

Obsah

1	Úvod	15
2	Principy shlukové analýzy	23
2.1	<i>Shlukování a klasifikace</i>	23
2.2	<i>Prvky shlukování</i>	24
2.2.1	Vstupní datová matice	24
2.2.2	Shluky	25
2.2.3	Míry podobnosti.....	26
2.2.4	Kritéria shlukování	27
2.2.5	Kritéria pro stanovení počtu shluků.....	28
2.2.6	Interpretace výsledků.....	29
2.3	<i>Předzpracování datového souboru</i>	29
2.3.1	Výběr proměnných	29
2.3.2	Transformace dat	31
2.3.3	Chybějící údaje a odlehlé objekty.....	34
2.4	<i>Problematika velkých datových souborů</i>	34
2.4.1	Indexování objektů	36
2.5	<i>Snižování rozměru úlohy</i>	38
2.5.1	Analýza hlavních komponent	38
2.5.2	Stanovení hlavních komponent.....	42
2.5.3	Metoda SVD	47
2.6	<i>Základní symboly použité při dalším výkladu</i>	49
3	Měření podobnosti	53
3.1	<i>Podobnost objektů</i>	53
3.1.1	Kvantitativní data.....	53
3.1.2	Dichotomická data	57
3.1.3	Nominální data.....	57
3.1.4	Kombinace datových typů	58
3.2	<i>Podobnost proměnných</i>	58
3.2.1	Principy zkoumání závislosti na základě analýzy variability.....	59
3.2.2	Závislost nominálních proměnných.....	62
3.2.3	Závislost binárních proměnných.....	66
3.2.4	Závislost ordinálních a kvantitativních diskrétních proměnných.....	69
3.3	<i>Podobnost v matici binárních dat</i>	73
3.3.1	Míry závislosti	74
3.3.2	Ostatní míry	76
3.4	<i>Podobnost kategorií</i>	78
3.5	<i>Podobnost shluků</i>	80

4	Metody shlukové analýzy	83
4.1	<i>Klasifikace metod shlukové analýzy.....</i>	83
4.1.1	Tradiční metody a jejich modifikace.....	83
4.1.2	Novější přístupy	85
4.2	<i>Metody rozkladu.....</i>	85
4.2.1	Metoda k -průměrů	85
4.2.2	Metoda k -medoidů.....	87
4.2.3	Metoda k -modů a k -histogramů.....	89
4.2.4	Fuzzy shluková analýza	89
4.2.5	Minimální kostra grafu.....	93
4.2.6	Aplikace matematického programování	94
4.2.7	Grafická reprezentace shluků.....	95
4.3	<i>Hierarchická shluková analýza.....</i>	96
4.3.1	Monotetické shlukování	97
4.3.2	Polytetické shlukování	98
4.3.3	Dvourozměrné shlukování	106
4.4	<i>Metody pro shlukování velkých souborů.....</i>	108
4.4.1	Modifikace metod rozkladu	108
4.4.2	Modifikace hierarchických metod.....	108
4.4.3	Metody založené na hustotě, mřížce a modelu	114
4.4.4	Shlukování podprostorů	118
4.4.5	Postupy pro kategoriální data.....	119
4.5	<i>Výběr vhodné metody a stanovení počtu shluků</i>	122
4.5.1	Ohodnocení disjunktčního shlukování.....	125
4.5.2	Ohodnocení fuzzy shlukování.....	137
5	Biologicky inspirované algoritmy	145
5.1	<i>Umělé neuronové sítě.....</i>	145
5.1.1	Úvod	145
5.1.2	Biologická inspirace	147
5.1.3	Formální neuron	149
5.1.4	Umělá neuronová síť	151
5.1.5	Učení	153
5.1.6	Kohonenovy mapy	154
5.1.7	Sítě založené na samoorganizujících se mapách s organizační fází	160
5.1.8	Soft konkurenční učení s možností změny architektury, neuronový plyn....	164
5.1.9	Neuronový plyn s konkurenčním Hebbovským učením.....	167
5.1.10	Rostoucí neuronový plyn.....	169
5.1.11	Rostoucí buněčná struktura	172
5.1.12	Algoritmus samoorganizujícího se stromu.....	176
5.1.13	Teorie adaptivní rezonance.....	184
5.1.14	Shrnutí	192
5.2	<i>Genetické algoritmy</i>	192
5.2.1	Genetický algoritmus k -průměrů.....	198
5.2.2	Shrnutí	200

Seznam literatury 203

Rejstřík 217

KAPITOLA I

ÚVOD