

# 以資料探勘技術建立宅配業之車輛維修及 預警決策支援系統

## A Vehicular Maintenance and Replacement Decision Support System in Distribution Services: A Data Mining Technique

吳怡瑾 I-Chin Wu 李睿傑 Ruei-Jie Li 陳子立 Tzu-Li Chen

天主教輔仁大學資訊管理系

Department of Information Management, Fu Jen Catholic University

(Received June 13, 2011; Final Version December 22, 2011)

**摘要：**物流宅配業隨著消費行為的改變及電子商務的興起正蓬勃的發展，但是當其業績不斷成長之時，人員薪資、車輛運作與維修費用也隨之攀升；因此，如何有效降低車輛相關的營運成本已是物流宅配業者當前之重要課題。研究由物流配送公司中車輛管理人員的角度出發，發展智慧型的車輛零件預防更換及危險車輛預警系統，透過網路平台協助非專業車輛維修技師的車輛管理人員，能輕易的進行預防性維修 (preventive maintenance)，即在車輛尚未失效前進行預防更換 (preventive replacement) 或謹慎地依據生產程序進行保養維修。實作上，研究首先透過資料探勘技術之關聯法則及序列型樣分析，找出不同車型維修零件的相關聯性及先後順序，建立預防性維修知識庫以輔助車輛管理人員進行車輛零件預防更換的決策分析。研究更進一步運用歷史維修記錄透過決策樹技術建立一套危險車輛預測模型並進行模型的驗證，研究結果顯示該危險車輛預測模型之平均正確度 (accuracy) 可以到達 96.98%。研究為使車輛零件維修預防更換及危險車輛預測知識與服務能夠不受時間及空間的限制，將建構 Web-Based 服務平台並利用 .Net 網頁建構技術建置此系統，讓車輛管理人員能透過平台以獲得即時最佳的車輛保養、零件更換建議與危險車輛預警服務。最後研究模型將落實在個案公司的車輛管理系統中以驗證方法之正確性與系統之可行性。

**關鍵詞：**關聯規則、資料探勘、決策支援系統、決策樹、車輛預防性更換

**Abstract:** As e-commerce has grown exponentially, the business of the distribution service is also growing and expanding quickly in recent years. Nevertheless, from the perspective of the merchant, the maintenance, repair and operations (MRO) fees of vehicles in distribution service is also increasing dramatically. In this project, we develop an intelligent maintenance and replacement system to help the manager and technicians conduct preventive maintenance and replacement for vehicles. Practically, we employ association rule mining and sequential pattern mining methods to analyze the relationships and the priorities among vehicles' components to execute preventive maintenance. Furthermore, we employ the decision tree algorithm to predict the vehicles dangerous based on the historical maintenance lists. In this work, we also validate the performance of the construct model which can achieve 96.98% in accuracy rate. Consequently, it can help the manager and technicians to make decisions on either repairing the vehicles or selling them out. For taking the advantage of Web-based Platform, we adopt the .Net technique to develop the intelligent system based on the proposed methods; therefore, employees can access the services anytime-anywhere via the Internet. Finally, we will realize the system in the distributions service to evaluate the accuracy and feasibility of the proposed model and system in the real operation condition.

**Keywords:** Association Rule Mining, Data Mining, Decision Support System, Decision Tree, Vehicular Preventive Replacement

## 1. 前言

物流宅配業隨著消費行為的改變及電子商務的興起正蓬勃的發展，根據行政院主計處 2001 年調查，台灣在流通運輸服務業，其生產總額約 2.67 兆元，占整體服務業生產總額之 36.27%。2010 年國內經濟情勢檢討報告，也指出 2010 年流通運輸服務業的實質成長為 6.06%，占服務業名目 GDP (gross domestic product；國內生產總值) 比率 34.59%，且對經濟成長貢獻 1.41 個百分點，故可見流通運輸服務業對全國服務業產值與經濟成長率影響很大。若以 2001 年政府相關資料估算，流通運輸服務業就業人員數量約達 204 萬人，約占服務業總就業人口數量之 59.7%。台灣早期並沒有類似講求服務及客制化的物流產業，比較接近宅配的服務業為郵局的包裹，而宅配業興起的分水嶺大約是在 2000 年，統一速達及台灣宅配通，將日本行之多年的宅配服務引進至台灣。近年風行的電子商務、電視購物，為宅配業者注入的驚人的成長力道，以電子商務來看 2004 年市場規模不到 600 億、2007 到達 1500 億的商機，而 2010 年已經挑戰 3000 億的商

機；而在 Yahoo!提供電子商務的興奇科技，2004 年寄件量為 70 萬件，到了 2007 年提高到 900 萬件，成長了 12 倍（楊玟欣，民 96）。台灣宅配業的發展可以概分為三個時期：市場快速成長期、市場調整選擇期、市場成熟穩定期。市場快速成熟期與郵購、電視購物與網拍的起步有很大的關聯，由於單一郵局的宅配服務已經不敷需求，故有不少宅配業進入市場；市場調整選擇期與市場成熟穩定期則與網拍市場的快速成長與宅配業務的多元化有直接關聯（林建煌，民 96）。因此，由目前線上購物的成長速度觀之，國內的物流宅配業也跟著蓬勃成長，這對台灣的物流宅配產業應該是一劑強心針；反之，若沒有物流宅配業者的完美配合，電子商務與線上購物也難成氣候。

台灣目前五大宅配業者分別為統一速達（黑貓宅急便）、台灣宅配通、新竹貨運、大榮一日配與郵局（宋建達，民 97）；各家業者背後資源各自不同，一方有龐大通路做支援，一方有貨運經驗、郵務體系為基礎，各有其競爭優勢。但不管其背後資源如何，就物流宅配業來看，是一個勞力密集、固定資產密集的行業，當業績不斷成長的時候，人員及固定資產也就會同步向上增高。以目前物流宅配業者來看每月所花費的成本當中最高占比就是人員薪資，再來便是油錢與車輛維修費用。以新竹貨運為例，從 2007 年以來油價上揚，造成每月成本增加一千五百萬元，因此車輛雖是生財工具但也是一項重大的成本及費用支出，當因車輛所延伸出的費用愈高的時候，針對整個物流宅配業者的毛利就壓縮的更多了，在當下競爭激烈的物流宅配業中，平均毛利大約是 2~3% 來看，只要能降低車輛相關延伸費用 1% 的話，那相對的整體獲利率就能提昇。以目前物流宅配業者龍頭公司 2008 年為例（即本研究之個案 T 公司），每月各型車輛的維修及保養的費用將近 800 萬，且油價飆漲的情況下，如何降低車輛相關的營運成本便是經營層級一直苦思的課題。由於車輛管理人員並非專責管理且也無相關的車輛維修專業知識，所以只能仰賴車輛原廠提供的通則性的保養規範及車輛維修技師的經驗以進行維修；但國內維修車廠良莠不齊、維修人員的維修能力不一，維修診斷的知識與經驗無法適時保存與利用。

Nemati *et al.* (2002) 曾提出知識倉儲 (knowledge warehouse) 的概念，作者由傳統資料倉儲的概念延伸提出整合知識管理與人工智慧相關技術於決策支援系統實作中。作者指出其提出知識倉儲 (knowledge warehouse) 的實作系統除保有傳統知識擷取與儲存功能外亦進一步考量知識搜尋與分享，透過該整合性的平台方能有效輔助企業進行決策分析。郭家齊 (民 93) 曾提出整合資料探勘與案例式推論技術於機台故障診斷，其目的希望透過自然語言處理程序建立維修案例，然後採用決策樹及倒傳遞類神經網做案例索引及推斷，讓機台設備廠商能立即查明故障原因，加速維修時間。維修形式主要區分成四種，置換方式取代維修 (replacement Instead of Maintenance)、計劃型換置 (planned replacement)、故障維修 (breakdown maintenance)、預防性維修 (preventive maintenance)，而故障機台案例式推論就屬於故障維修的一種 (Clifton, 1974; Edwards *et al.*, 1998)。蔡有藤 (民 88) 指出預防維護 (preventive maintenance) 於裝備開始運作後

藉由建立的程序來進行階段性的維護工作，以預防裝備在使用時發生功能性失效；而預防更換是指針對不可修理、或是經過幾次維修後不值得再修理的組件或系統，在尚未失效前進行更換。蔡燕純（民 93）應用馬可夫預防於維修保養策略之探討中指出，各式器具在有預防保養及定期修理的情況下，一定比無維護來得可靠及退化會減緩，但是如何在考慮成本的因素下，達到可以節省成本找出最佳的預防修理週期是一個重要的課題；該研究藉由馬可夫鏈的機率模式並結合動態可靠度模型以規劃系統的維護策略，觀察不同維護策略及修理週期下系統性能的變化，並引入成本觀念獲取最佳的維護時機。Buddhakulsomsiri and Zakarian (2009) 應用序列型樣演算法於汽車維修大量資料，該研究主要著重於分析發掘維修問題之間的規則並評估序列型樣演算法效能。該研究指出汽車維修決策支援系統應能包含資料整合、資料前處理與圖形化介面等模組方能有效應用在產業界。

本研究主要目的在建立預防性維修與預警決策系統，因此，如何運用商業智慧相關技術將過往的維修記錄做一個整理及挖掘，並配合專家建議及診斷，建立出針對不同車款、不同車齡的零件維修更換知識庫幫助個案業者的車輛管理人員獲得基本的維修知識，進而提供車輛預防維修或是車輛零件預防換置的決策分析，以有效達成車輛維修費用與營運成本降低為本研究動機之一。台灣的大型物流宅配業者，為了因應台灣多變的地形與街道設計，所以會準備多樣的車種來符合不同的配送環境。以國內物流宅配業龍頭之個案公司來看，便擁有 1 千 3 百台車及 16 種左右的車種，管理上勢必複雜許多，且為了節省成本，車輛會盡量延長使用年限，以不斷維修來代替購置新的車輛，但老舊車輛或是過度頻繁維修的車輛是否該繼續維修下去，或是報廢購置新車輛來替換，一直都沒有標準化決策模式。所以如何運用歷史維修記錄建立一套危險車輛預測模型，以幫助業者能迅速的做出車輛維修或出售的決策判斷，為另一重要研究動機。物流宅配業者目前大多無專屬的車輛維修管理人員，所以對車輛的維修及保養都只能仰賴通則性的定期保養規範及維修車廠技師的個人經驗做判斷，且無法有效做出預防更換的預測，所以本研究運用資料探勘中關聯法則、序列型樣與決策樹建構一套車輛維修與預測之系統，並且把系統建構在網際網路平台，讓車輛零件維修預防更換及危險車輛預測知識與服務能夠不受時間及空間的限制。以下簡述研究之主要目的：(1) 運用資料倉儲與資料探勘 (data mining) 技術，針對物流宅配業者車輛的維修記錄，整理出待分析的資料；(2) 利用關聯法則及序列型樣，找出維修零件的相關聯性及先後順序，做出車輛零件預防更換的建議；(3) 利用決策樹建立危險車輛模型，協助車輛管理人員能預測及判斷出維修車輛是否為危險車輛；(4) 建立 Web-Base 服務平台，讓車輛管理人員能透過網際網路平台能即時的獲得最佳化的車輛定期保養計畫及取得零件更換建議之決策。

## 2. 文獻探討

### 2.1 決策支援系統於車輛維修之應用

在現今激烈的競爭環境中，企業存活要件為能隨時掌握及了解世界的脈動與挑戰，並且能夠在第一時間正確及快速的反應。而商業智慧正是能幫助企業內決策者利用大量經過匯總、運算、整理後的資料並透過視覺化的呈現，以縮短決策時間的有利工具。美國 Gartner 公司的分析師 Howard Dresner (1989) 首次提出「商業智慧」(business intelligence) 名詞，主要概念是透過對長期而大量的資料分析，獲得穩定可靠的知識，讓企業決策者除了依據過往經驗外，更有充足的數據資料做決策時重要的參考 (王志堅，民 97)。因此，商業智慧的關鍵是從許多來自不同的企業運作系統中提取出有用的資料，進行資料清潔用以保證資料的正確性，然後經過資料擷取 (extraction)、轉換 (transformation) 和載入 (loading) 後，將企業資料轉入至資料倉儲之中。在此基礎之上利用線上即時分析處理 (OLAP)、資料探勘 (data mining) 等技術，對其進行分析及處理，將資訊變成輔佐決策的知識，最後再將知識透過不同的系統，例如：客戶關係管理系統、企業資源整合系統、供應鍊管理系統...等，呈現給企業內不同階層的管理者進行決策。目前提供企業之商業智慧解決方案的軟體廠商主要結合資訊科技與企業智慧資產提出整合性解決方案 (張紂微等，民 90)，如：延伸資料庫或資料倉儲系統功能、整合 ERP 資料，強化資料倉儲與報表分析功能、結合大量資料統計運算功能 (例如 SAS、SPSS)、由 OLAP 分析與報表系統向下整合。

Yam *et al.* (2001) 基於預知維修 (condition-based maintenance; CBM) 概念提出智慧型預測決策系統 (IPDSS)，該系統主要目標為：(1)設備維護；(2)智慧型錯誤偵測與(3)設備耗損趨勢預測。該研究以遞迴式類神經網路 (recurrent neural network) 建立預測模型並提供診斷介面協助發電廠人員預先規劃與維修設備。Smith *et al.* (2010) 使用類神經網路方法建立航空系統的機門損壞預知維修 (CBM) 系統；該系統雖屬於雛形階段，但已經初步證實所建立模型的可用性並預計進一步提出針對不同機門的客制化預測模型。Last *et al.* (2010) 根據汽車歷史保養記錄資料提出 Multi-Target Information-Fuzzy Network 之資料探勘預測模型，該研究顯示透過所建構模型可以發現危險車輛與預測車輛維修時間。本研究基於上述相關研究，根據不同車種提出建立預防性維修與預警決策支援系統，研究運用商業智慧相關技術將過往的維修記錄進行知識發掘程序並配合專家建議及診斷，建立零件維修更換知識庫以幫助個案業者的車輛管理人員獲得基本的維修知識並提供車輛預防維修或是車輛零件預防換置的決策分析輔助系統。

### 2.2 關聯法則與序列型樣

資料庫知識發掘 (knowledge discovery in databases; KDD) 或資料探勘 (data mining) 技術以自動化的探勘流程在大量資料中發掘有趣或有用的資訊或資料樣式 (patterns)，進而輔助企業

進行決策 (Han and Kamber, 2006)。資料探勘相關研究領域包含資料庫、知識學習、人工智慧、專家系統、特徵識別、統計學及資料視覺化等 (Chen *et al.*, 1996; Han and Kamber, 2006)。在資料探勘的步驟中，前置處理動作包含資料清潔 (data cleaning) 與整合 (integration)、資料轉換 (transformation) 與選擇 (selection)等。而資料探勘的技術包含關聯規則 (association rules)、序列樣本 (sequential patterns)、資料分類 (classification) 及資料分群 (clustering) ...等。

關聯法則是資料探勘主要的技術之一，關聯規則分析可以在大量資料中找尋資料項目 (如購買物品) 之間的相關性，典型的應用包含超市購買行為分析等。關聯式規則提供支持度 (表示項目出現的頻率)及信賴度 (表示項目間的強度)，來表示關聯式規則的關聯性，以瞭解關聯式規則的強度 (Agrawal *et al.*, 1993; Chen *et al.*, 1996; Han and Kamber, 2006)。Agrawal *et al.* (1993) 所提出關聯法則定義如下，令  $\mathbf{I} = \{i_1, i_2, \dots, i_m\}$  為一群項目的集合 (Items)， $\mathbf{D}$  是一群交易 (transaction) 所組成的集合，每筆交易  $T$  是由一群項目  $\mathbf{I}$  中任意項目的子集合，若一個項目組包含有  $k$  個項目，稱之為  $k$ -項目組 ( $k$ -itemsets)，以  $\text{itemset}_k$  表示之， $k \geq 1$ 。在項目組  $X$  與  $Y$  之間有一關聯規則被表示成  $X \rightarrow Y$ ， $X, Y \subseteq \mathbf{I}$ ，且  $X \cap Y = \emptyset$ 。其中  $X$  與  $Y$  分別表示在資料庫中不同的資料項目組，即若購買項目集合  $X$  時，可能會在購買項目集合  $Y$ ，每一條關聯法可透過支持度 (support) 及可信度 (confidence) 以判斷所找出的關聯法則是否有意義。一般支持度及可信度都是由使用者所自訂，一個有效的關聯法則其支持度及可靠度均須大於或等於使用者所訂的最小限制，則此關聯才被認為是有意義的。關聯法則最具代表性的方法是 Apriori (Agrawal and Srikant, 1994) 其主要執行過程分為兩個步驟：(1)逐漸擴展找出所有高頻率項目集，即找到支持度大於是先設定的最小支持度的項目集合；(2)透過分析高頻率項目集以建立合適的關聯規則，其公式如下：

$$\text{Support } (X \rightarrow Y) = P(X \cup Y) \quad (1)$$

$$\text{Confidence } (X \rightarrow Y) = P(Y | X) = \frac{\text{support}(X \cup Y)}{\text{support}(X)} \quad (2)$$

陳星壁 (民 95) 指出在國內軌道運輸業在營運過程中，常會碰到電路板件故障與損壞的問題，而在企業內部維修過程中，常因維修準則判不易及維修工程師經驗傳承等問題，導致增加維修時程與成本。所以作者運用資料探勘技術中關聯法則的技術，設計一套電路板件維修輔助系統，該法主要針對電路板件故障問題與維修準則之間的關聯進行分析，以提供企業內部維修工程師參考之用，且同時針對電路板件的潛在故障問題提供維修準則的建議，因此可以進行故障預防檢修的動作，降低零件未來的故障率。Lin *et al.* (2001) 將資料探勘技術應用於中風患者的臨床路徑 (clinical pathways) 的分析，該研究主要透過不同病患在不同時間的醫療資源使用活動以建構出臨床路徑圖型並探勘經常性臨床路徑。該研究結果有助於新進患者醫療流程的預測並進而有效利用改善醫療流程。資料探勘另一個有效而廣為應用的技術為序列型樣 (sequential pattern)，其原理與關聯規則相似，但是該法在計算物件組時會把時間的先後關係納入考量，曾

經發展出 Apriori-like、AprioriAll 與一般化的序列型樣探勘演算法 (Agrawal and Srikant, 1995; Srikant and Agrawal, 1996)。序列型樣 (sequential pattern) 所考量的是各種資料項目之間的順序性，所以如果把每個資料項目的前後位置視為一個特徵值的話，那麼在找出序列型樣時，就可以用來當作決定資料項後順序的主鍵值，例如交易時間，使其為一維向度的呈現，接著再利用相關演算法及使用者所給定的門檻值，找尋頻繁出現的序列 (frequent patterns)，而這些序列大量的在資料庫中出現，應該代表某種意涵。序列型樣資料探勘實作廣為運用在銷售分析、推薦或是預測系統之上。Buddhakulsomsiri and Zakarian (2009) 應用序列型樣演算法於汽車維修大量資料，該研究主要分析發掘維修問題之間的規則並評估演算法的效能。舒毓筵 (民 93) 在醫學報告資料中探勘連續序後，以序列型樣找出病患下次可能會就診的疾病，提供醫生診斷的參考及預測；謝智聿 (民 91) 以序列探勘理論實作之網路異常型樣驗證，預測網路會發生異常問題的先後順序及關聯性。

### 2.3 決策樹

決策樹是一種常用的預測模型的演算法，預測 (prediction) 技術乃依據某一特定對象屬性，觀察過去的行為或是歷史資料，用以推估其未來的可能性。決策樹是藉由一連串的問題和規則將資料分類，由相似的型態來推測相同的結果。由於決策樹是將資料依據不同的變數循序來產生分析結果，可藉由決策樹分析方式來分析目標特質與異同點。決策樹演算法是目前資料探勘中常用的方法之一，該方法吸引人之處在於有規則的運用樹狀圖來表達資料規則的路徑，另外也可以用文字表達或是轉換成 SQL 之類的資料庫語言，讓落在特定類別的資料記錄可以被搜尋。決策樹分析模式主要演算法包括 ID3、C4.5 及 CART (Han and Kamber, 2006; Turban *et al.*, 2005)。

決策樹建構的步驟首先選擇一個屬性，利用屬性分類，計算分類後平均分散度，若分散度降得最低即為最佳分隔屬性，沒有屬性可以降低分散度或是只有一個數值時即停止。決策樹建立之後，可以用樹的修剪 (tree-pruning) 步驟來縮減決策樹的大小，決策樹如果太大，很容易有過度學習 (overfitting) 的現象，所以修剪樹的動作將有助於改善決策樹類推至其他資料的能力。決策樹的優點除可產生易於瞭解的規則外，也可以處理連續性和類別性變項，不需要太多的計算就可以進行分類。但當樣本數目太少或是類別太多時容易造成混淆，則是決策樹的缺點。而決策樹模型的建立，常運用在診斷系統之上，例如在廖介銘 (民 92) 在決策樹應用於糖尿病之探勘研究，以決策樹演算法找出糖尿病患者的決定影響因子。郭家齊 (民 93) 整合資料探勘與案例式推論於機台故障診斷，將眾多的維修案例以決策樹演算法建立案例庫的故障索引表。

### 3. 研究方法

本章中將介紹如何運用資料探勘的技術以建構出車輛維修知識庫，並進而建立危險車輛預警模型與決策系統，讓各級的車輛管理人員都能受益。

#### 3.1 研究問題定義

車輛中各項零件故障是不可抗力的現象，但如何針對不具維修價值的組件或系統，在尚未失效前進行預防更換 (preventive replacement) 或謹慎地依據生產程序進行保養維修，以防止在運轉途中發生故障情況的預防性維修 (preventive maintenance)，是本研究主要探討的問題之一。車輛在經過重大事故發生後或車輛本身出廠時便有瑕疵存在，會導致日後維修成本大幅上升，所以如何運用歷史維修記錄將各車種中維修金額超出平均值的車輛建立危險車輛預測模型，是研究另一重要探討問題。本個案將車輛依據車齡切割為 2 年內、3~6 年及 7 年以上，為根據個案公司車輛會計折舊原則與使用年限資料進行前置分析。研究參考車輛維修專家建議，預期針對不同使用年限車輛進行維修規則與序列探勘以提供個案公司零件維修參考依據；此外，並對使用年限較長之車輛進行維修或置換之預警決策。無論是車輛零件的預防更換或是危險車輛的預測模型建立，都是必需靠本身產業知識的不斷累積才有辦法獲得，因此藉由資訊技術可協助業者有效累積本身有用的資料，再經過資料的整理及分析後，建構出一套車輛維修知識庫及決策支援系統，達到降低維修成本的目的。

在研究個案中車輛維修目前屬於無系統化管理狀態，所以對於維修知識的獲得或是傳承，通常只靠各級車輛管理人員彼此間詢問及溝通或是仰賴車輛原廠技師提供，因此，車輛維修知識無法有效累積更無法分享。本研究將利用資料探勘的技術建構出一套車輛零件維修預測建議及危險車輛預測系統，透過資料探勘的技術將過去只是條列式記錄的維修資料，經理整理及挖掘後，建立出車輛零件維修知識庫及建立危險車輛預測模型。研究並結合網際網路將維修知識分享，透過 WEB 化介面呈現維修預測的結果，讓各級車輛能真正有效管理車輛。研究首先使用關聯法則及序列分析資料探勘技術以探勘出車輛維修零件間的相關聯性及維修的先後順序關聯性，接著將探勘出的規則資料建立維修知識庫；研究並根據序列分析結果篩選重要零件以建立危險車輛決策樹預測模型。該車輛維修知識系統考量日後的維修預測及危險車輛預測服務之使用便利性，因此，透過 .NET 技術與 SQL Server 資料庫技術，讓智慧型車輛維修及預警系統可以在任何地點及時間於在網路平台上運作。

#### 3.2 系統架構

本決策支援系統建置分為兩大部份，分別為系統前端維修預測及危險車輛預測介面呈現和後端維修預測及危險車輛預測建構成的維修知識庫。



- **系統前端平台：**當車輛管理人員於每次輸入車輛維修零件資料後，系統將連結後端維修知識庫，於介面提供本次可能會伴隨一併維修的零件資料，或是下次可能會接續維修的零件資料，主要提供車輛管理人員做預防更換的決策判斷。另外系統也會透過後端建立起的危險車輛模型，判別該輛車輛是否符合危險車輛的狀況並通知車輛管理人員。
- **系統後端處理與分析：**後端系統將分為三大部份，分別為資料的前處理、車輛維修零件關聯法則和序列型樣資料挖掘及使用決策樹建立危險車輛模型。第一部份之資料前處理主要目的為透過專家的協助將歷史維修記錄中，定期保養資料，常識性資料及不完整或缺漏的資料移除，保留真正有意義的資料。第二部份車輛維修零件透過關聯法則 (association rule mining) 和序列型樣 (sequential pattern mining) 探勘維修零件的相關聯性。研究主要根據資料前處理完的車輛維修資料倉儲資料，然後再按照不同車種及不同車齡進行挖掘，找出每次維修零件間的相關聯性。第三部份也是依照不同車種及車齡，利用決策樹建立出危險車輛模型，並將結果回存至維修知識庫中，當下次維修時可與危險車輛模型進行比對。

### 3.3 系統運作流程

本研究的架構流程如圖 1 所示，車輛零件維修預防更換預測及危險車輛預測可分成三個部份，分別是車輛一般維修零件關聯性的建立、危險車輛的預測模型建立及車輛每次維修時維修預測系統的運作，各個流程編號所代表的義意如下說明：

- **車輛維修資料倉儲建立：**蒐集車輛維修資料庫及歷史維修記錄，透過資料倉儲中的資料提取、資料轉換及資料匯入三個步驟，可以匯集及整理 T 公司的車輛維修資料至資料倉儲中，並且經過資料前處理讓車輛維修資料無「雜質」後，才進行後續的資料探勘分析。
- **以關聯法則探勘維修零件間相關聯性：**車輛維修資料倉儲建立後，可以使用關聯法則推導所有車輛維修零件之間的規則。而考量到各類型車種間的零件的差異性及車齡不同會造成維修零件的關聯性不同，所以在使用關聯法則進行資料挖掘前，會將維修資料先依不同車種、再依車齡切割成不同群資料。而透過關聯法則挖掘出不同車種及不同車齡的零件相關聯性後會交由專家判定規則的價值，再將探勘後有用的資料回存至車輛維修知識庫中。
- **車輛維修零件序列型樣 (sequential patterns) 推導：**我們使用序列型樣資料挖掘技術找出車輛維修零件間維修的先後順序關係。進行序列型樣資料探勘時，也會採用與關聯法則相同的資料分割方式，依據不同的車種及不同的車齡進行資料的切割，然後分別進行序列型樣的資料挖掘。
- **建立車輛維修知識庫：**透過關聯法則與序列型樣分析所探勘的車輛維修零件關聯性及維修先後順序性的規則，需要透過車輛維修專家確認及修正後，才能回存到知識庫中。其中，部分挖掘出來的規則是屬於常識性規則，需先將常識性資料先排除後，才將真正有價值的知識存進知識庫中，以利後續前端車輛零件預防更換系統操作使用。

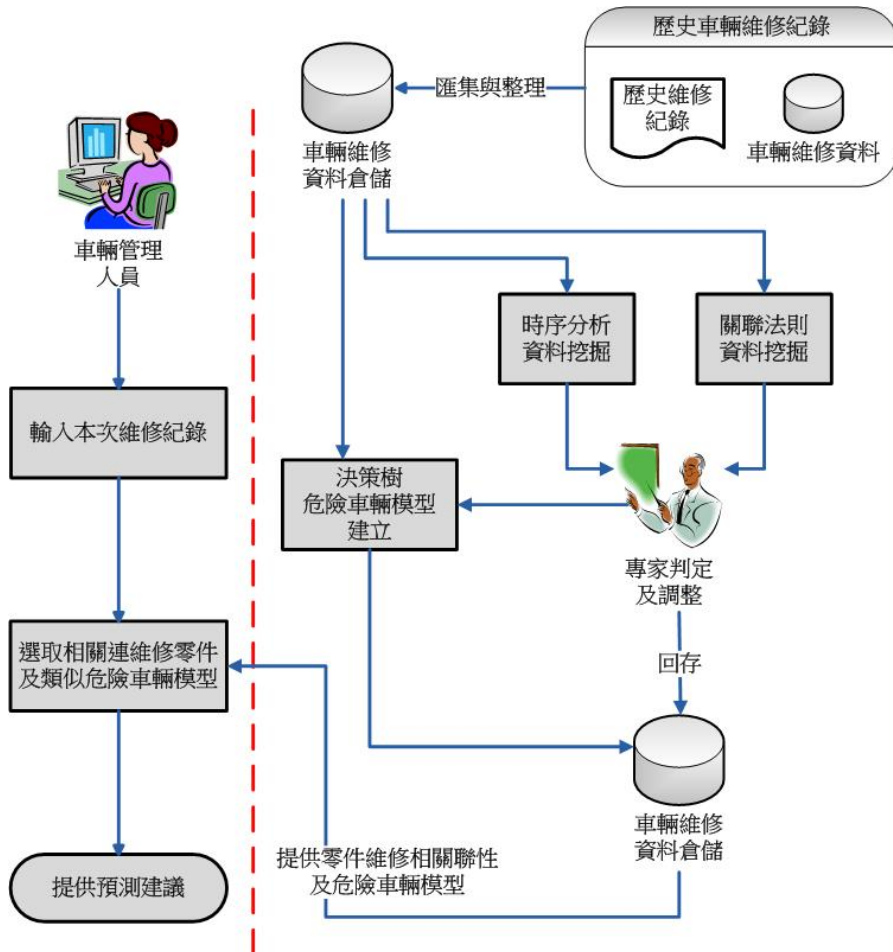


圖 1 車輛維修及預警模式系統架構及流程圖

- **利用決策樹建立出危險車輛預測模型：**運用決策樹我們可以找出在車輛維修資料倉儲中危險車輛的維修資料，研究將透過前一關聯規則分析步驟所得知結果，歸納分析出相關屬性變數(車輛零件)的重要度，及配合車輛的外部屬性資料挑選輸入變數，並將維修金額超過高標金額的車輛挑選成危險車輛，接著運用 C4.5 決策樹演算法建立危險車輛的預則模型。
- **智慧型車輛維修及預警系統：**車輛管理人員於每次車輛維修後，將可透過智慧型車輛維修及預警系統所提供之 WEB 化的輸入介面，輸入當次維修車輛的車種及車齡後，再輸入當此維修零件資料。系統將主動透過車輛維修知識庫的知識與規則，提供車輛零件維修的預防更換的預測建議；此外，系統也將透過相同車種的危險車輛決策樹預測模型，判別該維修車輛是危險車輛的可能性機率，以提供車輛管理作為車輛更換或維修之決策參考。

### 3.4 系統評估方法

研究透過關聯法則及序列型樣分析出來的車輛維修零件相關聯性，會不斷的與車輛維修領域的專家溝通討論，以解讀出正確的原因與意義，並驗證其正確性及可應用性。此外車輛維修專家也可提供本身的專業知識與經驗，以作為進一步改良的意見並評估資料探勘的模式是否需要改進。針對透過決策樹建立的危險車輛模型主要是以混亂矩陣 (confusion matrix) 來評斷，所謂混亂矩陣是把預測結果與實際結果情況進行比對，以矩陣的方式呈現，如表 1 所示。表中  $t\_pos$  表示為實際為安全車輛而預測亦為安全車輛之數量； $t\_neg$  表示為實際危險車輛而預測亦為危險車輛之數量。在實際應用中，每種不同的預測錯誤與預測正確的價值是不同的話，那價值最高的將是我們所要選擇的。本研究使用混亂矩陣得到正確率 (Accuracy)、正確辨識率 (TP-rate) 與錯誤辨識率 (TN-rate) 以判斷該模型是否準確。正確率 (Accuracy) 所代表意義為所建立的預測模型可以預測安全車輛與危險車輛的正確率；正確辨識率 (TP-rate) 為安全車輛之預測能力與錯誤辨識率 (TN-rate) 表示危險車輛的預測能力，如公式(3)、公式(4)與公式(5)所示。我們更進一步與車輛管理人員討論，提出以下評估得分的模式 (Scoring)，如表 1 所示。本研究以模型正確分類得分為 1 分為基礎，當安全車輛分類錯誤成危險車輛時，將可能導致後續過度維修或是提前報廢，造成損失，以扣 3 分表示。當能正確預測出危險車輛，而讓車輛管理人員能加強防範減少損失，以得分 10 分表示。至於是危險車輛而預測成安全車輛，則有可能造成後續重大損失，則以扣 7 分表示。本研究的分數計算為與該領域專家討論之結果並透過分數高低表示模型預測結果優劣。

$$Accuracy = \frac{(t\_pos + t\_neg)}{(t\_pos + f\_pos + f\_neg + t\_neg)} \quad (3)$$

$$TP - rate = \frac{t\_pos}{(t\_pos + f\_neg)} \quad (4)$$

$$TN - rate = \frac{t\_neg}{(t\_neg + f\_pos)} \quad (5)$$

表 1 混亂矩陣 (confusion matrix) 與危險車輛預警評估得分模式 (scoring)

實際 (Actual)	預測 (Predication)	
	安全車輛	危險車輛
安全車輛	$t\_pos$ Score=1	$f\_neg$ Score=-3
危險車輛	$f\_pos$ Score=-7	$t\_neg$ Score=10

## 4. 車輛維修及預警模式與系統建立

### 4.1 車輛維修資料倉儲建立

本系統之車輛維修資料倉儲透過資料倉儲中的資料萃取、資料轉換以及資料匯入三步驟產生與建立。資料探勘程序，首先資料需要夠正確以避免有“garbage in, garbage out”的狀況發生，若是資料中帶有「雜質」或是有遺漏值將會影響探勘的結果；因此，透過資料倉儲的三步驟，可以將不完整的車輛維修資料進行過濾與整理。本研究將車輛維修資料匯集到一個資料集，在匯集前的資料表內，有些資料筆數其中部分欄位值有遺漏，因此需要向資料庫管理員再確認是否完整或是人工補齊資料。若是此筆資料不完整，則不放進車輛維修資料倉儲內，最後在案例倉儲的資料是完整且無遺漏值。

本研究取得研究個案公司在 2008~2009 年間所有車輛的零件維修記錄，並儲存在車輛管理維修資料倉儲中，分別為維修零件資料表、維修主檔資料表及維修明細資料表；其中，資料倉儲中共有 43090 筆維修工單（即維修次數）與 182,249 筆零件維修明細。其中維修零件資料表是車輛維修零件基本資料表，是經專家整理，將相近或是同義的車輛零件，經過專家的整理編碼後，分類並定義出車輛零件大分類、車輛零件中分類及零件基本資料出來，如表 2 所示；維修主檔資料表代表這段期間每次車輛維修所產生的維修工單資料，包含工單編號、車號、車種、進廠日期等資料；而維修明細資料表記錄著每一張維修工單實際維修的零件資料。而為了進行

表 2 維修零件資料表內容

T125_T120_ID	T125_T121_ID	T125_T122_ID	T125_ID	T125_NAME
01	01	01	010101	汽門油封
01	01	02	010102	汽門頂桿
01	01	03	010103	汽門頂筒
01	01	04	010104	汽門彈簧
01	01	05	010105	汽門間隙調整...
01	01	06	010106	汽缸床墊片
01	01	07	010107	汽缸頭總成 (三期)
01	01	08	010108	汽缸頭總成 (四期)
01	01	09	010109	汽缸套組壹台份
01	01	10	010110	汽缸頭試水壓
01	01	11	010111	汽缸頭研磨
01	01	12	010112	更換汽缸頭工資
01	01	13	010113	活塞
01	01	14	010114	活塞環組
01	01	15	010115	機油泵浦

註: T125\_T120\_ID 為零件大分類代號; T125\_T121\_ID 為零件中分類代號; T125\_T122\_ID 為零件小分類代號; T125\_ID 為零件代號; T125\_NAME 為零件名稱。

車輛零件維修間的序列型樣資料挖掘，又將維修主檔資料表及維修明細資料表經過轉換處理，整理出車輛維修主檔資料表及車輛維修時序明細資料表。車輛維修主檔資料表指的是 2008~2009 年間有過維修記錄的車輛基本資料，記錄著車號、車種、車齡等資訊；而車輛維修時序明細資料表則代表該車輛在這 2008~2009 年間每次維修零件的記錄及何時維修，並依照維修時間由遠至近進行排序，如表 3 所示。

## 4.2 車輛維修零件關聯探勘

本研究個案公司擁有 16 種左右的車種，但各車種間的車輛數並不平均，以中華 CANTAR 3.49 噸及 ISUZU 3.49 噸為大宗。本研究考量各類型車種間的零件的差異性及車齡不同會造成維修零件的關聯性不同，因此採用(1)區分車種車齡與(2)不區分車種車齡兩種資料預處理模式以比較並分析探勘結果，最後並請車輛專家依據探勘的結果進行判斷與選擇哪一種模式進行探勘，並將分析與規則回存至維修知識庫中。

### 4.2.1 區分車種車齡探勘

首先我們以車種做為區分，然後再依不同的車齡切割成 2 年內、3~6 年及 7 年以上的三個群體，分別以關聯規則進行維修零件間關係的探勘。本研究根據資料特性以最低支持度 15 次以上及最低信心度 41% 為標準，透過關聯規則找出不同車種，不同車齡的維修零件關聯規則，如圖 2 與圖 3 所示為中華 Canter 3.49 車種之結果。其中最顯著的規則為：

(離合器片，離合器壓板) → 離合器釋放軸承 *Confidence=98.6%*

表 3 車輛維修時序明細資料表內容

t200_t170_car_jd	fix_datetime	t125_name	T201_PRICE	T201_T208_SYSNUM	T201_MEMO
426-JT	200802280301002	冷氣高壓軟管 (壓縮機-散熱排)	1350.00	BEH2008010200002	4 分
426-JT	200802280301003	冷氣高壓軟管 (壓縮機-散熱排)	1150.00	BEH2008010200002	3 分
426-JT	200802280301004	冷煤 R134A	1300.00	BEH2008010200002	1 台份
3K965	200802290305118	機油	800.00	BEH2008010200001	
3K965	200802290305119	機油芯	350.00	BEH2008010200001	
3K965	200802290305120	空氣芯	900.00	BEH2008010200001	
7T-582	200802270302003	水溫感知器/ 溫度表開關	557.00	BEH2008010200002	
6T-181	200802270302004	左後煞車分泵	900.00	BEH2008010200002	左後前分泵
6T-181	200802270302005	左後煞車分泵	900.00	BEH2008010200002	左後後分泵

註: T200\_T170\_CAR\_ID 為車號; FIX\_DATETIME 為維修日期; T125\_NAME 為維修零件名稱; T201\_PRICE 為維修金額; T201\_T208\_SYSNUM 為維修行為; T201\_MEMO 為維修備註。

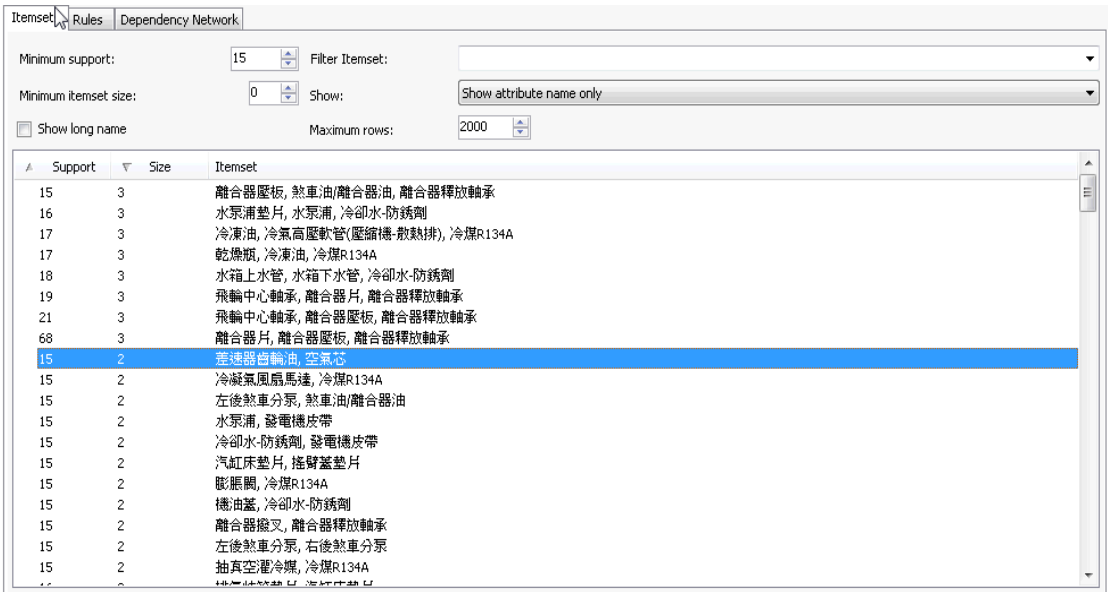


圖 2 部分通過最小支持度規則之示意圖 (中華 Canter 3.49 車種)

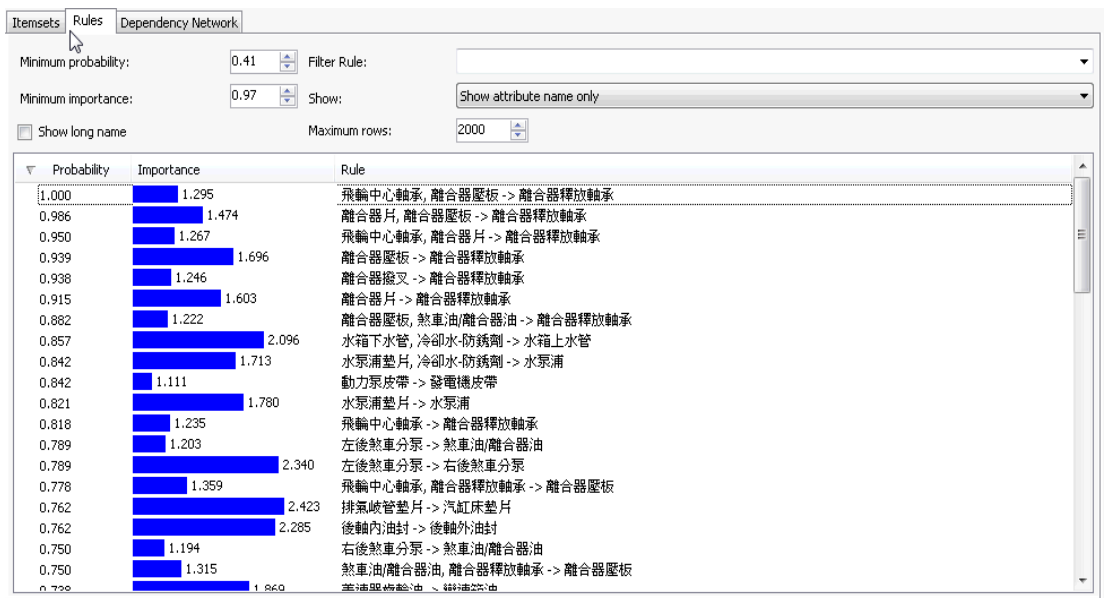


圖 3 部分通過最小信心度規則之示意圖 (中華 Canter 3.49 車種)

該規則代表當維修離合器片與離合器壓板時，通常也會維修離合器釋放軸承，且信心度高達 98.6%，重要度也有 1.474，是值得參考的資料。

### 4.2.2 整體車輛探勘

我們直接採用整體的維修資料，不針對車種及車齡做預先的資料切割，直接使用關聯法則做資料的探勘，只是需要將車種及車齡當做額外的輸入變數，一起納入資料探勘，我們以最低支持度 71 次以上及最小信心度 43% 為標準，結果如表 4 與圖 4 所示。其中最顯著的規則為：

(壓縮機皮帶/冷凍皮帶,車種=ISUZU 3.49 噸) → 動力泵皮帶

該規則代表的意思為當維修壓縮機皮帶/冷凍皮帶且車種是 ISUZU 3.49 噸時，通常也會維修動力泵皮帶也會一起維修，且信心度有 49%，且重要度也有 1.72。

表 4 關聯法則採用整體資料帶入車種及車齡變數探勘

Support	項目集合	Confidence	規則
0.01	壓縮機皮帶/冷凍皮帶, 車種= ISUZU 3.49 噸, 動力泵皮帶	0.49	壓縮機皮帶/冷凍皮帶, 車種= ISUZU 3.49 噸→動力泵皮帶
0.02	動力泵皮帶, 車齡=8 年以上, 發電機皮帶	0.64	動力泵皮帶, 車齡=8 年以上→發電機皮帶
0.01	離合器釋放軸承, 車種= 中華 CANTAR3.49, 離合器壓板	0.66	離合器釋放軸承, 車種=中華 CANTAR3.49 →離合器壓板
0.06	離合器釋放軸承, 車齡=4-7 年, 離合器壓板	0.71	離合器釋放軸承, 車齡=4-7 年→離合器壓板

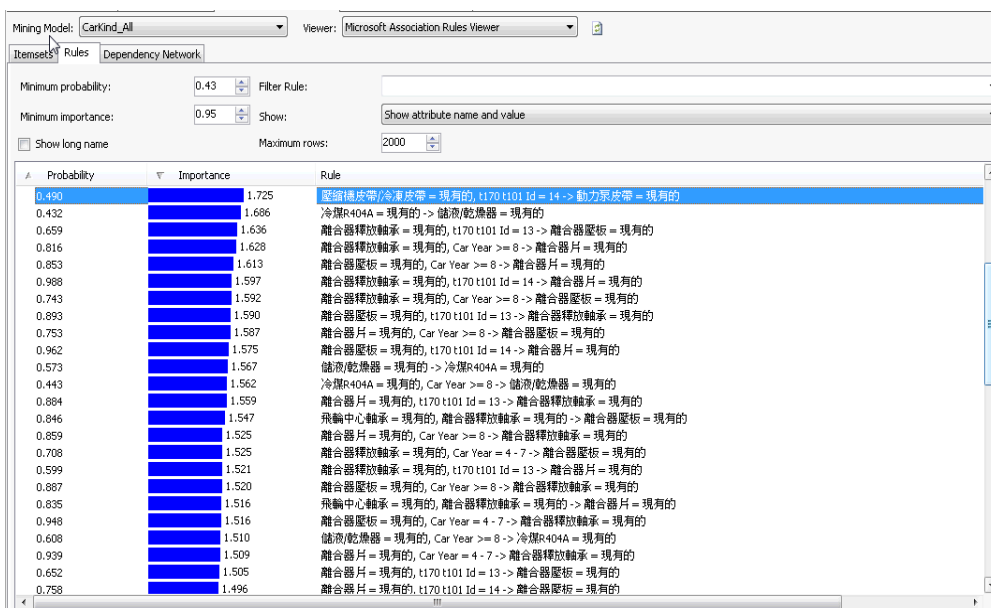


圖 4 車種車齡當作輸入變數的關聯規則挖掘

經過這兩種採用不同資料預處理方式的關聯法則資料探勘後，專家指出：(1)採用區分車種車齡探勘模式所挖掘出來的規則，是一些比較不易被發現的維修規則，(2)採用整體維修資料將車種及車齡納入輸入變數，發現的是通則性的關聯法則，且也比較偏向中華 CANTAR 3.49 噸及 ISUZU 3.49 噸這兩種車型，可能的原因是各車種間車輛數並不平均分佈。因此，本研究決定採用先將資料依不同車種及車齡預先切割的資料然後再進行關聯法則探勘的方式。

#### 4.2.3 車輛維修零件間的序列型樣探勘

本研究除了挖掘車輛零件間的關聯規則外並進一步考慮零件維修的時序關係，以期了解下次可能維修的零件，並考慮是否需做預防更換的動作。本研究比照關聯法則以車種及車齡做為區分，不同車種及車齡的序列型樣探勘的結果如表 5 所示。結果以中華 Canter 3.49 車種之先後順序規則數目較多且發生機率較高，ISUZU 3.49 車種則僅有少數維修順序規則；兩種車款之規則亦不相同，顯見區分車種與車齡分析有其必要性。

### 4.3 以決策樹建立危險車輛模型

本研究根據個案公司過去車輛進廠維修與報廢經驗以及在實務上對於安全與危險車輛的定義，採用維修金額高低標來判斷是否為危險車輛，若維修金額超出高標即認定為危險車輛，低於低標值則認定為安全車輛。研究以維修金額作為危險車輛判別的原因為(1)維修金額數據來源較意外事故的數據更完整與正確；(2)若維修金額超出高標表示有重大零件維修，其維修金額必定高於定期保養金額。研究透過先前關聯規則及序列型樣分析所得知結果，歸納分析出相關屬性變數（車輛零件）的重要度再加上車輛一些基本變數（如意外發生事故次數、車輛區域...等）。

表 5 不同車種顯著的維修零件序列機率表

先修零件	後修零件	機率
<b>車種：中華 Canter 3.49</b>		
離合器分泵	離合器釋放軸承	1.00
動力泵皮帶	冷氣皮帶	1.00
右前煞車分泵	冷氣皮帶	1.00
排檔繼動桿	儲液/乾燥器	1.00
儲液/乾燥器	冷媒 R404A	1.00
後輪內油封	前輪油封	1.00
前輪油封	軸承黃油	1.00
凸輪軸油封	軸承黃油	1.00
<b>車種：ISUZU 3.49</b>		
燃油管路及接頭	冷卻水-防銹劑	1.00
離合器壓板	離合器片	0.84
冷氣皮帶	發電機皮帶	0.83
方向機油	軸承黃油	0.61



本研究採用的決策樹演算法為 C4.5，由於決策樹產生的規則並非全部都是好的，所以會依據下列三種方法，衡量決策樹的有效程度：(1)進入該節點之資料數，設定為至少 10 筆；(2)若是葉部節點，將觀察資料分類的方式；(3)該節點將資料正確分類的比率，設定為至少 50%。為評估決策樹的分類效益，採用錯誤率與分類結果矩陣表呈現分類結果，對於不良的決策樹結果，將進行樹的深度、節點純度及節點記錄數的調整，重新訓練樹，直到萃取出有意義的規則。為了使建置模型的準確度具有一定的可信度，在資料準備階段，本研究將以 V 維交叉驗證 (V-fold cross-validation) 的方式，即先把資料隨機分成不相交的 V 份資料，將其中 V-1 份資料合起建立模型，把另外一份資料用來檢驗模型，得到模型的準確率，在以循序的方式進行，直到每一份資料均擔任過檢驗模型的資料集，而模型的準確率即為上面 V 次準確率的平均，本研究將以 V=4 來進行研究。先挑選維修最多的前兩名車種中華 CANTER 3.49 噸及 ISUZU 3.49 噸，分別建立不同的危險車輛模型，建立出來的危險車輛模型。

研究中危險車輛之定義是將各車種中維修金額超過高標金額的車輛挑選成危險車輛，再依照上述決策樹運算流程，我們先運用關聯法則及序列型樣技術挖掘出同車種但不同維修金額級距的車輛中具有高度關聯性的零件資料及車輛屬性資料，再透過專家挑選後，當成輸入變數，而是否為危險車輛為目標變數，建構本研究的決策樹危險車輛預警模型。研究進行決策樹建模的輸入變數挑選是經由專家透過先前將不同維修金額等級的車輛分別使用關聯法則與序列型樣挖掘出具有高度相關聯的零件資料，再配合車輛的外部屬性資料挑選而成，如：意外事故次數、車輛使用區域、車齡等。因此，研究之決策樹輸入變數為各項維修零件是否有維修及車輛屬性資料，而目標變數則為是否危險車輛，表 6 與圖 5 所示為建模危險車輛判別決策樹之輸入變數與模型。

#### 4.3.1 訓練模式結果與分析

本危險車輛預測模型採用混亂矩陣來驗證 (如 3.4 節評估方法所述)，矩陣的每一縱行代表一種目標分類，每一橫列代表一種推論的分類，矩陣的  $X_{ij}$  元素代表當屬於  $j$  種分類時，被推論是第  $i$  種的分類資料筆數。如此就能計算出各輸出類別的正確率，若正確率愈高，則代表分類結果良好，正確率公式如公式(3)所示。而危險車輛預測模型重點是要預測出危險車輛，所以危險車輛預測率愈高 ( $TN$ -rate)，則代表此模型正確率愈高，如公式(5)。此外研究並依據表 2 計算危險車輛預警得分 (Scoring)。

**訓練一：**訓練模式下得出此危險車輛決策樹預測模型的 *Accuracy* (正確率) 為 90.12%，該模型的預測危險車輛的  $TN$ -rate(危險車輛辨識率)值有 72.29%。此外，該模型的分數為 601 分  $((168 \times 1) + (60 \times 10) - (2 \times 3) - (23 \times 7) = 601)$ ，如表 7 所示。為了求得最佳的危險車輛預測模型，我們重新調整危險車輛判別的金額上限值，然後再重新建模，修正後的模型正確率有微幅提高，且  $TP$ -rate(正確辨識率)與  $TN$ -rate 也有提昇，重新建立模型後的資料如表 7 所示。

表 6 危險車輛建模輸入變數屬性示意說明

屬性代碼	屬性名稱	數值或類別
X1	意外事故次數	數值 0-沒發生過 1-三次內 2-三次以上
X2	車輛使用區域	數值 0-北 1-中 2-南 3-東 4-離島
X3	車齡	數值 0-1~2 年 1-3~6 年 2-7 年以上
X4	引擎本體系統是否維修	數值 0-沒有維修 1-有維修
X5	離合器釋放軸承是否維修	數值 0-沒有維修 1-有維修
X6	發電機是否維修	數值 0-沒有維修 1-有維修
X7	引擎點火系統是否維修	數值 0-沒有維修 1-有維修
X8	引擎進排氣系統是否維修	數值 0-沒有維修 1-有維修
X9	排檔繼動桿是否維修	數值 0-沒有維修 1-有維修
X10	引擎冷卻系統是否維修	數值 0-沒有維修 1-有維修
X11	底盤油品是否維修	數值 0-沒有維修 1-有維修
X12	底盤煞車系統是否維修	數值 0-沒有維修 1-有維修
Y1	是否為危險車輛	數值 1-危險車輛 2-安全車輛

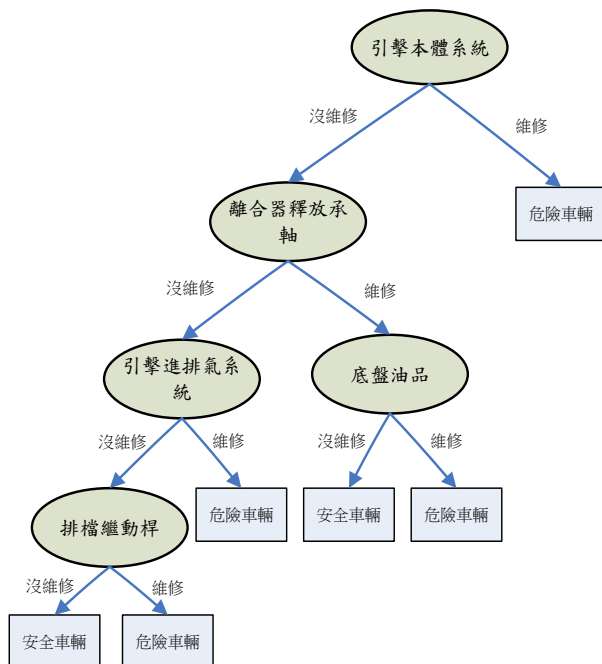


圖 5 決策樹建模示意圖

表 7 訓練模式混亂矩陣結果(一)

	安全車輛(預測)	危險車輛(預測)	
安全車輛(實際)	168	2	<i>TP-rate: 98.83%</i>
危險車輛(實際)	23	60	<i>TN-rate: 72.29%</i>
			<i>Accuracy:90.12%</i>

表 8 訓練模式混亂矩陣結果(二)

	安全車輛(預測)	危險車輛(預測)	
安全車輛(實際)	170	0	<i>TP-rate:100%</i>
危險車輛(實際)	19	58	<i>TN-rate: 75.32%</i>
			<i>Accuracy:92.31%</i>

表 9 訓練模式混亂矩陣結果(三)

	安全車輛(預測)	危險車輛(預測)	
安全車輛(實際)	169	1	<i>TP-rate: 99.41%</i>
危險車輛(實際)	6	56	<i>TN-rate:90.32%</i>
			<i>Accuracy:96.98%</i>

**訓練二：**由表 8 觀之，該模型的評分的分數為 617 分且第二次的模型 *TN-rate* (危險車輛辨識率) 有提昇。該模型與第一次建模的模型相比較，分數雖然沒有差異很大，但第二次的模型比第一次的模型在各方面預測率都有提升。為了達到我們預期的效果，我們再重新調整危險車輛的定義，將危險車輛的維修金額判斷門檻往上微調，已進行第三次的預測模型建立。

**訓練三：**由表 9 觀之，*TN-rate* (危險車輛辨識率) 已大幅提昇到 90.32%、整體的 *Accuracy* (正確率) 提昇至 96.98% 且該模型的評分的分數亦提升為 684 分。經過反覆的的建模，以此次的分數為最高，我們最後決定採用此訓練模型。

#### 4.3.2 測試模式實驗結果與分析

本研究採用 V 維交叉驗證 (*V-fold cross-validation*) 的方式，先把資料隨機分成沒有交集的 V 份資料，將其中 V-1 份資料當作訓練集合以建立模型，把剩餘的一份當作用來檢驗模型的測試資料；研究以循序的方式進行，直到每一份資料均擔任過檢驗模型的測試資料集。本研究將以 V=4 來針對第三次危險車輛的所建立之模型進行驗證，驗證的結果同樣採用 3.4 小節的評估方式以評估測試結果的表現，結果如表 10 所示。

**實驗結果：**實驗結果如表 10 所示，最佳測試案例 (*case*) 的總得分為 175 分，且此預測模型的整體正確率為 98.25%，*TP-rate* (正確辨識安全車輛) 能夠達到 100% 且 *TN-Rate* (正確辨識危險車輛) 也可以達到 93.33%，符合使用者的目標能成功且正確的預測出危險車輛。最差測試 *case* 的總得分有 164 分 (該表現值與最佳案例分數相覷不遠)，此預測模型的整體正確率亦可達到 9 成 (94.83%)，*TP-rate* (正確辨識安全車輛) 能夠達到 97.62%，*TN-Rate* (正確辨識危險車輛) 則略低為 87.50%。經過多次採用不同的測試資料驗證後，我們可以得出平均的結果數據如表 10 最後一欄所示，研究結果顯示該模型可得到與訓練模式最佳狀況相同的結果。我們發現各項數據均能達到九成以上，其中安全車輛預測為將近 100%。該結果顯示所建立得模型將有助於公司避免浪費資源維修安全車輛且也可成功預測危險車輛以進行預防性維修。

**零件維修規則歸納：**研究透過決策樹建立模型後共整理出下列七個主要規則：(1) **規則一：**當「汽缸頭總成」有維修，有 87.45% 的機率可能是危險車輛；(2) **規則二：**當維修「壓縮機」但沒有維修「汽缸頭總成」零件，有 91.4% 的機率可能是危險車輛；(3) **規則三：**若有維修「變速箱總成」，但沒有「維修汽缸頭總成」、「壓縮機」及「曲軸軸承」零件，有 98.39% 的機率是一般

表 10 測試階段之實驗結果

	<i>TP-rate</i>	<i>TN-rate</i>	<i>Accuracy</i>
最佳測試結果(best case)	100.00%	93.33%	98.25%
最差測試結果 (worst case)	97.62%	87.50%	94.83%
平均測試結果(average)	99.41%	90.32%	96.98%

車輛；(4) **規則四**：若維修「後軸內油封」及「前輪油封」，但沒有維修「汽缸頭總成」、「壓縮機」、「變速箱總成」及「沒有發生意外事故」，有 90.45%的機率是安全車輛；(5) **規則五**：若沒有維修「汽缸頭總成」、「壓縮機」、「變速箱總成」、「後軸內油封」、「冷氣壓縮機」、「風扇離合器」及「橫拉桿和尚頭右」零件，有 97.51%的機率是安全車輛；(6) **規則六**：有修「冷氣壓縮機」，但沒有修「汽缸頭總成」、「壓縮機」、「變速箱總成」及「後軸內油封」，有 74.33%的機率是危險車輛；(7) **規則七**：若有修「後內油封」，但沒有維修「汽缸頭總成」、「壓縮機」、「變速箱總成」及「前輪油封」，有 99.13%的機率是危險車輛。決策樹預測零件維修規則與危險車輛機率如表 11 與圖 6 所示。

#### 4.1 系統展示

透過後端機制，將所有的零件維修關聯資料挖掘出來和運用決策樹建立危險車輛模型後，車輛管理人員便可以透過瀏覽器進入車輛維修及危險車輛預測系統，只需要在車輛維修作業中，輸入車輛的車號、維修里程、維修日期等基本資料後，再輸入車輛維修零件後，系統便會透過維修知識庫，取得此次維修通常會伴隨的零件及下次可能維修的零件資訊，另外在整筆工單輸入完成後，系統也會比對危險車輛模型警示車輛管理人員，該台車輛是否類似危險車輛，

表 11 決策樹由零件維修規則推測危險車輛機率

規則	節點路徑	輸出屬性	case 數	機率
規則一	汽缸頭總成 = '1'	危險車輛	23	87.45%
		安全車輛	0	6.27%
規則二	汽缸頭總成 = '0' 及 壓縮機 = '1'	危險車輛	12	91.40%
		安全車輛	0	4.30%
規則三	汽缸頭總成 = '0' 及 壓縮機 = '0' 及 變速箱總成 = '1' 及 曲軸軸承 = '0'	危險車輛	8	98.39%
		安全車輛	0	0.86%
規則四	汽缸頭總成 = '0' 及 壓縮機 = '0' 及 變速箱總成 = '0' 及 後軸內油封 = '1' 及 前輪油封 = '1' 及 意外事故次數 = '0'	危險車輛	2	15.41%
		安全車輛	11	84.55%
規則五	汽缸頭總成 = '0' 及 壓縮機 = '0' 及 變速箱總成 = '0' 及 後軸內油封 = '0' 及 冷氣壓縮機 = '0' 及 風扇離合器 = '0' 及 橫拉桿和尚頭右 = '0'	危險車輛	4	2.49%
		安全車輛	157	97.51%
規則六	汽缸頭總成 = '0' 及 壓縮機 = '0' 及 變速箱總成 = '0' 及 後軸內油封 = '0' 及 冷氣壓縮機 = '1'	危險車輛	3	74.33%
		安全車輛	1	25.13%
規則七	汽缸頭總成 = '0' 及 壓縮機 = '0' 及 變速箱總成 = '0' 及 後軸內油封 = '1' 及 前輪油封 = '0'	危險車輛	5	99.13%
		安全車輛	0	0.43%

註: 0 表示沒有維修, 1 表示有維修。

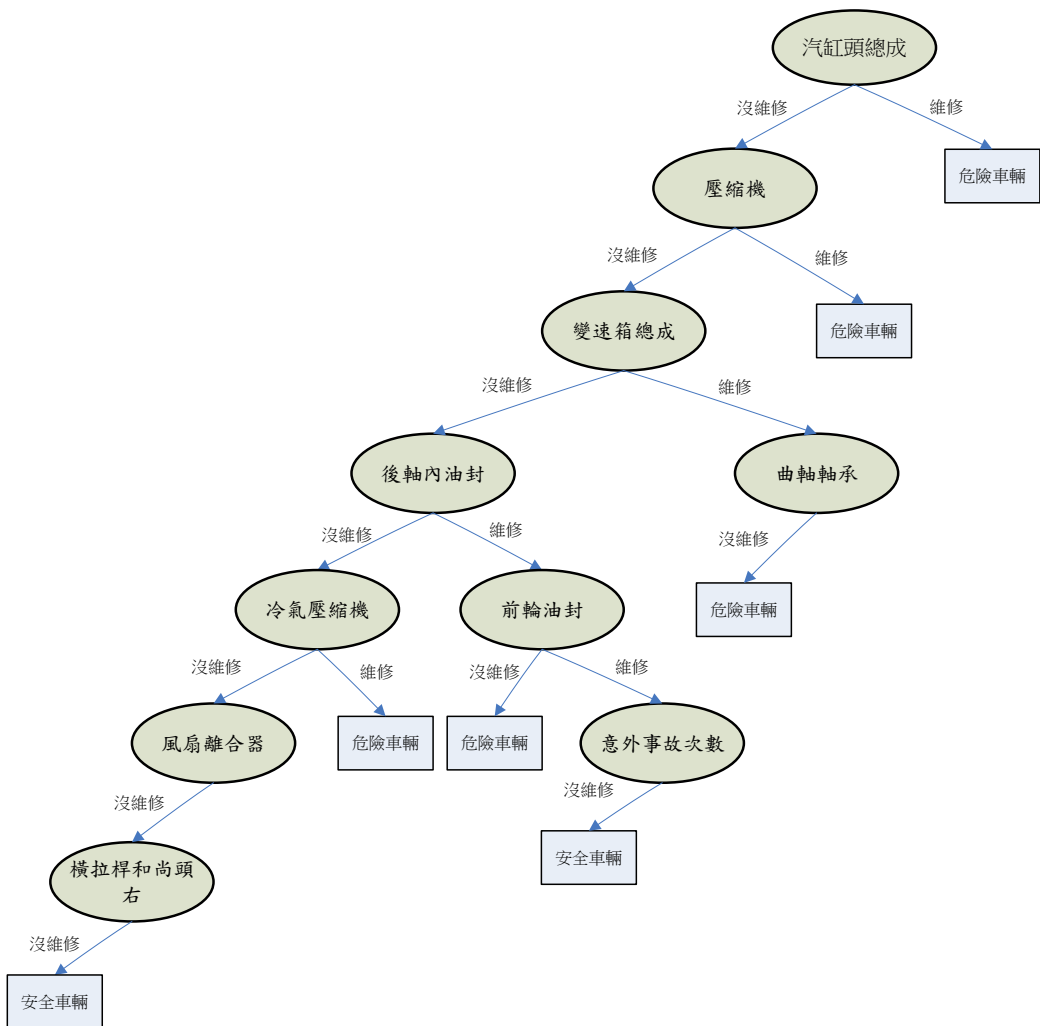


圖 6 零件維修規則推測危險車輛之決策樹示意圖

流程如圖 7 所示。本系統車輛維修的基本資料輸入模式如圖 8 所示，各級的车辆管理人員需輸入該次維修時車輛的基本資料，有車號、進廠維修里程、進廠維修日期等必要資訊。接著車輛管理人員將輸入正確的車輛維修零件基本資料，必要輸入的資訊包括：零件資料、價格及維修行為，車輛維修零件輸入模式如圖 9 所示。最後，前端系統會主動的回後端車輛維修知識庫抓取此次維修零件相關的預測資訊，其包括：本次通常會伴隨維修哪些零件，或是下次可能會維修哪些零件，並且依照此次維修的零件項目及過去維修零件項目，放入危險車輛模型中去預測，將預測結果呈現在畫面上，如圖 10 所示。

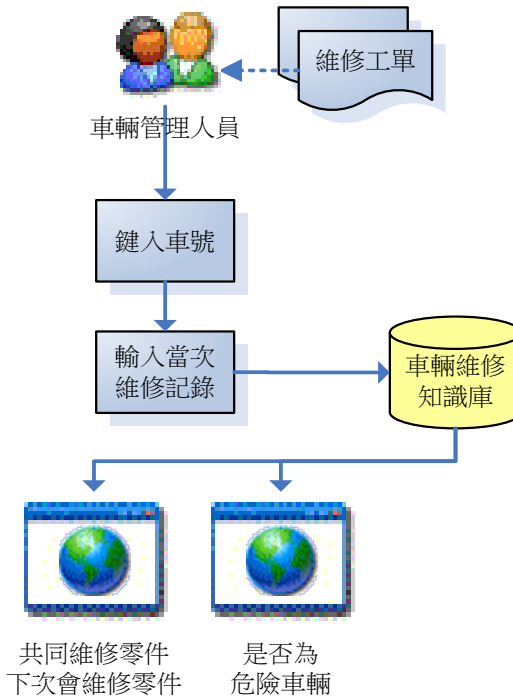


圖 7 前端系統資料輸入流程

**輸入維修工單** ■ 橘色欄位為必須資料，請正確填寫。

維修三聯單單號	8888888	(假設尚未取得新版維修三聯單，單號請用車號+年月日+兩碼流水號)	
工單號碼	8888888	進廠維修日期	2008/11/30
維修類型	<input checked="" type="checkbox"/> 一般維修 <input type="checkbox"/> 定期保養(可複選，如果你勾了定保那重新計算定保週期)	維修完工日期	2008/12/10
車號	6T-117	ISUZU 3.49噸(三期)	所屬單位 1010-北一區營業部
冷凍機	欣融Cold King		車廂 新三興
進廠里程數或小時數 (小時數是堆高機才用)	666666	(公里)/(小時)	<input checked="" type="radio"/> 未含稅 <input type="radio"/> 含稅價 工單稅額 注意:收據請選擇未稅!!
付款對象	<input checked="" type="radio"/> 廠商 80303054 長源汽車股份有限公司台北分公司 <input type="radio"/> 員工 admin-admin		

圖 8 車輛基本資料輸入畫面

零件大分類	零件中分類	零件	價格	折扣價	行為	備註	說明
引擎	燃料, 點火系統	010316 -燃油管路及接頭	0	0	一般週期性定期保養		清除
引擎	冷卻系統	010415 -冷卻水-防銹劑	0	0	一般週期性定期保養		清除
引擎	冷卻系統	010416 -正副水箱及冷卻水管路,接頭	0	0	一般週期性定期保養		清除
底盤	油品相關類	020104 -煞車油/離合器油	0	0	一般週期性定期保養		清除
電器	冷氣系統	030122 -冷氣皮帶	0	0	一般週期性定期保養		清除
電器	電氣系統	030224 -發電機皮帶	0	0	一般週期性定期保養		清除

上傳維修記錄照片

小計: 0

實際應付總金額: 0 元

新增完成 維修零件預防更換及危險車輛預測

圖 9 維修零件輸入畫面

零件大分類	零件中分類	零件	相伴隨維修零件	下次可能會更換零件
引擎	燃料, 點火系統	010316 -燃油管路及接頭	離合器釋放軸承	儲液/乾燥器
引擎	冷卻系統	010415 -冷卻水-防銹劑	離合器壓板	前輪油封
引擎	冷卻系統	010416 -正副水箱及冷卻水管路,接頭	離合器釋放軸承	冷氣皮帶
底盤	油品相關	020104 -煞車油/離合器油	離合器釋放軸承	冷媒R404A
電器	冷氣系統	030122 -冷氣皮帶	離合器釋放軸承	離合器釋放軸承
電器	電氣系統	030224 -發電機皮帶	冷氣皮帶	冷氣皮帶

可能是危險車輛!!

圖 10 預測及建議結果畫面

## 5. 結論與未來展望

物流宅配業者近來因為宅經濟的崛起，也連帶造成業績的快速成長，但業績成長的同時，物流宅配業者之配送車輛維修成本也隨之攀升。在本論文中，結合了實際企業中車輛維修、危險車輛資料及資料探勘及建模技術與理論，提出了一套智慧型車輛維修及預警決策系統。本決策支援系統的研究方向即利用資料探勘的技術將隱藏於車輛維修資料庫中有用的資訊進行探勘，建立車輛的維修知識庫與建立危險車輛預測模型。本研究提出之模式與程序除有助於物流司機或是車輛管理人員獲得車輛基本的維修知識，也能針對維修車廠提出適當的維修建議，進



而減少後續維修成本的支出。另一方面，系統透過決策樹而建立的危險車輛預警功能也提供了適度的示警及參考數據，能協助總部的車輛管理人員作出車輛是否要繼續維修、報廢或剩餘價值處份的決策，進而降低車輛維修的費用。整體而言，研究結果可以協助 T 公司的車輛管理人員獲得基本的維修知識，補足原先資訊不足的困擾，另外可以事先針對危險車輛示警與安全車輛預測，節省後續維修成本。

後續研究將(1)持續蒐集資料以建立更多車種的的預測模型與可調適之決策系統；(2)在零件維修關連規則探勘部分，將探討以不同車輛屬性為依據的關聯規則探勘，如：品牌、車種、車齡與里程數，與專家所建議之規則的差異，以進行零件維修建議之調整與相關驗證；(3)根據相關研究並結合本應用領域之資料特性改良研究預測模型，並比較不同模型，例如：邏輯斯迴歸 (logistic regression) 與支援向量機 (support vector machine) 等，預測結果差異，建立正確率更高的預測模型並將找出更有效的知識法則與(4)預期實際導入系統後將進行上線後的預測準確率評估以提供模型修正之依據；此外，將評估系統對個案公司節省維修成本的效益以發展適性化決策支援系統。

## 參考文獻

- 王志堅，「架構式商業智慧系統需求的分析模型研究」，中山大學資訊管理研究所未出版碩士論文，民國 97 年。
- 宋建達，「台灣宅配業未來發展趨勢研究」，網路社會學通訊，第七十三期，民國 97 年。
- 林建煌，「台灣宅配業經營型態與經營作為之分析」，中央大學高階主管企業管理研究所未出版碩士論文，民國 96 年。
- 張紂微、潘宜龍、黃錦祥，製造業電子化教戰手冊 e-Business-電機電子產業，臺北市：工業局，民國 90 年。
- 舒毓筌，「在醫學報告資料中探勘連續序列」，東華大學資訊工程研究所未出版碩士論文，民國 93 年。
- 楊玟欣，「透視台灣五大宅配商 購物通路又多又新，宅配業乘勢崛起」，財訊月刊，第三百期，民國 96 年，250-253 頁。
- 郭家齊，「整合資料探勘與案例式推論與機台故障診斷維護系統之研究」，雲林科技大學工業工程與管理系未出版碩士論文，民國 93 年。
- 陳星壁，「以關聯法則探勘為基礎之電路板件維修輔助系統」，中華大學資訊工程學系未出版碩士論文，民國 95 年。
- 廖介銘，「決策樹應用於糖尿病之探勘」，華梵大學資訊管理學系未出版碩士論文，民國 92 年。

蔡有藤，「系統預防維護作業之研究」，國立中央大學機械工程研究所未出版博士論文，民國 88 年。

蔡燕純，「應用馬可夫預防維護模型於維修保養策略之探討」，國立中央大學機械工程研究所未出版碩士論文，民國 93 年。

謝智聿，「以序列探勘理論實作之網路異常型樣驗證」，私立逢甲大學資訊工程研究所未出版碩士論文，民國 91 年。

Agrawal, R., Imielinski, T., and Swami, A., "Mining Association Rules between Sets of Items in Large Databases," In *Proceedings of the ACM SIGMOD International Conference on Management of Data*, Washington, D.C.: ACM Digital Library, 1993, pp. 207-216.

Agrawal, R. and Srikant, R., "Fast Algorithms for Mining Association Rules," In *Proceedings of the 20th International Conference on Very Large Data Bases*, Santiago de Chile, 1994, pp. 487-499.

Agrawal R. and Srikant R., "Mining Sequential Patterns," In *Proceedings of the International Conference Data Engineering*, Taipei, Taiwan: IEEE Computer Society Press, 1995, pp. 3-14.

Buddhakulsomsiri, J. and Zakaria, A., "Sequential Pattern Mining Algorithm for Automotive Warranty Data," *Computers & Industrial Engineering*, Vol. 57, No.1, 2009, pp.137-147.

Chen, M. S., Han, J., and Yu, P. S., "Data Mining: An Overview from a Database Perspective," *IEEE Transactions on Knowledge and Data Engineering*, Vol. 8, No.6, 1996, pp. 866-883.

Clifton, R. H., *Principles of Planned Maintenance*, London: Edward Arnold, 1974.

Edwards, D. J., Holt, G. D., and Harris, F. C., "Predictive Maintenance Techniques and their Relevance to Construction Plant," *Journal of Quality in Maintenance Engineering*, Vol. 4, No.1, 1998, pp.25-37.

Han, J. and Kamber, M., *Data Mining: Concepts and Techniques*, 2nd ed., San Francisco: Morgan Kaufmann, 2006.

Last, M., Sinaiski, A., and Subramania, H. S., "Predictive Maintenance with Multi-target Classification Models," In *Proceedings of the Second International Conference on Intelligent Information and Database Systems*, Hue City, Vietnam. Nguyen, N.-T., Le, M. T., Swiatek, J. (Eds.), *Springer-Verlag, Lecture Notes in Artificial Intelligence*, Vol. 5991, 2010, pp.368-377.

Lin, F. R., Chou, S. C., Pan, S. M., and Chen, Y. M. "Mining Time Dependency Patterns in Clinical Pathways," *International Journal of Medical Informatics*, Vol. 62, No.1, 2001, pp.11-25.

Nemati, H. R., Steiger, D. M., Iyer, L. S., and Herschel, R. T., "Knowledge Warehouse: An Architectural Integration of Knowledge Management, Decision Support, Artificial Intelligence and Data Warehousing," *Decision Support System*, Vol. 33, No. 2, 2002, pp.143-161.

- Smith, A. E., Coit, D. W., and Liang, Y. C., "Neural Network Models to Anticipate Failures of Airport Ground Transportation Vehicle Doors," *IEEE Transactions on Automation Science and Engineering (TASE)*, Vol.7, No. 1, 2010, pp.183-188.
- Srikant, R. and Agrawal, R., "Mining Quantitative Association Rules in Large Relational Tables," In *Proceedings of the ACM-SIGMOD International Conferences Management of Data*, Montreal, Canada: ACM Digital Library, 1996, pp.1-12.
- Turban, E., Aronson, J. E., and Liang, T. P., *Decision Support Systems and Intelligent Systems*, 7th ed., New Jersey: Prentice Hall, 2005.
- Yam, R. C. M., Tse, P. W., Li, L., and Tu, P., "Intelligent Predictive Decision Support System for Condition-based Maintenance," *International Journal of Advanced Manufacturing Technology*, Vol.17, No. 5, 2001, pp. 383-391.