

# Data Mining

in der Medizin und Medizintechnik

von  
Ralf Mikut



Universitätsverlag karlsruhe

# Inhaltsverzeichnis

<b>Kurzfassung</b>	<b>i</b>
<b>Danksagung</b>	<b>Mi</b>
<b>1 Einführung</b>	<b>1</b>
1.1 Bedeutung und Einordnung . . . . .	.1
1.2 Entwicklungsstand . . . . .	2
1.3 Ziele und Aufgaben. . . . .	5
<b>2 Medizinische Problemstellungen und Rahmenbedingungen</b>	<b>9</b>
2.1 Datenquellen und -archivierung. . . . .	9
2.1.1 Datenquellen. . . . .	9
2.1.2 Standardisierung und Archivierung. . . . .	10
2.2 Problemstellungen. . . . .	12
2.2.1 Übersicht . . . . .	12
2.2.2 Klinische Studien. . . . .	13
2.2.3 Modellbildung, Simulation, Regelung und Steuerung biologischer Systeme. . . . .	19
2.3 Einsatzpotenziale und Anforderungen für Data-Mining-Verfahren . . . . .	24
<b>3 Data Mining: Verfahrensübergreifende Begriffe</b>	<b>27</b>
3.1 Übersicht . . . . .	27
3.2 Daten und Datenstrukturen. . . . .	28
3.3 Problemstellungen. . . . .	31
3.3.1 Systembegriff. . . . .	31
3.3.2 Prognoseorientierte Probleme. . . . .	34
3.3.3 IVlerkmalstransformationen als Hilfsverfahren. . . . .	41
3.3.4 Rollen von Bewertungsmaßen. . . . .	45
3.4 Informationsdefizite und Ungewissheit . . . . .	45
3.5 Bewertungsmaße . . . . .	48
3.5.1 Wahrscheinlichkeiten und Häufigkeiten. . . . .	48
3.5.2 Distanzmaße und Separationsmaße. . . . .	50

# Inhaltsverzeichnis

3.5.3	Informationstheoretische Maße. . . . .	52
3.5.4	Klassifikationsorientierte Maße. . . . .	54
3.5.5	Regressionsorientierte Maße. . . . .	58
3.5.6	Merkmalskosten, Interpretierbarkeits- und Komplexitätsmaße . .	60
3.6	Optimierungsprobleme und-verfahren. . . . .	65
3.7	Validierungstechniken. . . . .	67
<b>4</b>	<b>Einsatzszenario für Data-Mining-Verfahren in der Medizin</b>	<b>73</b>
4.1	Ablauf. . . . .	73
4.2	Problemformulierungen. . . . .	76
4.3	Lerndatensätze und Datenvorverarbeitung. . . . .	84
4.4	Merkmalsextraktion, -Selektion und -transformation. . . . .	88
4.5	Klassifikation und Regression. . . . .	95
4.6	Visualisierung. . . . .	97
4.7	Implementierungsaspekte für Medizingeräte. . . . .	100
<b>5</b>	<b>Data Mining: Spezielle Verfahren</b>	<b>103</b>
5.1	Übersicht . . . . .	103
5.2	Statistische Verfahren. . . . .	105
5.2.1	Verteilungsannahmen. . . . .	105
5.2.2	Statistische Signifikanztests. . . . .	107
5.2.3	Hauptkomponenten- und Diskriminanzanalyse. . . . .	111
5.2.4	Distanz-und Separationsmaße. . . . .	123
5.2.5	Bayes-Klassifikatoren und deren Vereinfachungen. . . . .	124
5.2.6	Nearest-Neighbor-Verfahren. . . . .	131
5.3	Support-Vektor-Maschinen. . . . .	135
5.4	Entscheidungsbäume. . . . .	144
5.5	Fuzzy-Systeme. . . . .	153
5.5.1	Prinzip. . . . .	153
5.5.2	Zugehörigkeitsfunktionen und Fuzzifizierung. . . . .	154
5.5.3	Fuzzy-Inferenz . . . . .	158
5.5.4	Defuzzifizierung. . . . .	165
5.5.5	Entwurf von Fuzzy-Systemen. . . . .	167
5.5.6	Automatisierte Generierung von Erklärungstexten. . . . .	186
5.5.7	Implementierungsaspekte. . . . .	189
5.5.8	Anwendungsfelder in der Medizin. . . . .	191
5.6	Künstliche Neuronale Netze. . . . .	193
5.6.1	Aufbau und Einteilung. . . . .	193

5.6.2	Entwurfsverfahren. . . . .	197
5.6.3	Anwendungsfelder in der Medizin. . . . .	205
5.7	Clusterverfahren. . . . .	206
5.8	Regressionsverfahren. . . . .	215
5.8.1	Übersicht . . . . .	215
5.8.2	Schätzung von Überlebenszeiten. . . . .	218
5.8.3	Logistische Regression. . . . .	221
5.8.4	Regressionsansätze für dynamische Modelle. . . . .	222
5.9	Wertung. . . . .	225
<b>6</b>	<b>Data-Mining-Softwarelösungen in der Medizin</b>	<b>231</b>
6.1	Ausgangssituation. . . . .	231
6.2	Gait-CAD als Plattform für interaktive Analysen. . . . .	232
6.2.1	Motivation. . . . .	232
6.2.2	Leistungsumfang und Bedienkonzept. . . . .	233
6.2.3	Repräsentation der Problemstellungen. . . . .	235
6.2.4	Versuchsautomatisierung. . . . .	236
6.2.5	Erweiterbare Merkmalsextraktion mit Plugins. . . . .	236
6.3	Diskussion. . . . .	237
<b>7</b>	<b>Anwendungen</b>	<b>239</b>
7.1	Übersicht . . . . .	239
7.2	Entscheidungsunterstützung bei Bewegungsanalysen. . . . .	240
7.2.1	Aufgabenstellung . . . . .	240
7.2.2	Methoden und Ergebnisse. . . . .	243
7.2.3	Diskussion. . . . .	259
7.3	Individuelle Anpassung von myoelektrischen Unterarmprothesen . . . . .	260
7.3.1	Aufgabenstellung. . . . .	260
7.3.2	Methoden und Ergebnisse. . . . .	261
7.3.3	Diskussion. . . . .	270
7.4	Brain Machine Interfaces. . . . .	271
7.4.1	Aufgabenstellung. . . . .	271
7.4.2	Methoden und Ergebnisse. . . . .	274
7.4.3	Diskussion. . . . .	284
7.5	Bildgestützte Diagnose bei Mammakarzinomen. . . . .	286
7.5.1	Aufgabenstellung. . . . .	286
7.5.2	Methoden und Ergebnisse. . . . .	288
7.5.3	Diskussion. . . . .	292

## Inhaltsverzeichnis

8	Vorgehensweise bei medizinischen Datenanalysen	293
9	Zusammenfassung und Ausblick	301
A	Wichtige Symbole und Bezeichnungen	305
B	Abbildungsverzeichnis	319
C	Tabellenverzeichnis	321
D	Literaturverzeichnis	323
E	Index	349