

目 錄

摘要	
目 錄	
圖 索 引	
表 索 引	
第一章、緒論	1
第一節 前言	1
第二節 研究背景	2
第三節 研究目的	4
第四節 論文流程	4
第五節 研究限制	7
第六節 論文架構	8
第二章、文獻探考	9
第一節 資料庫姿識發掘	9
第二節 資料探勘	10
第三節 關聯法則	21
第四節 決策樹	28
第三章、研究設備及方法步驟	35

第一節 研究方法	35
第二節 研究流程步驟	39
第四章、實驗成果	40
第五章、結論與後續研究	64
第一節 結論	64
第二節 後續研究及建議	65
參考文獻（中文）	67
參考文獻（英文）	69
附錄一	71

圖索引

1. 圖 1-1 研究步驟流程圖 5
2. 圖 2-1 資料庫知識發掘之程序10
3. 圖 2-2 將顧客分為四個群集17
4. 圖 2-3 資料探勘的過程 18
5. 圖 2-4 決策樹29
6. 圖 2-5 決策樹建立流程34
7. 圖 3-1 研究流程圖 39
8. 圖 3-2 Large Item Sets 產生過程43
9. 圖 3-3 決策樹範例 49
- 10.圖 4-1 資料篩選流程圖 51

表索引

1. 表 2-1 Apriori 演算法	26
2. 表 2-2 Apriori-gen function.....	26
3. 表 2-3 交易資料庫範例	27
4. 表 2-4 pass 1 candidate 和 large item sets 的產生	27
5. 表 2-5 pass 2 candidate item sets 和 large item sets 的產生	27
6. 表 2-6 pass 3 candidate item sets 和 large item sets 的產生	28
7. 表 3-1 部分門診科別表	36
8. 表 3-2 門診資料表範例	37
9. 表 3-3 病患資料庫原始欄位說明	40
10.表 3-4 無效門診資料表範例	41
11.表 3-5 計算文字頻率	44
12.表 3-6 病患主觀描述文字代號資料範例	45
13.表 3-7 代號後的 Large Item Sets 範例	45
14.表 3-8 病患門診資料的疾病名稱和疾病代碼之欄位表	47
15.表 3-9 屬性產生表	48
16.表 4-1 『疾病代碼 1』原始 ICD-9 碼	51
17.表 4-2 『疾病代碼 1』ICD-9 碼大分類後	52

18.表 4-3 有效資料病患門診紀錄欄位表.....	52
19.表 4-4 『主觀描述』欄位裏原始資料.....	54
20.表 4-5 高頻文字	55
21.表 4-6 主觀描述文字內容代號化後	60
22.表 4-7 Large Item Sets	61
23.表 4-8 決策樹.....	52