

Analiza algoritma uniclass

Jure Ham - 63080514

May 30, 2011

Kazalo

1	Opis programa	2
1.1	Branje podatkov	2
1.2	Računanje matrike razdalj	2
1.3	Simuliranje sil	2
2	Vizualizacija podatkov	2
3	Klasifikacija	2
4	Clustering	2
5	Zaključek	2

1 Opis programa

Program uniclass je celovit paket namenjen testiranju in nadgradji algoritma uniclass. Grafični vmesnik omogoča nalaganje testnih primerov, spreminjanje osnovnih nastavitev algoritma ter sproten prikaz delovanja. Ker je algoritem računsko zahteven, je program implementiran v javi, ki omogoča dober kompromis med hitrostjo izvajanja in hitrostjo razvoja.

1.1 Branje podatkov

Program podpira nalaganje testnih primerov v formatu tab, ki je osnoven format programskega paketa orange. Implementacija je omejena na zvezne in neurejene diskretne vrednosti ter na eno samo meta vrednost, ki postane ime entitete.

Manjkajoči podatki so obravnavani kot posebne vrednosti, kar pomeni, da je razdalja med entiteto z vsemi podatki in entiteto brez podatkov največja, razdalja med dvema entitetama brez podatkov pa je najmanjša. Razlog za to odločitev je zelo enostavna implementacija.

1.2 Računanje matrike razdalj

1.3 Simuliranje sil

2 Vizualizacija podatkov

3 Klasifikacija

4 Clustering

5 Zaključek