|  |
| --- |
| **TECHNICKÁ UNIVERZITA V KOŠICIACH**  **FAKULTA ELEKTROTECHNIKY A INFORMATIKY** |
| **k-means: Generovanie zhlukov**  **Systemova prirucka** |
|  |
|  |

|  |  |
| --- | --- |
| **2020 Košice** | **Bc. Vitaliy Replyuk** |

**Obsah**

[Uvod 3](#_Toc58238895)

[1. K – means 4](#_Toc58238896)

[2. Programovy modul 5](#_Toc58238897)

Uvod

Dokumentacia popisuje algoritm klasifikacie metodov k-means. Zadanie je vypracovane v jazyku C, v programovom prostredi Dev-C++.

1. K – means

Je to najobľúbenejšia klasifikačná metóda.

Činnosť algoritmu je taká, že sa snaží minimalizovať celkovú štvorcovú odchýlku bodov klastra od stredov týchto klastrov:

kde *k* je počet zhlukov, sú výsledné zhluky, *i* = 1, 2, *k*, a sú centrá hmotnosti všetkých vektorov *x* z klastra .

Popis algoritmu:

1. Vopred musíme vedieť, koľko klasterov potrebujeme na distribúciu údajov.

2. Musíme si zvoliť náhodné začiatočné centrá našich klasterov.

3. Vypočítajme vzdialenosť od každého stredu ku každému bodu. Táto vzdialenosť sa považuje za euklidovskú vzdialenosť medzi bodmi v priestore.

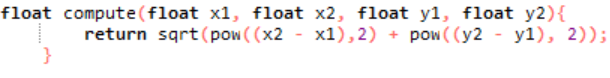
4. Po prepočítaní všetkých stredov sme body rozdelili do zhlukov a porovnali sme vzdialenosť od každého bodu zo stredom. Body je umiestnený v zhluku, do stredu ktorého je umiestnený bližšie ako do zvyšku centrov.

1. Programovy modul

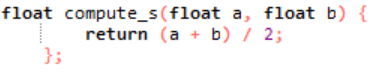
Obmedzujúce podmienky - pre fungovanie programu je potrebne zadať na začatku počet bodov.

Riešenie obsahuje dve funkcie:

сompute – spočíta euklidovskú vzdialenosť medzi bodmi v priestore.

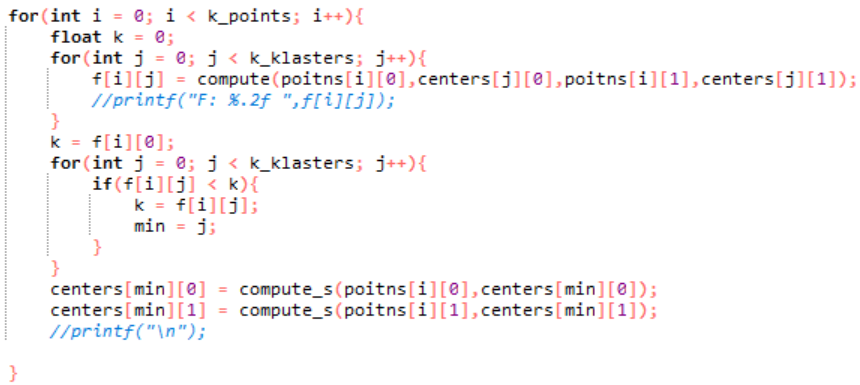


compute\_s – spočíta novy stred

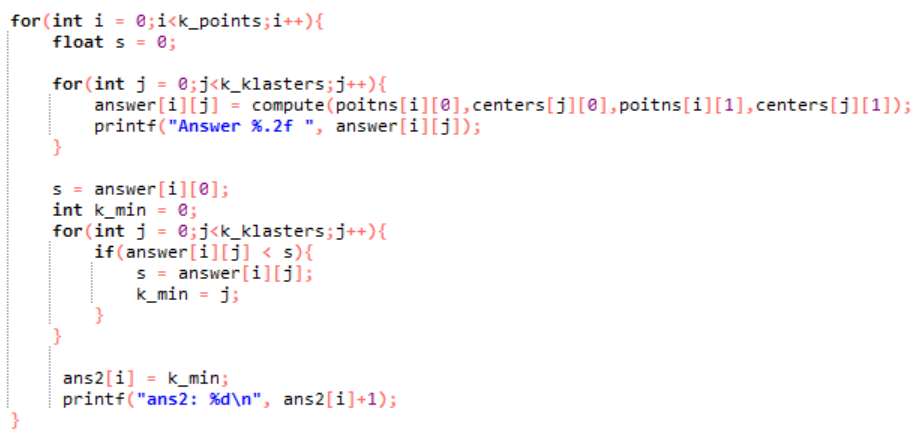


Telo základných výpočtov:

Dana čast kodu pre kazdy bod nachádza vzdialenosť od neho ku každému stredu. Definuje minimálnu vzdialenosť a fixuje číslo stredu pre ďalší prepočet.



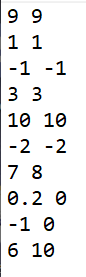
Dana čast kodu klasifikuje body podľa zoskupenía.



Program funguje, kým sa centrá novej iterácie sa líši od starej, alebo kým nie je prekročený limit v 30 (dasa zmenit) iterácií.

* 1. Ukážka funkcionality programu

Vstup:



Vystup:

