

Obsah

1	Statistické metody v systémech pro dobývání znalostí z dat	11
1.1	Typy softwarových systémů používaných pro dobývání znalostí z dat	11
1.1.1	Univerzální systémy pro dobývání znalostí.	12
1.1.2	Systémy pro podporu okamžitého rozhodování na základě dat.	13
1.1.3	Systémy s jiným primárním určením.	14
1.2	Clementine	15
1.2.1	Popisné statistiky	18
1.2.2	Regresní metody	19
1.2.3	Mnohorozměrná statistická analýza	20
1.3	DecisionSite	22
1.4	Matlab	26
1.4.1	Popisné statistiky	30
1.4.2	Lineární regrese a její zobecnění	30
1.4.3	Mnohorozměrná statistická analýza	35
1.4.4	Testování hypotéz	40
2	Statistické principy procedury 4ft-Miner	45
2.1	Metoda GUHA a získávání tvrzení observační logiky	45
2.1.1	Observační logika	45
2.2	Observační kvantifikátory založené na odhadech	48
2.2.1	Implikační kvantifikátory	49
2.2.2	Další kvantifikátory založené na odhadech	50
2.3	Observační kvantifikátory založené na testování hypotéz	51
2.3.1	Kvantifikátory založené na binomickém testu	52
2.3.2	Kvantifikátory založené na testech nezávislosti	53
2.3.3	Kvantifikátory neodpovídající zamítání nulové hypotézy	55

3	Moderní klasifikační a regresní metody v dobývání znalostí	58
3.1	Společné principy klasifikace a regrese	58
3.1.1	Hlavní rysy klasifikace	58
3.1.2	Hlavní rysy regrese	59
3.1.3	Regularizační operátory	61
3.1.4	Jednotlicí rámec klasifikace a regrese	62
3.2	Klasifikace pomocí perceptronů	64
3.2.1	Perceptron a jeho učení	64
3.2.2	Věta o konvergenci učicího algoritmu perceptronu	67
3.2.3	Modifikace perceptronu	68
3.3	SVM-klasifikátory pro lineárně separabilní problémy	69
3.3.1	Klasifikátory spočívající na úloze kvadratické optimalizace	71
3.4	Konstrukce klasifikátorů pro lineárně neseparabilní problémy	73
3.4.1	Jádrové funkce	74
3.4.2	Klasifikátory spočívající na jádrových funkcích	76
3.5	Využití principu SVM-klasifikátorů v regresi	78
3.5.1	SV-regrese	78
3.5.2	Řešení optimalizační úlohy pro SV-regresi	79
3.6	Nelineární regrese pomocí vícevrstvých perceptronů	83
3.6.1	Princip vícevrstvého perceptronu	83
3.6.2	Učení vícevrstvého perceptronu	86
3.6.3	Důvod používání vícevrstvých perceptronů k regresi	88
3.6.4	Prostory funkcí aproximovaných vícevrstvámi perceptrony	89
3.6.5	Věta o univerzální aproximační schopnosti vícevrstvých perceptronů	92

4	Statistický přístup k umělým neuronovým sítím	93
4.1	Základní principy statistického přístupu k umělým neuronovým sítím	93
4.1.1	Učení založené na náhodném výběru a na střední hodnotě	94
4.2	Použití silného zákona velkých čísel při učení neuronových sítí	96
4.2.1	Silný zákon velkých čísel pro učení vícevrstvých perceptronů	98
4.3	Použití centrální limitní věty při prořezávání sítě	99
4.3.1	Centrální limitní věta pro vícevrstvé perceptrony	100
4.3.2	Testovací procedura pro vícevrstvé perceptrony, založená na centrální limitní větě	102