

# Obsah

<b>1</b>	<b>Úvod</b>	<b>13</b>
<b>2</b>	<b>Principy shlukové analýzy</b>	<b>21</b>
2.1	<i>Shlukování a klasifikace</i>	21
2.2	<i>Prvky shlukování</i>	22
2.2.1	Vstupní datová matice	22
2.2.2	Shluky	23
2.2.3	Míry podobnosti	24
2.2.4	Kritéria shlukování	25
2.2.5	Kritéria pro stanovení počtu shluků	26
2.2.6	Interpretace výsledků	27
2.3	<i>Předzpracování datového souboru</i>	27
2.3.1	Výběr proměnných	27
2.3.2	Transformace dat	29
2.3.3	Chybějící údaje a odlehlé objekty	32
2.4	<i>Problematika velkých datových souborů</i>	33
2.4.1	Indexování objektů	34
2.5	<i>Snižování rozměru úlohy</i>	36
2.5.1	Analýza hlavních komponent	36
2.5.2	Stanovení hlavních komponent	40
2.5.3	Metoda SVD	45
2.6	<i>Základní symboly použité při dalším výkladu</i>	47
<b>3</b>	<b>Měření podobnosti</b>	<b>51</b>
3.1	<i>Podobnost objektů</i>	51
3.1.1	Kvantitativní data	51
3.1.2	Dichotomická data	55
3.1.3	Nominální data	55
3.1.4	Kombinace datových typů	56
3.2	<i>Podobnost proměnných</i>	56
3.2.1	Principy zkoumání závislosti na základě analýzy variability	57
3.2.2	Závislost nominálních proměnných	60
3.2.3	Závislost binárních proměnných	64
3.2.4	Závislost ordinálních a kvantitativních diskrétních proměnných	67
3.3	<i>Podobnost v matici binárních dat</i>	71
3.3.1	Míry závislosti	72
3.3.2	Ostatní míry	74
3.4	<i>Podobnost kategorií</i>	76

<b>4</b>	<b>Metody shlukové analýzy .....</b>	<b>81</b>
4.1	<i>Klasifikace metod shlukové analýzy.....</i>	<i>81</i>
4.1.1	Tradiční metody a jejich modifikace .....	81
4.1.2	Novější přístupy .....	83
4.2	<i>Metody rozkladu.....</i>	<i>83</i>
4.2.1	Metoda $k$ -průměrů .....	83
4.2.2	Metoda $k$ -medoidů.....	85
4.2.3	Metoda $k$ -modů a $k$ -histogramů.....	87
4.2.4	Fuzzy shluková analýza .....	87
4.2.5	Minimální kostra grafu.....	91
4.2.6	Aplikace matematického programování.....	92
4.2.7	Grafická reprezentace shluků .....	93
4.3	<i>Hierarchická shluková analýza.....</i>	<i>94</i>
4.3.1	Monotetické shlukování .....	95
4.3.2	Polytetické shlukování .....	96
4.3.3	Dvourozměrné shlukování .....	104
4.4	<i>Metody pro shlukování velkých souborů.....</i>	<i>106</i>
4.4.1	Modifikace metod rozkladu.....	106
4.4.2	Modifikace hierarchických metod.....	106
4.4.3	Metody založené na hustotě, mřížce a modelu.....	112
4.4.4	Shlukování podprostorů .....	116
4.4.5	Postupy pro kategoriální data .....	117
4.5	<i>Stanovení optimálního počtu shluků .....</i>	<i>119</i>
<b>5</b>	<b>Biologicky inspirované algoritmy .....</b>	<b>125</b>
5.1	<i>Umělé neuronové sítě.....</i>	<i>125</i>
5.1.1	Úvod.....	125
5.1.2	Biologická inspirace .....	127
5.1.3	Formální neuron .....	129
5.1.4	Umělá neuronová síť .....	131
5.1.5	Učení .....	133
5.1.6	Kohonenovy mapy .....	134
5.1.7	Sítě založené na samoorganizujících se mapách s organizační fází .....	140
5.1.8	Soft konkurenční učení s možností změny architektury, neuronový plyn....	144
5.1.9	Neuronový plyn s konkurenčním Hebbovským učením .....	147
5.1.10	Rostoucí neuronový plyn.....	149
5.1.11	Rostoucí buněčná struktura .....	152
5.1.12	Algoritmus samoorganizujícího se stromu .....	156
5.1.13	Teorie adaptivní rezonance.....	163
5.1.14	Shrnutí .....	171
5.2	<i>Genetické algoritmy .....</i>	<i>172</i>
5.2.1	Genetický algoritmus $k$ -průměrů.....	177
5.2.2	Shrnutí.....	180



Rejstřík .....	183
Seznam literatury .....	185

## KAPITOLA I

### ÚVOD