

Contents

CHAPTER 1	MINE2.COM	1
	The Idle Portal	1
	Mining Data	4
	Nothing New	7
	Mining Online	8
	How Data Mining Answers Key Business Questions	10
	Classification/Prediction	10
	Segmentation	11
	Association	13
	Clustering	15
	Visualization	15
	Optimization	17
	Data Mining Drives Relational Marketing	19
	Web Privacy	20
	Benefits and Incentives	21
	Ten Steps to Mining Your Web Data	22
	A Bookseller Example	25
	Look at the Data	26
	Split the Data	28
	Model the Data	30
	Integrate the Solution	32
	Compiling an Identity	33
	The Architecture for Analysis	35
	Personalized Future	36
	Why Bother?	39

CHAPTER 2	WHAT IS DATA MINING?	41
	It's about Predators and Patterns	41
	A Definition	42
	Drowning in Data	43
	The Complexity of Loyalty	45
	Traditional Data Mining Applications	46
	New Data Mining Applications	48
	Statistics	48
	Data Mining vs. Statistics	50
	Online Analytical Processing	51
	Data Mining vs. OLAP	52
	Website Analysis Tools	54
	Website Analysis Tools vs. Data Mining Tools	61
	Visual Data Mining	65
	Visualization Packages	65
	Link Analysis Systems	66
	Quantitative Display	67
	Data Mining Solutions	68
	Why and Who, Not What	71
	The Online Corner Store	72
	Selling as a Service	73
	Web Marketing	74
	Online Coupons	76
	Mining Customer Value	77
CHAPTER 3	NETWORKS AND ALGORITHMS	83
	They're about Learning and Evolving	83
	The AI War	84
	The Top-Down Approach	84
	The Bottom-Up Approach	85
	Neural Networks	86

From Slugs to VISA	87
Back-Propagation Networks	88
Supervised Learning	90
Training a Network	91
Preparing Your Data	92
Webbing Your Features	93
Purchasing Propensity	94
Training Cycle	95
Self-Organizing Maps (SOM)	95
Calibrating Clusters	96
Neural Network Tools	97
Machine-Learning Algorithms	98
A Measure of Information	99
The Rules	100
The Trees	102
The Trees vs. the Rules	103
Automatic Segmentation	104
Machine Learning Your Site	105
Machine-Learning Tools	106
Genetic Algorithms	107
Code of Survival	107
Reproduction	108
Crossover	109
Mutation	110
A Step-by-Step Example Using a Binary-Coded GA	110
Data Mining with Genetic Algorithms	112
Genetic Algorithm Applications	113
The AI Cousins: Case-Based Reasoning and Fuzzy Logic	114
Case-Based Reasoning (CBR)	114
Fuzzy Logic	115
Conclusion	117

CHAPTER 4 TEN STEPS TO MINING YOUR DATA 119

Identify Your Business Objective	119
Baseline	121
Select Your Data	123
Creating Customer Data	124
Brand Sites	124
Banners	126
Buttons	127
Pop Ups	128
Push	129
E-mail	129
Chat	130
Agents	130
Dealing with Transactional Data	131
Prepare Your Data	132
Enhancing Your Data	134
Evaluate Your Data	134
Utility Tools for Preparing and Viewing Your Data	134
Format the Solution	135
Select the Tools	136
Scalability	137
Accuracy	137
Formats	138
Solutions	139
Preprocessing	139
Connectivity	140
Import/Export	140
Memory Management	141
Performance	141
Noise	142
Paradigms	142
Efficiency	143

Tool Selection	143
A Tool Shopping Tip: Kick the Tires before Buying	145
Construct the Models	145
Validate the Findings	147
Deliver the Findings	148
Integrate the Solutions	149
Integrating Your Website and Data Warehouse	149

CHAPTER 5 THE TOOLS 155

Criteria	155
Single Paradigm Tools	159
Database Mining Marksman	159
4Thought	161
KnowledgeSEEKER	163
NGO NeuroGenetic Optimizer	168
WizWhy	170
The Data Mining Suites	173
Clementine	174
DataEngine	178
Decisionhouse	181
Scorecard Builder	182
Other Tools	184
Final Tool Tip	188

CHAPTER 6 THE DATA COMPONENTS 193

Server Log Files	194
Transfer Log File	194
Host Field	195
Identification Field	195
Authuser Field	196
Time Stamp	196
HTTP Request	196

Status Code Field	196
Transfer Volume	197
Error Log	197
Referrer Log	198
Agent Log	198
Extended Log Format	198
Configuring Your Server	199
Server Configuration Tips	199
Apache and NCSA Server Configurations	200
CERN Server Configurations	201
Netscape Server Configurations	201
WebSite Server Configurations	202
The Multiplicity of a Visit	202
Cookie Files	203
A Cookie Future	206
Cookie Applications	207
Configuring Cookies	210
Cookies and CGI	211
Open Profiling System (OPS)	213
Forms	218
Common Gateway Interface	221
Assembling the Components	223
Components Considerations	225

CHAPTER 7 THE PROVIDERS 227

Marketing Data Providers	228
Applied Geographic Solutions	228
Acxiom	229
CACI (ACORN)	233
Claritas	238
Equifax	239
Experian	241

Polk	242
Trans Union	246
Web Services and Software Providers	247
Art Technology Group	248
Aptex	249
Autonomy	250
BroadVision	250
DoubleClick	252
Engage Technologies	253
Firefly	255
GuestTrack	256
Gustos	257
Internet Profiles Corporation (I/PRO)	257
IMGIS	258
InText	259
LikeMinds	259
MatchLogic	260
MicroMass	261
NetPerceptions	262
Vignette	263
Your Data Warehouse	264
Closing the Loop	265

CHAPTER 8 E-RETAILING 267

The Product Is the Message	269
Sale Segmentation	272
The Digital Marketplace	275
Product Customization	277
Consumer Profiling	279
Real-Time Buying	282
Organize Your Storefront	288
Retailing and Mining	290
Electronic Commerce	294

CHAPTER 9	E-MINING	299
	Mining Your Site	299
	Capturing the Look	301
	A Sample Website Session	303
	Interactive by Design	305
	Solid State	305
	The Data Set	308
	Clustering (Association and Sequencing)	310
	Segmentation	318
	Classification and Prediction	325
	Optimization	332
	Visualization	335
	It's an Integrated Iterative Process	339
	Data Mining Future	345
	The Personalization of Self	346
	Right Turns, Hand Grenades, and Integrity	347
APPENDIX	Privacy Consortiums, Standards, and Legislation	349
	Glossary	355
	Index	361