

目 錄

第 1 篇 程序篇	1
第 1 章 資料探勘的概念	2
1-1 前言	3
1-2 資料探勘的定義 (What)	6
1-3 資料探勘的目的 (Why)	12
1-4 資料探勘的方法 (How)	17
1-5 資料探勘的演進 (When)	19
1-6 資料探勘的用途 (Where)	20
1-7 資料探勘的人員 (Who)	24
1-8 資料探勘與知識管理	25
1-9 本書的軟體	27
1-10 本書的個案	27
1-11 本書的結構	28
第 2 章 資料探勘的程序	31
2-1 前言	32
2-2 任務的理解	34
2-3 資料的理解	38
2-4 資料的準備	46
2-5 知識的建模	52
2-6 知識的評價	56
2-7 知識的布署	63

第 3 章	資料探勘的原理	73
3-1	前言	75
3-2	產生測試設計	78
3-3	建構知識模型	83
3-4	評估知識模型	88
3-5	整合知識模型	98
3-6	實例：書局行銷個案	102
第 2 篇	模式篇	105
第 4 章	變數特性與關係的分析	106
Part A	變數特性分析	107
4-1	簡介	107
4-2	變數敘述統計	107
4-3	機率分布型態之性質	108
4-4	機率分布參數之估計	109
4-5	機率分布參數之測試	110
4-6	機率分布型態之測試	110
4-7	資料視覺化	111
4-8	實例一：旅行社個案	114
4-9	實例二：書局行銷個案	116
Part B	變數關係分析	120
4-10	簡介	120

4-11	連續輸入與連續輸出的關係	120
4-12	離散輸入與連續輸出的關係	123
4-13	離散輸入與離散輸出的關係	125
4-14	連續輸入與離散輸出的關係	129
4-15	資料視覺化	131
4-16	實例一：旅行社個案	143
4-17	實例二：書局行銷個案	146
第 5 章	聚類分析 (一): 均值聚類分析	153
5-1	模型架構	154
5-2	模型建立	158
5-3	實例一：旅行社個案	162
5-4	實例二：書局行銷個案	164
5-5	實例三：公民意見分析個案	166
	實作單元 A：Excel 資料探勘系統——均值聚類分析	175
第 6 章	聚類分析 (二): 階層聚類分析	187
6-1	模型架構	189
6-2	模型建立	189
6-3	實例一：旅行社個案	198
6-4	實例二：書局行銷個案	200
6-5	實例三：公民意見分析個案	202
第 7 章	分類與迴歸 (一): 最近鄰居	211
Part A	最近鄰居：分類	212

7-1	模型架構	212
7-2	模型建立	212
7-3	實例一：旅行社個案	216
7-4	實例二：書局行銷個案	221
7-5	結論	223
Part B	最近鄰居：迴歸	226
7-6	模型架構	226
7-7	模型建立	226
7-8	實例一：旅行社個案	230
7-9	實例二：房地產估價個案	233
7-10	結論	237
	實作單元 B：Excel 資料探勘系統——最近鄰居分類與迴歸	239
第 8 章	分類與迴歸 (二)：迴歸分析	251
Part A	邏輯迴歸	252
8-1	模型架構	252
8-2	模型建立	254
8-3	實例一：旅行社個案	260
8-4	實例二：書局行銷個案	262
8-5	結論	266
Part B	迴歸分析	268
8-6	模型架構	268
8-7	模型建立	269

8-8	多項式迴歸分析	285
8-9	非線性迴歸分析	288
8-10	定性變數迴歸分析	293
8-11	逐步迴歸分析	295
8-12	實例一：旅行社個案	296
8-13	實例二：房地產估價個案	298
8-14	結論	303
	實作單元 C：Excel 資料探勘系統——迴歸分析	306
第 9 章	分類與迴歸 (三)：神經網路	321
	Part A 神經網路：分類	322
9-1	模型架構	322
9-2	模型建立	325
9-3	實例一：旅行社個案	336
9-4	實例二：書局行銷個案	340
9-5	結論	343
	Part B 神經網路：迴歸	346
9-6	模型架構	346
9-7	模型建立	346
9-8	實例一：旅行社個案	347
9-9	實例二：房地產估價個案	350
9-10	結論	353
	實作單元 D：Excel 資料探勘系統——神經網路	355

第 10 章 分類與迴歸 (四): 決策樹	369
Part A 決策樹：分類	370
10-1 模型架構	370
10-2 模型建立	373
10-3 實例一：旅行社個案	382
10-4 實例二：書局行銷個案	391
10-5 結論	397
Part B 決策樹：迴歸	398
10-6 模型架構	398
10-7 模型建立	398
10-8 實例一：旅行社個案	400
10-9 實例二：房地產估價個案	406
10-10 結論	411
實作單元 E：Excel 資料探勘系統——決策樹	413
第 11 章 關聯分析	429
11-1 模型架構	431
11-2 模型建立	437
11-3 實例一：商店購物個案	443
11-4 實例二：書局行銷個案	444
11-5 實例三：人才專長關聯分析	445
11-6 實例四：證券漲跌關聯分析	448

第 3 篇 實作篇	457
第 12 章 個案集 (一): 聚類探勘	458
12-1 前言	459
12-2 個案 1: 暖氣系統市場聚類分析	461
12-3 個案 2: 休旅車市場聚類分析	464
12-4 個案 3: 汽車保險市場聚類分析 (光碟)	3
12-5 個案 4: 健身俱樂部會員聚類分析 (光碟)	5
12-6 個案 5: 在職班學生滿意度聚類分析 (光碟)	8
12-7 個案 6: 公民對公共事務意見聚類分析 (光碟)	11
12-8 個案 7: 上市公司的信用評等聚類分析 (光碟)	13
12-9 個案 8: 台灣上市股票基本面聚類分析 (光碟)	15
12-10 個案 9: 台灣上市股票技術面聚類分析 (光碟)	23
12-11 個案 10: 企業貸款違約風險聚類分析 (光碟)	28
12-12 個案 11: 農會信用部風險聚類分析 (光碟)	31
第 13 章 個案集 (二): 分類探勘	471
13-1 前言	472
13-2 個案: 休旅車的潛在顧客開發	475
13-3 個案 2: 汽車保險潛在顧客開發 (光碟)	2
13-4 個案 3: 健身俱樂部會員開發 (光碟)	7
13-5 個案 4: 通信業潛在顧客開發 (光碟)	16
13-6 個案 5: ERP 系統潛在顧客開發 (光碟)	22
13-7 個案 6: 軟體維護合約續約顧客開發 (光碟)	24

13-8	個案 7：賽馬比賽勝負預測（光碟）	26
13-9	個案 8：在職班學生的滿意度評估（光碟）	32
13-10	個案 9：上市公司的信用評等（光碟）	33
13-11	個案 10：企業貸款違約風險預測（光碟）	39
13-12	個案 11：房屋貸款違約風險預測（光碟）	44
13-13	個案 12：信用卡逾期風險預測（光碟）	48
13-14	個案 13：農會信用部風險評估（光碟）	49
13-15	個案 14：台灣上市股票報酬率預測（基本面）（光碟）	55
13-16	個案 15：網購退貨顧客偵測（DMC 2004）（光碟）	58
13-17	個案 16：網購詐欺顧客偵測（DMC 2005）（光碟）	60
13-18	個案 17：捐血者捐血預測（光碟）	66
第 14 章	個案集（三）：迴歸探勘	483
14-1	前言	484
14-2	個案 1：休旅車市場潛在顧客開發	486
14-3	個案 2：在職班學生滿意度分析（光碟）	2
14-4	個案 3：上市公司的信用評等（光碟）	8
14-5	個案 4：選擇權價格預測（光碟）	13
14-6	個案 5：法拍屋拍賣價預測（光碟）	15
14-7	個案 6：台灣上市股票報酬率預測（基本面）（光碟）	17
14-8	個案 7：晶圓不良率預測（光碟）	19

第 15 章 個案集 (四) : 關聯探勘	491
15-1 前言	492
15-2 個案 1 : 商品銷售——以 FoodMart 2000 資料庫為例	493
15-3 個案 2 : 商品銷售——以化妝品銷售為例	497
15-4 個案 3 : 商品銷售——以資訊類教科書為例 (光碟)	2
15-5 個案 4 : 網頁資訊——以網路書局為例 (光碟)	5
15-6 個案 5 : 網路新聞——以台灣股市為例 (光碟)	18
15-7 個案 6 : 證券投資——以台灣股市基本面為例 (光碟)	26
15-8 個案 7 : 產品維修——以 Cable Modem 為例 (光碟)	28
15-9 個案 8 : 產品維修——以印表機為例 (光碟)	32
15-10 個案 9 : 製程診斷——以導線架為例 (光碟)	39
15-11 個案 10 : 製程診斷——以 Touch Panel 為例 (光碟)	44
第 16 章 資料探勘的展望	503
16-1 資料探勘的重要觀念	504
16-2 資料探勘的現況調查	505
16-3 資料探勘的面臨困難	506
16-4 資料探勘的社會衝擊	508
16-5 資料探勘的研究方向	509