

Inhaltsverzeichnis

1	Vorwort	11
1.1	Überblick zum Buch	15
1.2	Welcher Leserkreis wird angesprochen?	18
1.3	Zu den Autoren	19
2	Grundlagen	23
2.1	Was ist Data Mining?	23
2.2	Prozeßmodell	30
2.3	Data Mining als Geschäftsprozeß	32
2.3.1	Rollenverteilung im Data Mining-Prozeß	33
2.3.2	Organisatorische Einbettung	35
2.3.3	Möglicher Einsatz in einem Rechenzentrum	36
2.3.4	Datenschutzrechtliche Aspekte	38
2.4	Voraussetzungen und Grenzen	40
2.4.1	Methodisch begründete Voraussetzungen	40
2.4.2	Datentechnisch begründete Voraussetzungen	41
2.4.3	Spezifische Problemstellungen der Datenanalyse	44
3	Architekturüberlegungen zu Data Mining und Data Warehouse	49
3.1	Data Warehouse-Architekturen	51
3.2	Einbindung von Data Warehouse-Werkzeugen	52
3.3	Architekturkriterien für Data Mining-Werkzeuge	55
4	Methoden des Data Mining	59
4.1	Induktives maschinelles Lernen	59
4.2	Überwachtes Lernen	61
4.2.1	Neuronale Netze	64
4.2.2	Entscheidungsbäume und Regelinduktion	69
4.2.3	k-nächste-Nachbarn-Technik und Diskriminanzanalyse	74
4.3	Unüberwachtes Lernen	78
4.3.1	Verfahren zur Aufdeckung von Assoziationen	81
4.3.2	Neuronale Netze zur Clusteranalyse	82

4.3.3	Demographisches Clustern (IBM-Verfahren)	84
4.3.4	k-means Clustering und hierarchische Clusterung	86
4.4	Statistik und andere Methoden	89
4.5	Gegenüberstellung	94
5	Anwendung von Data Mining im Bankwesen	99
5.1	Database Marketing I: Suche nach typischen Kunden im Retail Banking	101
5.2	Database Marketing II: Werden geringere Zinsen zu höheren Geschäftserträgen führen?	104
5.3	Betrugserkennung: Möglichkeiten der Aufdeckung von Kreditkartenmißbrauch	106
5.4	Risikoanalyse bei Kreditbewertung: Bonitätsanalyse	107
5.5	Financial Forecasting: Devisenkursprognose mit neuronalen Netzen	110
6	Aspekte bei der Wahl von Projektpartnern	115
6.1	Anbieter	116
6.1.1	Kriterien	116
6.1.2	IBM	118
6.1.3	SAS	118
6.2	Software	118
6.2.1	Kriterien	119
6.2.2	SAS	120
6.2.3	Klassische Statistikwerkzeuge und INSIGHT	122
6.2.4	Automatische Regression (ARA)	124
6.2.5	Entscheidungsbäume (Treedisc)	127
6.2.6	Neuronale Netze (NNA)	129
6.2.7	IBM	131
6.2.8	Klassifikation (Intelligent Miner)	133
6.2.9	Clusterung (Intelligent Miner)	137
6.2.10	Predicting Values (Intelligent Miner)	138
6.2.11	Assoziationen, sequentielle Muster, ähnliche Zeitreihen (Intelligent Miner)	140
7	Aspekte der Projektdurchführung	145
7.1	Untersuchungsgegenstand	146
7.2	Angebot und Datenbereitstellung	149
7.3	Aufbau des Analysebestandes und Preprocessing	151

Inhaltsverzeichnis

7.4	Allgemeine Projektergebnisse	153
7.5	Ausgewählte Ergebnisse	163
8	Literaturverzeichnis	167
9	Ressourcen im Internet	171
	Stichwortverzeichnis	173