

Inhaltsverzeichnis

1	Data-Mining-Prozess	1
2	Daten und Relationen	5
2.1	Beispiel	5
2.2	Maßskalen	7
2.3	Matrixdarstellung	9
2.4	Relationen	10
2.5	Unähnlichkeitsmaße	10
2.6	Ähnlichkeitsmaße	12
2.7	Sequenz- und Textrelationen	14
2.8	Abtastung und Quantisierung	17
3	Datenvorverarbeitung	21
3.1	Fehlerarten	21
3.2	Filterung	26
3.3	Standardisierung	31
3.4	Datenkonsolidierung	34
4	Visualisierung	35
4.1	Diagramme	35
4.2	Hauptachsentransformation	37
4.3	Mehrdimensionale Skalierung	41
4.4	Histogramme	47
4.5	Spektralanalyse	50
5	Korrelation	55
5.1	Lineare Korrelation	55
5.2	Chi-Quadrat-Unabhängigkeitstest	61
6	Regression	65
6.1	Lineare Regression	65
6.2	Neuronale Netze	69
6.3	Radiale Basisfunktionen	75
6.4	Training und Validierung	76
6.5	Merkmalsselektion	79

7	Zeitreihenprognose	81
8	Klassifikation	85
8.1	Naiver Bayes-Klassifikator	89
8.2	Lineare Diskriminanzanalyse	91
8.3	Support-Vektor-Maschine	93
8.4	Nächster Nachbar Klassifikator	96
8.5	Lernende Vektorquantisierung	96
8.6	Entscheidungsbäume	99
9	Clustering	105
9.1	Sequentielles Clustering	106
9.2	Prototypbasiertes Clustering	109
9.3	Fuzzy Clustering	111
9.4	Relationales Clustering	117
9.5	Clustervalidität und -tendenz	122
9.6	Hierarchisches Clustering	124
9.7	Selbstorganisierende Karte	126
9.8	Regelerzeugung	128
10	Zusammenfassung	135
	Übungsaufgaben	141
	Symbolverzeichnis	145
	Literaturverzeichnis	147
	Sachwortverzeichnis	157