

Inhaltsverzeichnis

1	Einführung	1
2	Unsupervised Learning	7
2.1	Klassische Clusteranalysen	7
2.1.1	Ähnlichkeitsmaße und Distanzbegriffe bei Objekten . . .	10
2.1.2	Ähnlichkeitsmaße und Distanzbegriffe bei Objektmen- gen	26
2.1.3	Homogenität bzw. Inhomogenität einer Objektmenge . .	31
2.1.4	Partitionen	36
2.1.5	Optimalitätskriterien	39
2.1.6	Varianzkriterium	46
2.1.7	Determinantenkriterium	50
2.1.8	Spurkriterium	53
2.1.9	Minimierung gegebener Kriterien	53
2.1.10	Qualitative Merkmale	62
2.1.11	Wahl der Anfangspartitionen	62
2.1.12	Modifikationen der obigen Optimierungsverfahren	64
2.2	Rekursive Verfahren	67

2.2.1	Stochastische Approximationsverfahren	67
2.2.2	Motivation	70
2.2.3	Das rekursive Verfahren	73
2.3	Sequentielle Verfahren	81
2.4	Medoidverfahren	83
2.5	Nearest Neighbor Clustering	84
2.6	Simulated Annealing und Clusterung	86
2.7	Mischungsverteilungen	90
2.8	Nichtparametrische Verfahren	100
2.8.1	Dichteschätzung	100
2.8.2	Konsistenz der Dichteschätzer	107
2.8.3	Ein zentraler Grenzwertsatz	109
2.8.4	Optimale Kerne	110
2.8.5	Beste Konvergenzrate	111
2.8.6	Multivariate Dichteschätzung	112
2.8.7	Anwendung der Dichteschätzung	116
2.9	Wahl der Anzahl m der Klassen	119
2.10	Validierung von Klassen	123
2.11	Anwendungsbeispiele	123
2.11.1	Ein physikalisch-chemisches Problem	123
2.11.2	Ein betriebswirtschaftliches Problem	125
2.12	Hierarchische Algorithmen	127
2.12.1	Partitionenhierarchien	127
2.12.2	Mächtigkeit von Partitionenhierarchien	131
2.12.3	Agglomerative Verfahren	132

2.12.4	Divisive Verfahren	137
2.12.5	Tarifierung in der Versicherungsbranche	140
2.12.6	Hierarchische Verfahren: Die Praxis	143
2.13	Clusteranalysen und neuronale Netze	144
2.13.1	Self-Organizing Maps	144
2.13.2	Learning Vector Quantization	148
2.14	Erfahrungen aus der Praxis	150
3	Supervised Learning	155
3.1	Bayes'sche Entscheidungstheorie	160
3.2	Diskriminanzanalyse	165
3.2.1	Parametrische Diskriminanzanalysen	166
3.2.2	Die multivariate Normalverteilung	168
3.2.3	Multivariate t-Verteilungen	174
3.2.4	Anmerkungen aus der Praxis	176
3.2.5	Variablenselektion: quantitative Daten	176
3.2.6	Ein Beispiel aus der Praxis	184
3.2.7	Nichtparametrische Diskriminanzanalysen	187
3.2.8	Variablenselektion	193
3.2.9	Nichtparametrisch versus parametrisch	193
3.3	Verallgemeinerte lineare Modelle	194
3.3.1	Logistische Regressionen	201
3.3.2	Eine Anwendung in der Praxis	230
3.4	Neuronale Netze	232
3.4.1	Grundlagen der neuronalen Netze	232

3.4.2	Additive Modelle	244
3.4.3	Projection Pursuit Regression	245
3.4.4	Radial Basis Functions	246
3.4.5	MARS	247
3.4.6	Feed-forward neuronale Netze	248
3.4.7	Einige Anmerkungen	248
3.5	Entscheidungsbäume	251
3.5.1	Die Menge $\mathcal{E}$	259
3.5.2	Die Bewertungsfunktion	261
3.5.3	Zuordnungsregeln für Klassen	265
3.5.4	Pruning	269
3.6	Regressionsbäume	291
3.7	Bayes-Risiko Konsistenz	296
3.8	Entscheidungsbäume: Ausblick	301
3.8.1	Maße für Splits	302
3.8.2	Multivariate Splits	303
3.8.3	Die richtige Größe	303
3.8.4	Wichtige Variable	304
3.9	Erfahrungen aus der Praxis	304
3.10	Abgrenzung der Klassifikationsverfahren	305
4	Literatur	309
4.1	Überblick	309
4.1.1	Clusteranalysen	309
4.1.2	Parametrische Diskriminanzanalysen	310

- 4.1.3 Nichtparametrische Diskriminanzanalysen 308
- 4.1.4 Neuronale Netze: Grundlagen 308
- 4.1.5 Verallgemeinerte lineare Modelle 308
- 4.1.6 Neuronale Netze und Pattern recognition 309
- 4.1.7 Selforganizing Maps, rekursive Verfahren 309
- 4.1.8 Entscheidungsbäume 310
- 4.1.9 Versicherungsmathematik 310
- 4.2 Weitergehende Literatur 310