

Bertram Küppers

Data Mining in der Praxis

**Ein Ansatz zur Nutzung
der Potentiale von Data Mining
im betrieblichen Umfeld**



PETER LANG
Europäischer Verlag der Wissenschaften

Inhaltsverzeichnis

1 Einleitung	13
1.1 Problem.....	13
1.2 Ziel der Arbeit.....	14
1.3 Forschungsmethodik.....	14
1.4 Aufbau der Arbeit.....	15
2 Stand der Data-Mining-Forschung - Begriffsabgrenzung	17
2.1 Informationsbegriff.....	17
2.2 Ursprüngliche Data-Mining-Definition.....	19
2.2.1 Ansätze in der Literatur.....	19
2.2.2 Modell.....	22
2.3 Modifizierte Data-Mining-Definition.....	23
2.3.1 Vergleich zu bisherigen Definitionsansätzen.....	23
2.3.1.1 Knowledge Discovery vs. Data Mining.....	23
2.3.1.2 Data Mining - Definition.....	24
2.3.2 Modifiziertes und erweitertes Modell.....	25
2.4 Bewertung.....	28
3 Bedeutung von Data Mining im betrieblichen Umfeld	29
3.1 Individueller Informationsbedarf.....	30
3.1.1 Ansätze zur Ermittlung des Informationsbedarfs.....	30
3.1.2 Unterstützungspotential durch Data Mining.....	34
3.2 Informationssysteme zur Unterstützung des Managements.....	35
3.2.1 Ursprünge der Management-Informationssysteme.....	35
3.2.2 Weiterentwicklungen der Management-Informationssysteme.....	37
3.2.3 Aktuelle Konzepte zur Informationsversorgung.....	40
3.2.3.1 Integrativer Ansatz.....	40
3.2.3.2 Data-Warehouse-Ansatz.....	41
3.2.4 Ausblick für die Entwicklung von Informationssystemen.....	43
3.2.5 Einordnung des Data-Mining-Ansatzes.....	44
3.3 Informationsmanagement als Bestandteil der Unternehmensführung.....	44
3.3.1 Wesen des Informationsmanagements.....	45
3.3.2 Der HPO-Ansatz.....	46
3.3.3 Bedeutung von Data Mining für das Informationsmanagement.....	48

3.4 Bewertung	49
4 Data-Mining-Methoden.....	51
4.1 Abgrenzung zu verwandten Gebieten	51
4.2 Neuronale Netze.....	53
4.3 Maschinelles Lernen.....	55
4.3.1 Entscheidungsbäume	56
4.3.2 Entscheidungsregeln	57
4.3.3 Konzeptionelles Clustern	57
4.3.4 Case-Based Reasoning (Fallbasiertes Schließen)	59
4.4 Heuristische Ansätze.....	61
4.4.1 Automatische Navigation	61
4.4.2 CLUSMIN.....	64
4.4.3 Assoziativ-Regeln	65
4.4.4 Gruppenvergleich.....	66
4.5 Sonstige Ansätze	67
4.5.1 Rough Sets	67
4.5.2 Visualisierungsansätze.....	69
4.6 Unterstützende Ansätze.....	70
4.6.1 Clusteranalyse	70
4.6.2 Fuzzy-Set-Theorie.....	71
4.6.3 Bayes-Ansatz	73
4.6.4 Optimierung.....	75
4.7 Einordnungsansätze in der Literatur.....	77
4.7.1 Analogie zum Maschinellen Lernen	78
4.7.2 Ansatz von Bissanz.....	78
4.7.3 Ansatz von Fayyad u. a.	79
4.7.4 Bewertung der Einordnungsansätze.....	80
5 Ansatz einer problemorientierten Systematisierung von Data-Mining-Methoden.....	81
5.1 Ergebnisorientierte Einordnung von Data-Mining-Methoden	82
5.1.1 Prinzip der ergebnisorientierten Einordnung.....	82
5.1.2 Anwendung der ergebnisorientierten Einordnung auf die vorgestellten Methoden.....	84
5.2 Auswahlkriterien für die Bewertung von Data-Mining-Methoden.....	87
5.2.1 Anwenderorientierte Kriterien.....	88

5.2.1.1	Interessantheit	88
5.2.1.2	Interpretierfähigkeit/Verständlichkeit	90
5.2.1.3	Autonomiegrad	92
5.2.2	Methodenorientierte Kriterien	93
5.2.2.1	Charakterisierung von Unsicherheit	93
5.2.2.2	Explizite und implizite Annahmen	93
5.2.2.3	Regularisierung (Über-/Unteranpassung)	95
5.2.3	Datenorientierte Kriterien	96
5.2.3.1	Datendeformation	96
5.2.3.2	Datenqualität	97
5.2.3.3	Verarbeitbare Datenmengen/Laufzeitverhalten	98
5.3	Kombination verschiedener Methoden	99
6	Ansatz zur Implementierung von Data-Mining-Methoden im betrieblichen Umfeld	103
6.1	Analyse der beabsichtigten Anwendung	104
6.1.1	Anwendungszweck	105
6.1.1.1	Anwendungsbereich	105
6.1.1.2	Art der Ergebnisse	107
6.1.1.3	Häufigkeit der Analyse	107
6.1.2	Einbindung in das betriebliche Umfeld	109
6.1.3	Rechtliche Rahmenbedingungen	110
6.2	Technische Voraussetzungen	112
6.2.1	Infrastruktur	113
6.2.2	Datenbasis	113
6.3	Wirtschaftlichkeitsbetrachtung	117
6.4	Methodenauswahl und -prüfung	121
7	Einsatzgebiete und Potentiale von Data-Mining-Methoden	123
7.1	Überblick	123
7.2	Funktionale Anwendungen	125
7.2.1	Marketing	125
7.2.1.1	Segmentierung	126
7.2.1.2	Preisfindung	128
7.2.1.3	Database Marketing	132
7.2.1.4	Warenkorbanalyse	134
7.2.2	Controlling	134

7.2.3 Personal.....	136
7.2.4 Einkauf.....	137
7.2.5 Produktion.....	139
7.2.6 Service	140
7.3 Branchenspezifische Anwendungen.....	142
7.3.1 Finanzdienstleistungen.....	142
7.3.2 Handel/Konsumgüterindustrie	145
7.3.3 Gesundheitswesen.....	146
7.4 Bewertung	148
8 Zusammenfassung und Ausblick.....	151
9 Literaturverzeichnis	153
10 Anhang: Data-Mining-Anbieter	171