

目次

まえがき	i
この本の構成	iii
謝辞	v
訳者 まえがき	vii
第1章 はじめに	1
拡大し続けるデータ	1
生産性の係数としての情報	2
学ぶことができるコンピュータシステム	3
データマイニング	6
KDD とデータマイニング	6
データマイニング 対 問合せツール	8
市場におけるデータマイニング	10
データマイニングの実際的な応用	12
まとめ	13
第2章 学習とは何か	15
はじめに	15
学習とは何か	16

自己学習コンピュータシステム	17
機械学習と科学の方法論	19
概念の学習	23
分類の正確さ	24
透過性	25
統計学的有意	25
情報の内容	26
探索空間の複雑さ	26
霧の中のカンガルー	27
まとめ	31
第3章 データマイニングとデータウェアハウス	33
はじめに	33
データウェアハウスとは何であって、なぜそれを必要とするのか	33
メタデータ	37
意思決定支援システムの設計	37
一般ユーザからの要求	38
意思決定支援システムのハードウェア製品とソフトウェア製品	38
データマイニングにおける統合	39
クライアント／サーバとデータウェアハウス	41
マルチプロセッシングコンピュータ	43
コストの正当化	45
まとめ	49
第4章 知識の発見段階	51
はじめに	51
知識発見の手順の詳細	53
データの選択	54
データの洗浄	55
補強	57
コード化	58
データマイニング	63
伝統的な問合せツールを使ったデータの予備的分析	64
視覚化技術	70

もってもらしきと距離	72
OLAP ツール	74
k -NN	75
決定木	78
連想規則	84
ニューラルネットワーク	90
遺伝的アルゴリズム	96
報告	103
まとめ	103
第5章 KDD の環境を設定する	105
はじめに	105
知識の異なった形	105
とりかかり	108
データの選択	110
データの洗浄	111
データの補強	112
データのコード化	114
データマイニング	114
報告	118
KDD 環境	119
10 の黄金則	120
1 非常に大きなデータを操作すること	120
2 混成学習を可能とすること	120
3 データウェアハウスを構築すること	121
4 データを洗浄する能力を導入すること	121
5 コード化を動的に行ないやすくすること	121
6 意思決定システムと統合すること	122
7 拡張可能な構造を選ぶこと	122
8 非均質データベースを使えること	122
9 クライアント／サーバーアーキテクチャを導入すること	122
10 キャッシュの最適化を導入すること	123
まとめ	123

第6章 実生活での応用例	125
はじめに	125
顧客調査	126
パイロットの乗務入札の予測	134
外部キーという関係の発見	139
結果	143
まとめ	145
第7章 学習アルゴリズムにおける形式的側面	147
はじめに	147
データ集合の圧縮としての学習	148
メッセージの中の情報内容	153
ノイズと冗長性	156
ノイズの重大性	157
ファジィデータベース	158
関係データベースの伝統的な理論	158
関係から表へ	161
キーから統計的な依存関係へ	161
非正規化	164
データマイニングの基礎技術	166
まとめ	168
この本のまとめ	169
用語解説	173
読み進めるべき文献	187
索引	191

