

壹、緒論

資料探勘融合了多個包括統計、機器學習及人工智慧之領域 (Zhan & Luan, 2006)，亦被視為是知識管理的新技術，不僅是單純的統計分析，更是資料蒐集與處理、問題識別、統計／預測模式及輸出顯示的過程，以利進行配置或決策 (Streifer & Schumann, 2005)。此外，資料探勘更可用於支援決策制訂、預測及估計的模式識別和獲取可利用資訊 (Chang, 2006)，亦揭露「隱藏訊息」、類別及大量數據資料來預測行為和成果之過程 (Luan, 2002)，資料探勘為一跨領域之資料分析方式，可提供更有效的資料檢索和知識來源 (Zhan & Luan, 2006)。

雖然資料探勘技術已廣泛且有效地應用於商業界，在教育上，它仍是相對較新的使用方式 (Chang, 2006)，然而，一些教育的相關研究已經著手進行，如〈資料探勘於高等教育的應用〉 (Luan, 2002)、〈以資料探勘技術檢查學生本位及教師本位課程之線上學習模式〉 (Hung & Crook, 2009)、〈以資料探勘預測學業表現〉 (Vandamme, Meskens, & Superby, 2007)、〈以資料探勘探討學生使用先進技術縮短畢業時間之學習型態〉 (Eykamp, 2006)、〈資料探勘於心理教育研究之貢獻：自組織型態分析闡明資優生福利〉 (Thuneberg & Hotulainen, 2006)、〈以資料探勘建立學生參與的組織類型〉 (Luan, Zhao, & Hayek, 2009)、〈擴充小規模私立學院之研究：使用資料探勘之招生管理案例研究〉 (Antons & Maltz, 2006)，以及〈資料探勘雜誌及書籍：以網路科學發掘教育研究社群之結構〉 (Carolan & Natriello, 2005)。

然而，資料探勘卻較少應用於教育行政，根據Luan與Zhao (2006)所述，與傳統實驗原理不同的是，資料探勘於粒狀數據記錄中尋找具體且獨特的模式，它為最小的分析單位，故資料探勘為數據分析中之新技術，主要用於計算和資料儲存。

與傳統統計技術相比，資料探勘以搜尋為主，因此不需設定初步假設，僅需使用數據資料找出其中之模式及關係 (Chrysostomou, Chen, & Liu, 2009)。Pechenizkiy、Puuronen與Tsymbol (2008)曾提出現有資料探勘架構的全面探討，並將之分為三種類型：理論導向、程序導向及基礎導向，而

本研究因專注於研究教育行政，故採用上述程序導向之架構。

貳、資料探勘技術的方法與步驟

資料探勘通常應用於商業性質，其概念非常簡單：因已積累相當之數據，故無論其價值為何，值得嘗試存取，而資料探勘之最終目標為發掘知識，此類知識通常以生成之模式，如規則集、決策樹或回歸公式之形式展現（Charest, Delisle, Cervantes, & Shen, 2008），使探勘之數據得以應用於達成研究目的（Van Horn, 1998）。然而，也有人認為，其可加快在各領域科學數據的採集，並形成一個幾乎無法逾越的數據崩落（Borne, Becla, Davidson, Szalay, & Tyson, 2008）。為了增加對現存教育行政研究的了解，本研究從兩個論文資料庫中進行資料探勘，一個位於臺灣；另一個則位於北美，本研究採王如哲（1998）的教育行政主題分類，共分為11個主題：教育組織、教育計畫、教育決策、教育領導、教育溝通、教育制度、教育財政、教育法令、教育政策、教育視導及教育評鑑，共統計1920～2009年此二資料庫之論文。

參、資料探勘技術在教育行政的應用

資料探勘過程包括資料選擇、資料預先處理、資料轉換、模式搜尋、解釋及評定，而資料之提取和驗證模式則是此過程之核心（Pechenizkiy et al., 2008），本研究於中文資料庫使用臺灣博、碩士論文知識加值系統，其包含所有於臺灣完成之博、碩士學位論文，另外，在英文資料庫使用ProQuest論文資料庫，其包含所有於北美以英文撰寫的博、碩士論文。

一、用論文名稱之資料探勘結果

表1及表2顯示國內外博、碩士論文名稱之資料探勘結果，圖1顯示國內資料庫中教育行政11個論文主題類別分布情形，分別為：「教育政策」（51.6%）出現頻率最高，其次為「教育制度」（13.7%）及「教育計畫」（12.6%）；而國外資料庫中，出現頻率最高者為「教育政策」（26.5%），其次為「教育制度」（25.8%）、「教育組織」（16.0%）及教