

Istraživanje podataka - vežbe 11, 2021.

1 Zadaci

1. U programskom jeziku Python izvršiti klasterovanje nad skupom *dogs* primenom algoritma K-sredina za 2, 3 i 4 klastera na osnovu visine i težine pasa. Rezultate klasterovanja prikazati pomoću grafika sa raštrkanim elementima.

Skup *dogs* ima attribute:

- *breed* - rasa psa
- *height* - visina psa
- *weight* - težina psa

2. Izvršiti klasterovanje predmeta primenom algoritma K-sredina u alatu IBM SPSS Modeler. Skup *podaci_o_predmetima.csv* ima attribute:

- *predmet* - naziv predmeta
- *upisalo* - broj studenata koji su upisali predmet
- *polozilo* - broj studenata koji su položili ispit iz predmeta
- *prosek* - prosečna ocena na položenim ispitima iz predmeta. Za predmete koje nijedan student nije položio, prosek je 5.

3. Izvršiti klasterovanje država na osnovu podataka o ishrani stanovništva primenom algoritma K-sredina u alatu IBM SPSS Modeler. Skup *ishrana.csv* sadrži podatak koliko su određene namirnice zastupljene u ishrani stanovništva nekih država. Vrednosti numeričkih atributa su izražene u procentima. Atributi skupa su:

- *Country* - država
- *RedMeat* - crveno meso
- *Eggs* - jaja
- *Milk* - mleko
- *Fish* - riba
- *Cereals* - žitarice
- *Starch* - skrob
- *Nuts* - koštunjavo voće
- *Fr&Veg* - voće i povrće