety: $Q = (x, y, ?)$.

Klasyfikacja metodą k najbliższych sąsiadów (k nearest neighbours, kNN) obiektu Q polega na zlokalizowaniu w zbiorze treningowym jego k najbliższych sąsiadów i przyporządkowaniu mu najliczniej pojawiającej się etykiety spośród owych sąsiadów.

Zdefiniuj funkcję `void knn(const Obiekt treningowy[], int n, Obiekt& Q, int k)`, która dokona klasyfikacji obiektu Q (uzupełni jego etykietę) wykorzystując jego k najbliższych sąsiadów. Założyć, że $k \ll n$ (k jest dużo mniejsze od n) oraz $1 \leq k \leq K$, gdzie K to pewna znana w programie stała. Jeżeli więcej niż jedna etykieta będzie najliczniej występującą, wybrać dowolną. Składowe oceny:

- wyznaczenie odległości obiektu Q od wszystkich elementów zbioru treningowego [1 pkt],
- znalezienie k najbliższych sąsiadów obiektu Q [2 pkt],
- zliczenie wystąpień etykiet najbliższych sąsiadów [1 pkt],
- znalezienie najliczniej pojawiającej się etykiety i ustawienie jej w obiekcie Q [1 pkt].

W razie potrzeby wykorzystaj funkcje zdefiniowane w poprzednich zadaniach. **Nie wolno** wykorzystywać funkcji biblioteki standardowej.