

## 顧客流失分析模式之個案研究—以臺灣 H 健康休閒俱樂部為例

陳麒文

國立體育學院

### 摘要

本研究之主要目的包括瞭解臺灣某家健康休閒俱樂部（以下均以 H club 稱之）之顧客組成結構、利用資料採礦分類技術（鑑別分析、羅吉斯迴歸、人工類神經網路、多元適應性雲形迴歸）來建立 H club 之顧客流失分析模式、與經由顧客流失分析模式來瞭解流失顧客之重要特徵。本研究利用 H club 所提供的資料來進行實證研究，共計 1,287 筆，在剔除資料不全與資料內容不合理的資料後，共有 1,152 筆。本研究之結果如下：

- 一、H club 之顧客組成結構以女性會員（61.72%）佔大多數、平均年齡為 38.16 歲且多集中在 25 歲～45 歲（69.03%）、平均會齡為 2.63 年且多集中在二年以下（52.86%）、入會金額以 0 元為最多（50.52%）、月費金額以 2,500 元為最多（36.37%）、居住地區則以臺北市居多（81.68%）、付款方式則多為現金（59.46%）。
- 二、本研究所提出之顧客流失分析模式之建構程序，主要的目的是希望透過四種不同分類方法的比較，來求得一個最佳的區別模式。此外，為驗證所提模式之有效性，本研究利用 H club 所提供的資料來進行實證研究，結果顯示，多元適應性雲形迴歸的整體分類績效為 86.52%，具有極佳之分類效果，是一項值得建議使用的工具。
- 三、總體而言，以整體正確判別率為最高之多元適應性雲形迴歸所建構之顧客流失分析模式，其流失顧客之特徵為年齡介於 30～35 歲、會齡為一年以下、月繳 2,500 元的月費、且以現金支付費用的會員。

**關鍵詞：**資料採礦、顧客流失、鑑別分析、羅吉斯迴歸、人工類神經網路、多元適應性雲形迴歸





## **A Case Study of Customer Churn Model—An Example of H Health Club in Taiwan**

Chi-Wen Chen

National College of Physical Education and Sports

### **Abstract**

The main purpose of this study was to understand the customer structure in a certain Taiwanese health club (the following substituted as H club) and to establish its customer churn model through data mining classification technology (discriminant analysis, logistic regression, artificial neural networks, multivariate adaptive regression splines) as well as to realize the significant characteristics of churn customer through the customer churn model. The data used was provided by H club to perform the empirical research, and the original data were 1,287 records. After eliminating the incomplete data and the unreasonable data, there were 1,152 records totally. The results were as followed:

1. In the customer structure of H club, the majority of gender was female member taken 61.72 percent of total. While the average age was 38.16-year-old, the range of age was from 25-year-old to 45-year-old taken 69.03 percent. The duration time of membership was 2.63-year in average and most centralized in two-year-below (52.86%). The initiation fee was 0 New Taiwanese dollar taken the major part of total (50.52%). Otherwise, the monthly fee was 2,500 New Taiwanese dollars in majority (36.37%). With the residential area, the most members were living in Taipei city (81.68%). Finally, the method of payment was in cash (59.46%).
2. The constructional process of churn model in this study was by way of four classification methods to obtain the one best discriminating mode. Besides, in order to verify the effectivity of the discriminating mode, this study used the data provided by H club to perform the empirical research. The result found that whole





correct classification rate was 86.52% in the churn model using the multivariate adaptive regression splines, which had the best utility and been suggested as a worth tool to use.

3. By all accounts, the characteristics of churn customer through the customer churn model using the best way of multivariate adaptive regression were the members of the age between 30-year-old and 35-year-old, the duration time of membership of one-year-below, the monthly fee of 2,500 New Taiwanese dollars, and the method of payment of cash.

**Keywords: Data Mining, Customer Churn, Discriminant Analysis, Logistic Regression, Artificial Neural Networks, Multivariate Adaptive Regression Splines**





## 壹、緒論

### 一、研究背景

隨著資訊科技 (Information Technology, IT) 的腳步不斷地躍進, 利用其技術的導入來輔助健身俱樂部之會員相關業務的進行, 早已行之有年。相關的技術, 不外乎會員資料的建立、管理, 與財務報表的製作、列印, 然而這些都只是第一步驟; 要如何從中瞭解顧客的需求及特性, 藉以提高健身俱樂部的利潤及降低顧客流失率, 並進一步做好顧客關係管理, 才是未來努力的目標。

國內健身俱樂部的發展, 由有氧運動的先鋒推動者—姜慧嵐女士於 1986 年開啓了此一產業的戰國風雲; 此後, 唐雅君女士於 1991 年起積極在臺灣全省各地設立亞力山大據點; 到了 2000 年, 以「健身結合娛樂」為經營概念的加州健身中心, 成功地切入目標市場為 25 至 35 歲的年輕族群; 同時, 由國內耐斯企業和美國金牌俱樂部合資的金牌健身俱樂部 (Gold's Gym), 也吸引了不少上班族。健身俱樂部從以往多位於郊區、且費用昂貴, 到如今位在市中心、交通方便, 會費大多下降, 月繳 1000~3000 元不等的清潔費, 便可以享有多種運動設施與服務。

消費者在面對低價與多種選擇的情況下, 健身俱樂部的業者往往面臨了經營上的一大窘境—高顧客流失率 (High Churn Rate)。而國內在「1999 健康體能與有氧運動大會」中, 相關業者也提出健身俱樂部目前所面臨的一大窘境是—會員流失率高, 與美國健身俱樂部會員的流失率高達 54% (Stotlar, 1999) 的情況相類似。此外, 也有許多的研究報告指出, 開發新顧客的成本遠比保留舊顧客的成本高出 3 到 10 倍以上 (羅家德, 2001; Peppers & Rogers, 1993; Walters, 1994; Newell, 2000)。因此, 如何制定有效的企業策略並進一步地維持顧客人數以減少開發新顧客的成本, 實考驗著企業主的智慧。

### 二、研究動機

傳統上利用市場調查與問卷方式來瞭解顧客滿意度的方式, 可能會因為抽樣方式的不佳、受試者填答的真實性...等等, 而導致取樣誤差或分析偏差, 進而影響結果的準確性; 但是, 從現有企業的資料庫中利用資料採礦的技術來進行資料分析, 在許多的研究中都證明了可得到較佳的結果, 此為本研究動機之一。





所以，本研究將針對健身俱樂部所能提供之顧客資料庫做一深入的分析，期望能利用資料採礦技術中的鑑別分析(Discriminant Analysis)、羅吉斯迴歸(Logistic Regression)、人工類神經網路(Artificial Neural Networks, ANNs)與多元適應性雲形迴歸(Multivariate Adaptive Regression Splines, MARS)方法來對顧客進行分類，並評估此模型的適當性，此為本研究動機之二。

### 三、研究目的與問題

本研究的主要目的在於建立與評估臺灣某家健身俱樂部（以下均以 H club 稱之）之顧客流失模型；而主要的研究問題，除了瞭解 H club 之顧客組成結構外，同時也著重在如何建立 H club 之顧客流失模型？並希望能由 H club 所提供之顧客資料庫，挖掘出流失顧客的特性。

### 四、研究範圍與限制

本研究係以 H club 之全部會員為研究範圍。而由於本研究的顧客流失模型是針對此家健身俱樂部而設計，因此只能作為該家健身俱樂部來評估顧客流失的依據，但是其它各家健身俱樂部可利用本研究之類似觀念，採用相同的研究程序來建立屬於各家健身俱樂部的顧客流失模型，此為本研究之限制。

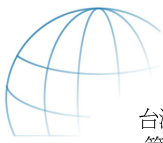
## 貳、文獻探討

### 一、顧客流失

根據各行業平均統計值顯示，每年約有 85%的既有顧客會留下；相對地，每年會有 15%的顧客流失，形成一個旋轉門理論，有進有出（夏載，2001）。而從成本考量的觀點來看，挽留一個舊顧客的成本幾乎只有開發新顧客的十分之一，這驗證了顧客關係管理的名言—「若能將顧客流失率減少 5%，利潤將會有 100%的成長」（數博網，2001）。問題是，如何挽留即將流失的舊顧客？這就必須透過資料採礦技術，從以往顧客流失紀錄中找出共同的規律，來預測現有顧客流失的意向了。

針對上述顧客流失(Churn)或顧客保留(Customer Retention)問題，實務上一般的作法是根據顧客的人口統計變數(Demographic Variables)及相關的資訊將





顧客分為「會流失」及「不會流失」兩大類；也有將顧客分為「低度流失」、「中度流失」、與「高度流失」三大類，因此，顧客流失問題屬於資料採礦中廣義的分類問題。但是，在實際應用的層面上，資料採礦能夠給予每個顧客一個流失評等積分（Churn Score），透過這組分數，企業內的行銷人員便能夠輕易掌握未來半年內、一季內、甚至是一個月內可能流失的顧客名單（數博網，2001）。

## 二、資料採礦

由突尼西亞人 Usama Fayyad 所發現的 Pattern 辨識演算法，不但成了他 1991 年論文的主題，也衍生出後來資料採礦的發展。資料採礦最早是被應用在天文學上，藉由機器學習、類型辨識、與統計等技術，在短短 4 小時內所發現的行星勝過 20 多位天文學家 4 年的研究成果（謝邦昌，2001）。而美國麻省理工學院（Massachusetts Institute of Technology, MIT）在 2001 年初的《科技評論雜誌（Technology Review）》中，更將資料採礦評選為 21 世紀影響人類生活最大的 10 項新興科技之一（表一）。

表一 科技評論雜誌預測未來影響人類之十大新興科技趨勢

●人腦機器介面 （Brain-Machine Interface）	●自然語言辨識 （Natural Language Processing）
●塑膠晶片（Flexible Transistors）	●微光電學（Microphotonics）
●資料採礦（Data Mining）	●公開程式碼（Untangling Code）
●數位權利管理 （Digital Rights Management）	●機器人設計（Robot Design）
●生物辨識學（Biometrics）	●微流體學（Microfluidics）

Note. From “The Technology Review Ten,” by MIT Technology Review, 2001, *MIT Technology Review*. Retrieved December 11, 2001, from [http://www.technologyreview.com/magazine/jan01/tr10\\_toc.asp](http://www.technologyreview.com/magazine/jan01/tr10_toc.asp)

目前，資料採礦是一個當紅的話題。而關於資料採礦的定義，根據許多學者的解釋（Grupe & Owrang, 1995; Berry & Linoff, 1997; Cabena, Hadjinaian, Stadler, Verhees & Zanasi, 1997），皆認為它是一種應用在龐大資料庫中的技術，尋找及分析一些隱藏難見與可能認為毫無意義的資料，並自動萃取出可預測的資訊，找到







長期趨勢變化的相關模式 (Relevant Patterns)；這樣的過程也被稱為「知識發掘」(Knowledge Discovery in Database, KDD)的技術(Piatetsky-Shapiro, 1993; Matheus, Chan & Piatetsky-Shapiro, 1993; Fayyad, Piatetsky-Shapiro & Symth, 1996a)；其主要的貢獻在於，它能從資料庫中找出有價值的隱藏事件，以及對資料能歸納出一個有結構的模式，以作為企業的決策支援系統(陳麒文、陳鴻雁，2001)。

基本上，資料採礦與傳統資料分析的不同點在於資料採礦不需事先對資料提出假設，而是在其分析的過程中發掘出資料存在的各項條件，因此更能真實地反映出資料的隱藏特性(Cios, Pedrycz & Swiniarski, 1998)。此外，由於近年來資料採礦的深受重視，其所應用的範圍也越來越廣，目前其應用的領域包含了科學、行銷、工業、商業、體育...等各產業(謝邦昌，2001)。

### 三、資料採礦之分類技術

#### (一) 鑑別分析

鑑別分析是傳統方法中最常使用於處理有關分類問題的統計技術之一，乃由 Fisher 始於 1936 年所提出的一種劃分群體之技術，主要目的在於尋找預測變數最佳之線性組合，而其組間差異平方和相對於組內差異平方和(或總差異平方和)之比值最大，且每一個線性組合皆與先前已經獲得之線性組合均不相關(邱志洲、李天行、周宇超、呂奇傑，2002)。換言之，鑑別分析即於所有資料點歸屬群組為已知的情況下，求取最能將各群組資料點區別清楚之鑑別函數，並且以此函數對新的資料點做歸類或預判其最可能之結果(Cooper & Emory, 1998)。

根據 Johnson and Wichern (1998)說明，鑑別分析必須滿足相關之統計假設，其最主要之必要條件為：所欲分析之資料必需符合常態性、獨立性及均質性等假設條件。此外，亦要求分析變數之變異矩陣須具備均質性與常態性之特性，即鑑別分析假設在各群體變數間的共變異矩陣須符合均質性。若當共變異矩陣具備均質性時，則可採用 Fisher 所提之線性鑑別模式(Linear Discriminant Function, LDA)作為分析工具。

本研究採取逐步選取變數法建立鑑別模式，以保留重要的區別變數，而變數選擇的標準則採用最小的 Wilks' Lambda 值作為選取變數和淘汰變數的標準。鑑別分析的應用範圍相當廣泛，舉凡醫學、商業、化學、生物、教育、生理、行銷研





究、以及考古學等都有其應用之處。例如 Trevino and Daniels (1995)將鑑別分析用於分辨投資績效，及判斷有無直接投資到美國市場對公司績效的影響；Desai, Crook and Overstreet (1996)運用鑑別分析在信用卡及銀行領域中，建立信用分數的鑑別模型；Sung, Chang and Lee (1999)則利用鑑別分析進行產業的破產預測，並指出在破產預測方面，鑑別分析是最常應用的分析方法；Kim, Kim, Kim, Ye and Lee (2000)使用鑑別分析對韓國的房地產市場進行市場區隔分析，並預測消費者的購買行為。

## （二）羅吉斯迴歸分析

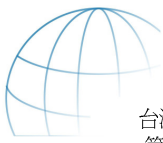
羅吉斯迴歸分析是一個簡單的有母數統計方法，其與傳統的迴歸分析性質類似，而最大的不同則在於反應變數的差異。因此，羅吉斯迴歸分析在應用上也必須符合一些傳統迴歸分析的假設，例如避免殘差項存在自我相關、避免自變數間存在共線性問題、以及要求資料符合常態分配等相關統計假設。

羅吉斯迴歸分析為將已知的分類變數（二元類別數值），利用一系列的數值建立一個最精簡和最能擬合資料的模式，找出哪些變數對判別各群體較有分辨能力，與各群體的分類規則。羅吉斯迴歸分析與鑑別分析一樣，皆是用來處理應變項為名目尺度時，自變項和應變項之間的關係；所不同的是：鑑別分析要求變項需要滿足常態性的分配以及共變異數矩陣相等的假設，如此方能求出最適值(Efron, 1975)。然而羅吉斯迴歸分析並不需要這些假設，即使這些假設全部滿足，羅吉斯迴歸分析仍然可以有很高的預測性(Hosmer & Lemeshow, 1989)。

Odum (1971)指出羅吉斯方程式最早是由 Verhulst 於 1838 年所提出，而由 Lotka 首先加以推廣，在沉寂一段時間後，Pear 及 Peed 重新發現。由於羅吉斯方程式有成長的極限值，因而引發研究人員將羅吉斯方程式的觀念引進迴歸模式中，而建立羅吉斯迴歸(龔昶元，1998)。而在醫學、生物統計、破產預測、財務預警、信用評等、金融房貸...等的統計分析上，羅吉斯迴歸分析的應用早已普遍。例如 Browne and Hoyt (1995)利用羅吉斯迴歸分析法以 1970 年至 1990 年之產險公司為樣本，發現保險公司破產機率與損失率及費用率呈正相關；Hwang, Lee and Liaw (1997)也針對 1985 年至 1988 年間美國 FDIC Annual Income and Call Report 上的銀行進行預測倒閉模型研究；Dahya, Lonie and Power (1998)則應用羅吉斯迴歸分析探討公司績效、股權結構、及董事會組成與高階管理者離職之關係，其樣本是以英







國企業中，1989 年至 1992 年間宣告替換公司董事長或是總經理者之 105 家公司。此外，在體育運動領域的應用上，施致平（1999）利用羅吉斯迴歸分析預測觀眾未來參與中華職籃的運動參與情形。

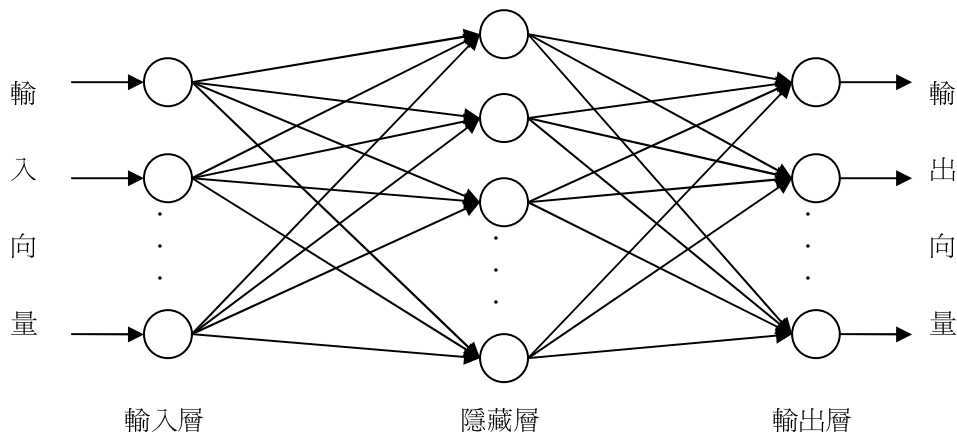
### （三）人工類神經網路

近年來常被討論的人工類神經網路模式，由於沒有傳統統計方法在建構模式時需要滿足許多假設條件的要求，以及其在建構非線性模式方面的優越能力（Zhang, Patuwo & Hu, 1998），因此相對於傳統統計之鑑別分析上的應用範疇而言，人工類神經網路主要應用於變數間較符合非線性資料型態的分類問題上。

人工類神經網路是一種以電腦來模擬人類腦神經細胞網路的科學，簡單地說，人工類神經網路是一個模仿生物神經網路的資訊處理系統，它使用了大量簡單的相連人工神經元來模仿生物神經網路的能力，以使電腦能夠模擬人類的神經系統結構與資料處理方式（Freeman & Skapura, 1992）。而在一個網路模型當中，一個人工神經元將從外界環境或其他人工神經元取得資訊，並依據資訊的相對重要程度給予不同的權重（Weight），予以加總後再經由人工神經元中的激發函數（Activation Function）轉換，再將其結果輸出到外界環境或其他的人工神經元中（Berry & Linoff, 1997）。

人工類神經網路發展至今已有許多模式被提出，其中葉怡成（1992）將之分為四類：監督式（Supervised）、非監督式（Unsupervised）、聯想式（Associate）、與最適化應用網路（Optimization Application Network）。在眾多的網路模式中，以屬於監督式學習的倒傳遞類神經網路（Back-propagation Network, BPN）最具代表性、應用也最為廣泛，此模式是由 Werbos 於 1974 年在哈佛大學提出。根據 Vellido, Lisboa & Vaughan (1999) 的研究指出，在 1992 到 1998 年間，商業領域中使用類神經網路作為研究方法者，約有 78% 的高比例使用倒傳遞類神經網路。倒傳遞類神經網路架構如圖一所示，共包括了輸入層（Input Layer）、隱藏層（Hidden Layer）、與輸出層（Output Layer）。一般而言，輸入層神經元數目即為欲輸入的變數數目；輸出層神經元的數目則為欲得之結果變數個數；而在隱藏層方面，通常包含一個或多個隱藏層。





圖一 單一隱藏層倒傳遞類神經網路模式之構造

倒傳遞類神經網路的演算流程是由輸入層將資料傳給隱藏層，經計算與轉換後送到輸出層（Forward Pass，順向傳遞），在計算完與實際值之差距後，再回傳給隱藏層去修正連接鍵的權數（Reverse Pass，逆向傳遞），使得輸出層之輸出值與真實值的均方誤差最小，此訓練過程會重複至誤差收斂至假設之條件為止。由於人工類神經網路具有嚴謹的數學基礎、巨量平行的處理能力、容錯能力、高聯想力、與能過濾雜訊...等特性，因此在相關領域的應用中，已有許多的文獻被發表，例如市場區隔、股價指數預測、匯率／利率預測、破產預測、信用預測、信用評估、與保險問題中的道德危機...等（Fish, Barnes & Aiken, 1995; Berry & Linoff, 1997; Lee, Jo & Han, 1997; Zhang et al., 1998; Vellido et al., 1999）。

#### （四）多元適應性雲形迴歸

多元適應性雲形迴歸是另一個新興且漸受重視的分類工具。MARS 是由史丹福大學統計與物理學家 Friedman 於 1988 年間所提出的一種多變量無母數迴歸處理的程序，是用來解決多元資料問題的新方法，其基本概念是運用數段解釋方程式加總組合出一個較具彈性的 MARS 預測模型。而其函數中的解釋方程式（Spline Basis Function, BF）之個數則是根據其資料本身參數間之交互關係所決定，並經由評估其損適性（Loss of Fit, LOF）之判斷標準同時獲得最佳及最適合的變數組合、折點數（knots）及交互作用以解決高維度資料的各種問題，它可以自動建立準確的模型以推測其連續和間斷的反應變數（Friedman, 1990）。

MARS 的優點在於，它可以藉由較佳的演算程序來快速找出隱藏在高維度資





料的複雜資料結構其中的最佳變數轉換和交互作用 (Friedman, 1991)。而這個新的迴歸模型建立程序，可以從上述的過程中，有效地在資料中發掘出在過去的一些處理方法難以發現的重要特性及關係。其目前所較被廣泛運用的領域，大部份為資料探勘上的預測及分類問題，如 Smarajit (1996)將 MARS 運用在三種資料上，並將其分類的結果與類神經網路、CART (Classification and Regression Tree)、以及 CUS (Classification Using Splines)進行比較，結果顯示 CART 和 CUS 在運算速度上的優勢，同時在其錯分率上的結果亦不輸於類神經網路的表現，意即 CART 和 CUS 較類神經網路更適用於實務上的應用，此外在 Griffin, Fisher, Friedman and Ryan (1997)於鑽石分類上的應用，他們採用最近鄰居法 (Nearest Neighbour)、CART 以及 MARS 來區分兩類鉻鐵礦以及鑽石的四種等級。而 CART 以及 MARS 同樣有相當優異的表現；許峻源 (2001) 則將 MARS 運用在資料探勘中信用卡分類模式的建構，顯示 MARS 之表現優於傳統之鑑別分析及羅吉斯迴歸，且有不亞類神經網路的表現。

## 參、研究方法

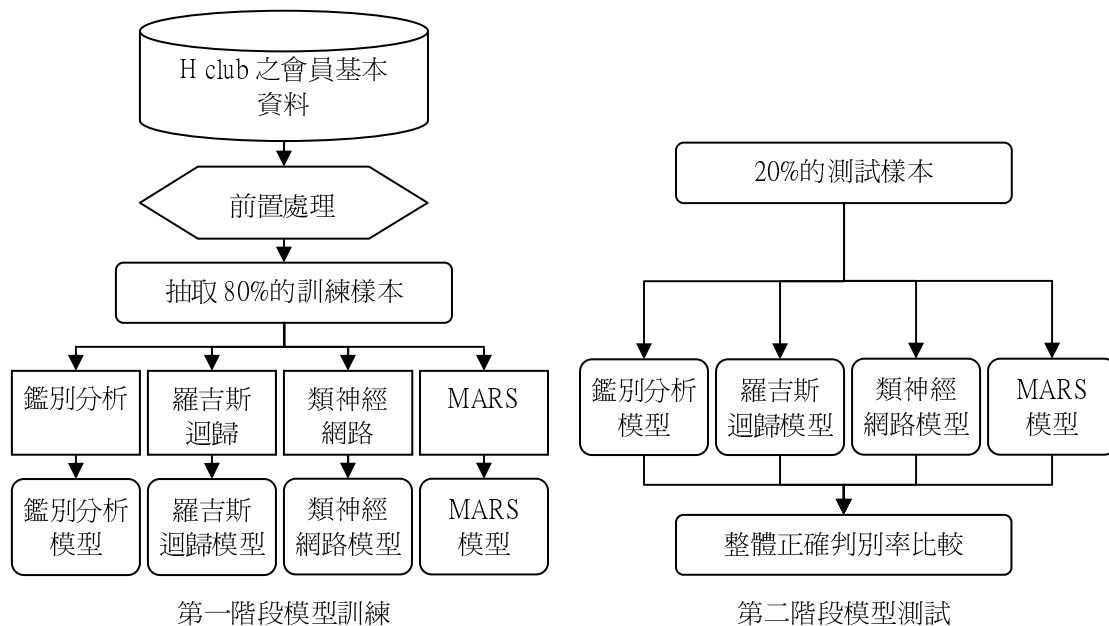
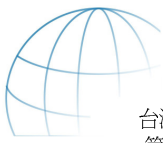
### 一、研究對象

本研究之顧客資料庫由 H club 所提供，內含流失和未流失之會員共 1,287 筆，剔除資料不全與資料內容不合理的資料後，共有 1,152 筆。

### 二、研究架構

而為了驗證模型的適用性，本研究之研究架構分為模型訓練階段與模型測試階段進行 (圖二)。在第一階段之模式訓練階段，主要在建立四種顧客流失分析模式；而第二階段之模式測試階段，主要是利用先前建立好之模式，來預測未來可能流失顧客之正確率。





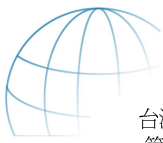
圖二 研究架構圖

而在上述模型的訓練與測試部分，其樣本比例數並無一放諸四海皆準的準則，有些研究將訓練資料與測試資料的比例定為 50：50（彭慧雯，2001；蘇文娟，2000），也有些研究將其定為 67：33（許峻源，2001；邱志洲、李天行、周宇超、呂奇傑，2002）、或者 70：30（吳國禎，2000；呂奇傑，2001；賴欣儀，2000）、85：15（江世傑，2001；周慶華，2001）、甚至 90：10（蔡裕春，2001），而也有研究者認為測試資料筆數須為訓練資料筆數的 10%~30%（黃雅蘭，2001）；但為了不使訓練資料筆數與測試資料筆數太過接近或太過極端而使最後結果產生偏誤，因此參酌大部份的研究（王信勝，2001；張政一，2001；劉嘉鴻，2001；魯振華，2001），以 80：20 的比例隨機抽出訓練模式的樣本及測試模式的樣本個數。

### 三、資料處理

本研究之資料分析軟體，在鑑別分析與羅吉斯迴歸分析方面，使用 SPSS 公司所出版的 SPSS for Windows 13.0 版統計套裝軟體；在人工類神經網路網路分析方面，則採用 Vesta 公司所出版之 Qnet 97 軟體（Vesta Service, 1998）進行分析；而在多元適應性雲形迴歸分析方面，則採用 Salford Systems 公司所出版的 MARS 2.0 版分析軟體（Salford Systems, 2000）。

而在 H club 所提供的顧客資料庫中，經資料前置整理與資料剔除後共剩下 8



個變數可用，分別為性別、年齡、會齡、入會金額、月費金額、付款方式、居住地區、與流失狀況；按流失及未流失之比例共抽出 922 筆資料作為訓練模式之用，而剩下的 230 筆資料則保留為測試之用。

## 肆、實證資料分析

### 一、H club 之顧客組成結構

在 H club 的顧客組成結構描述上，如表一所示。其中平均年齡為 38.16 歲，最大為 73.96 歲，最小為 11.65 歲；平均會齡為 2.63 年，最大為 15.69 年，最小為 0.06 年。

表一 H club 之顧客組成結構描述

人口統計變項		人數	百分比 (%)
性別	男	441	38.28
	女	711	61.72
年齡	20 歲以下	15	1.30
	20 歲以上~25 歲	56	4.86
	25 歲以上~30 歲	156	13.54
	30 歲以上~35 歲	220	19.10
	35 歲以上~40 歲	250	21.70
	40 歲以上~45 歲	168	14.58
	45 歲以上~50 歲	159	13.80
	50 歲以上~55 歲	89	7.73
	55 歲以上	39	3.39
會齡	1 年以下	360	31.25
	1 年以上~2 年	373	32.38
	2 年以上~3 年	142	12.33
	3 年以上~4 年	56	4.86
	4 年以上~5 年	73	6.34
	5 年以上~6 年	37	3.21
	6 年以上~7 年	28	2.43
	7 年以上~8 年	17	1.48
	8 年以上	66	5.73
入會金額	0 公關卡	582	50.52
	2,000 一年期會員折扣一	3	0.26
	3,400 一年期會員折扣二	20	1.74
	3,600 一年期會員折扣三	6	0.52
	4,000 一年期會員折扣四	29	2.52
	4,800 一年期會員折扣五	13	1.13
	5,000 一年期會員折扣六	5	0.43
	6,000 一年期會員折扣七	7	0.61
	入會金額(續) 6,400 一年期會員折扣八	17	1.48
	6,800 一年期會員折扣九	29	2.52







A Case Study of Customer Churn Model—An Example of H Health Club in Taiwan

	8,000	一年期會員折扣十	97	8.42
	10,000	一年期會員	30	2.60
	12,800	團體長期會員折扣一	19	1.65
	15,000	團體長期會員折扣二	35	3.04
	18,000	團體長期會員折扣三	3	0.26
	20,000	團體長期會員折扣四	247	21.44
	25,000	團體長期會員折扣五	5	0.43
	30,000	長期會員	5	0.43
	0	公關卡	160	13.89
	1,250	一年期每月月費團體折扣一	18	1.56
	1,500	一年期每月月費團體折扣二	4	0.35
	1,995	一年期每月月費團體折扣三	1	0.09
	2,500	一年期每月月費	419	36.37
	5,000	一年期每月月費（年繳）團體折扣一	2	0.17
	12,500	一年期每月月費（年繳）團體折扣二	43	3.73
	14,175	一年期每月月費（年繳）團體折扣三	1	0.09
月費金額	15,000	一年期每月月費（年繳）團體折扣四	16	1.39
	16,500	一年期每月月費（年繳）團體折扣五	16	1.39
	19,950	一年期每月月費（年繳）團體折扣六	2	0.17
	25,000	一年期每月月費（年繳）	386	33.51
	27,500	終身會員團體折扣一	11	0.95
	30,000	終身會員團體折扣二	2	0.17
	45,000	終身會員團體折扣三	12	1.04
	66,000	終身會員團體折扣四	57	4.95
	88,000	終身會員團體折扣五	2	0.17
		現金	685	0.35
		自動轉帳	234	0.35
付款方式		自動刷卡	225	19.53
		支票	4	20.31
		手動刷卡	4	59.46
居住地區		臺北市	941	81.68
		非臺北市	211	18.32
流失狀況		未流失	988	85.76
		已流失	164	14.24

## 二、鑑別分析

為取得較精簡的輸入變數，本研究使用逐步鑑別模式來進行鑑別分析，並且依據各變數之 Wilks' Lambda 值（又稱為 U 統計量）作為刪減變數的準則。根據逐步鑑別分析的結果，在 4 次步驟後發現 7 個變數中有 5 個被刪減（性別、年齡、會齡、入會金額、付款方式）。因此，利用其他 2 個變數來建構顧客流失模型之判別模式。

而此模式的適合度檢定係檢定 Wilks' Lambda 值，在本研究中 Wilks' Lambda





值為 0.928，F 值為 17.877 且  $p < .01$ ，因此訓練模式為顯著。此外，第一種顧客類型（未流失）的區別函數： $F_1 = -4.587 + 1.092 \times 10^{-5} \times \text{月費金額} + 5.197 \times \text{居住地區}$ ；而第二種顧客類型（已流失）的區別函數： $F_2 = -5.202 - 2.479 \times 10^{-5} \times \text{月費金額} + 5.778 \times \text{居住地區}$ 。最後，將測試樣本代入前述之訓練模式，由表二可知，整體的正確判別率為 73.04%，而個別的判別正確率以 {1-1} 的比率最高，為 75.63%，意即原始群體為第 1 類（未流失）的樣本被正確地判別到第 1 類（未流失）的比率為 75.63%。

表二 鑑別分析結果

原始群體 流失狀況	判別後群體		總和
	1 未流失	2 已流失	
1 未流失	149 (75.63%)	48 (24.37%)	197
2 已流失	14 (42.42%)	19 (57.58%)	33
總和	163	67	230
整體正確判別率：73.04%			

### 三、羅吉斯迴歸分析

將顧客組成結構變數（性別、年齡、會齡、入會金額、月費金額、付款方式、居住地區）投入羅吉斯迴歸分析，再根據最大概似估計值的最大概似比率統計量是否達到顯著性作為判斷準則，以 LR 式前進逐步刪選法將變數選入模型中。結果發現有 4 個變數被選入，分別為會齡、月費金額、付款方式與居住地區。而由 Hosmer and Lemeshow 檢定得知，p 值（0.04）小於 0.05，表示模型配適度不高。最後，由表三可知，測試樣本的整體正確判別率為 68.70%，而個別的判別正確率以 {1-1} 的比率最高，為 70.05%，意即原始群體為第 1 類（未流失）的樣本被正確地判別到第 1 類（未流失）的比率為 70.05%。

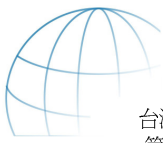
表三 羅吉斯迴歸分析結果

原始群體 流失狀況	判別後群體		總和
	1 未流失	2 已流失	
1 未流失	138 (70.05%)	59 (29.95%)	197
2 已流失	13 (39.39%)	20 (60.61%)	33
總和	151	79	230
整體正確判別率：68.70%			

### 四、人工類神經網路分析

由於人工類神經網路對於變數的選取有較大的自由度，沒有如迴歸分析般的





限制。因此，在網路結構中的隱藏層數目、隱藏層中神經元數目、訓練的學習率大小等，都需要以主觀邏輯判斷，或以不同組合加以測試，並沒有一定道理可資遵循。茲將本研究所需之網路參數設定說明如下：

- (一) 輸入層之輸入變數：共有 7 個自變數，分別為性別、年齡、會齡、入會金額、月費金額、付款方式、與居住地區，因此設定 7 個輸入層之輸入變數。
- (二) 隱藏層數目：考慮網路的學習速率，本研究採用的隱藏層數目為 1。
- (三) 隱藏層中神經元數目：本研究設定隱藏層中神經元數目為：12、13、14、15 與 16。
- (四) 訓練與測試樣本比例：訓練樣本與測試樣本的比例為 8：2，各為 922 與 230 筆。
- (五) 訓練的學習率：以 0.002、0.004、0.006、0.008、與 0.010 五種組合測試。
- (六) 輸出層之輸出變數：本研究為分類型的網路問題，因此輸出層之輸出變數為 1（未流失）、2（流失）之二元變數。

在網路參數的相關設定中，學習率將測試 0.002、0.004、0.006、0.008、與 0.010 等五種組合，而有關網路停止訓練準則方面則以訓練資料的 RMSE 值（Root Mean Squared Error, RMSE，均方根誤差）小於或等於 0.0001，或最多訓練 10,000 次為條件，並以擁有最小測試資料 RMSE 值之網路結構為最佳的網路模式。在本研究中則以隱藏層中神經元數目為 16、學習率為 0.008 之 RMSE 為最小（如表四），利用此一組合再進行 10 次之測試，發現其 RMSE 值均介於 0.229287~0.229623 之間，其穩定性甚佳。而由表三中得知，整體的正確判別率為 84.78%，而個別的判別正確率以 { 1-1 } 的比率最高，為 95.43%，意即原始群體為第 1 類的樣本（未流失）被正確的判別到第 1 類（未流失）的比率為 95.43%。圖三為測試樣本之 RMSE 趨勢圖，由圖三可知建構模式 RMSE 值之收斂情況良好。

表四 不同網路參數組合之 Testing RMSE 值

網路節點	學習率	Testing RMSE	網路節點	學習率	Testing RMSE
12	0.002	0.233709	15	0.002	0.234525
	0.004	0.233062		0.004	0.232095
	0.006	0.230227		0.006	0.233828
	0.008	0.231002		0.008	0.233403
	0.010	0.231697		0.010	0.234475
13	0.002	0.235265	16	0.002	0.235196
	0.004	0.232206		0.004	0.233037





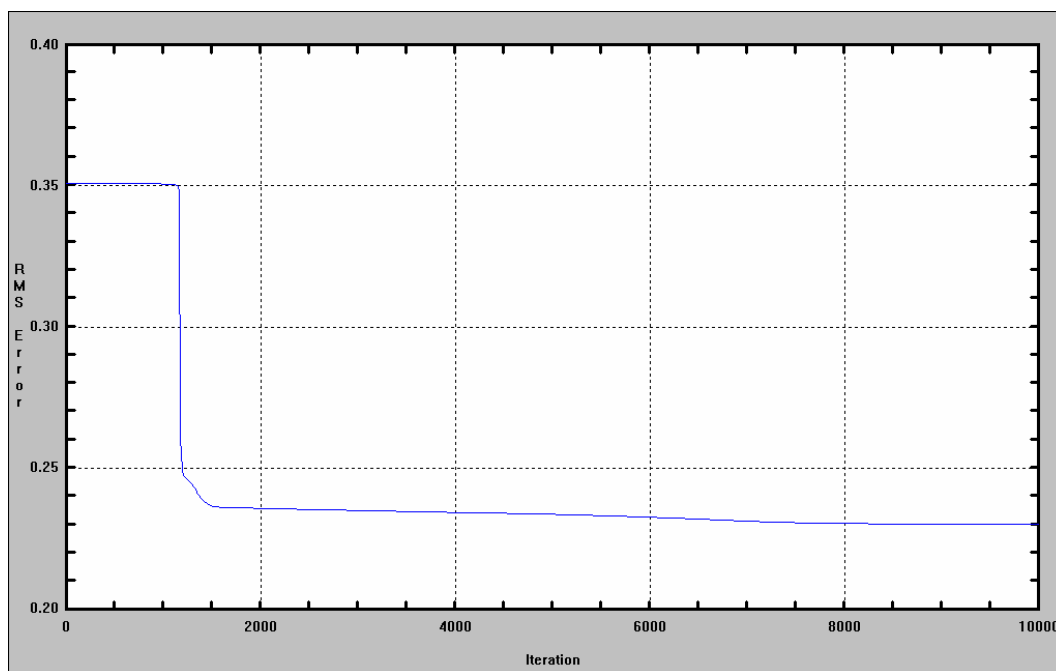
A Case Study of Customer Churn Model—An Example of H Health Club in Taiwan

14	0.006	0.229724		0.006	0.231409
	0.008	0.231440		0.008	0.229310
	0.010	0.236315		0.010	0.230905
	0.002	0.235913			
	0.004	0.231897			
	0.006	0.229901			
	0.008	0.231899			
	0.010	0.233690			

表五 人工類神經網路分析結果

原始群體 流失狀況	判別後群體		總和
	1 未流失	2 已流失	
1 未流失	188 (95.43%)	9 (4.57%)	197
2 已流失	26 (78.79%)	7 (21.21%)	33
總和	214	16	230

整體正確判別率：84.78%

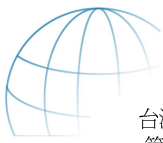


圖三 最終模式測試樣本之 RMSE 趨勢圖

## 五、多元適應性雲形迴歸分析

在 MARS 的資料分析下，根據變異數分析（ANOVA）以及 LOF-GCV 值的評估準則，篩選出顯著的 Basic Function，並根據包含在其中的變數重要性，來決定所需的變數個數，以得到較具貢獻的輸入變數，而相對的變數重要性如表六所示。





表六 相對變數重要性

變數	重要性	log-gcv
月費金額	100.000	0.112
會齡	69.898	0.109
付款方式	59.431	0.108
年齡	23.567	0.106
性別	0.000	0.106
入會金額	0.000	0.106
居住地區	0.000	0.106

根據表六的結果可知，原本所蒐集到的相關變數為 7 個，在經過 MARS 的分析過程後則縮減為 4 個較為顯著的變數。而表七則顯示了整體的正確判別率為 86.52%，而個別的判別正確率以 { 1-1 } 的比率最高，為 100.00%，意即原始群體為第 1 類的樣本（未流失）被正確的判別到第 1 類（未流失）的比率為 100.00%。

表七 MARS 分析結果

原始群體 流失狀況	判別後群體		總和
	1 未流失	2 已流失	
1 未流失	197 (100.00%)	0 (0.00%)	197
2 已流失	31 (93.94%)	2 (6.06%)	33
總和	228	2	230

整體正確判別率：86.52%

## 六、四種模式之整體正確判別率比較

為了比較 4 種模式（鑑別分析、羅吉斯迴歸、人工類神經網路、多元適應性雲形迴歸）的整體正確判別率，本研究將上述各模式的整體正確判別率整理於表八。

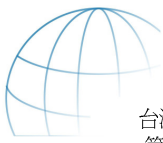
表八 四種模式之判別結果比較表

模式	整體正確判別率
鑑別分析	73.04%
羅吉斯迴歸	68.70%
人工類神經網路	84.78%
多元適應性雲形迴歸	86.52%

由表八可知，整體正確判別率以多元適應性雲形迴歸模式為最高，達 86.52%；而以羅吉斯迴歸模式最低，為 68.70%。而人工類神經網路模式雖然與多元適應性雲形迴歸模式之最終結果相距不大，但是人工類神經網路模式在建構模式時之收







斂情況尚可，也無法如多元適應性雲形迴歸模式般可以挑出重要的變數。故在變數間關係較複雜、較符合非線性關係的資料上，多元適應性雲形迴歸模式成功地建構較佳的判別模式，符合本研究先前在第貳章中所探討之結果。

而在變數的挑選方面，經整理後如表九所示。其中「月費金額」變數在每一個模式中均被挑選到，顯示出這個變數的重要性。而「付款方式」變數、「會齡」變數與「居住地區」變數在其他任三種模式中也均被選擇到，因此其重要性也不容忽視。

表九 四種模式所挑選之變數比較表

模式	鑑別分析	羅吉斯迴歸	人工類神經網路	多元適應性雲形迴歸
所挑選之變數	月費金額 居住地區	月費金額 會齡 付款方式 居住地區	(全部)	月費金額 會齡 付款方式 年齡

## 七、流失顧客之特徵

在經由上述之分析後，本研究發現在鑑別分析部份，其流失顧客之重要特徵為居住在臺北市內、且月繳 2,500 元的會員。

在羅吉斯迴歸部份，其挑選出來之重要變數為月費金額、會齡、付款方式、與居住地區，這也代表了居住在臺北市內、月繳 2,500 元、以現金付款方式繳費、且會齡為 1 年以下的會員為最容易流失的一群。

而在人工類神經網路部份，由於其工具本身的限制，並無法挑選出重要的變數以驗證流失顧客之特徵，因此這部份有待後續的研究者來深入探討；最後在多元適應性雲形迴歸的部份，其流失顧客之特徵為年齡介於 30~35 歲、會齡為 1 年以下、月繳 2,500 元的月費、且以現金支付費用的會員。

因此，對於以上之分析結果，各個分析方法均可以找出對流失顧客的重要特徵，並以多元適應性雲形迴歸的整體正確判別率為最高。而如何在此一流失顧客群中利用適當的行銷方法進行顧客的保留工作，乃是對 H club 的另一重大課題。

此外，鑑於以往使用經驗法則來辨別顧客是否流失、進而採取保留顧客措施的手法不甚符合科學研究之精神，因此本研究所提出之顧客流失模式建構程序，不但在辨識顧客流失的正確率高，且執行效率極佳，對於企業在進行顧客保留措施的決策上提供相當重要的線索，是為一項值得建議使用的工具。





## 伍、結論與建議

### 一、結論

在快速資訊化以及電子化的商業環境中，企業的經營發展已與資訊運用的良好與否牽連上了密切的關係。透過資料採礦的技術與觀念，企業能快速且有效的從資料庫中將有用的資訊與知識發掘出來，以反應市場或消費者的需求，進而提昇企業的競爭優勢。而本研究所提出之顧客流失模式建構程序，主要的目的是為發展一個更為快速、精確的分類模式。此外，為驗證所提模式之有效性，本研究利用 H club 所提供的資料來進行實證研究，結果顯示，多元適應性雲形迴歸的整體分類績效為 86.52%，具有極佳之分類效果，所以運用多元適應性雲形迴歸於健身俱樂部以執行顧客流失分析是一項值得建議使用的工具。

而在流失顧客之重要特徵上，總體而言，為年齡介於 30~35 歲、會齡為 1 年以下、月繳 2,500 元的月費、且以現金支付費用的會員。

### 二、建議

#### （一）對 H club 的建議

就運動管理上而言，由於本研究具資料採礦分類模式之能力，且具備顧客流失之分類、與一般化的建構程序，除了達到預先預警之功能外，對於健康休閒俱樂部普遍存在應如何正確且迅速的處理顧客之流失狀況也提供一個精確、簡易合理的解決途徑；以管理層面上而言，除了具有危機管理的涵義存在外，對於健康休閒俱樂部利用資料採礦達到預先研究、預先處理的過程，以保留顧客或節省大量的顧客取得成本實有相當大的助益。因此，未來 H club 應將此一程序納入原本之管理體系，以建立一個架構完整、且有效率的管理系統。

而就運動行銷上而言，根據研究結果得知，流失顧客之特徵為年齡介於 30~35 歲、會齡為 1 年以下、月繳 2,500 元的月費、且以現金支付費用的會員。因此，H club 應就此一結果進行市場調查與行銷區隔，確切地瞭解此一流失顧客之生活型態、消費特性與心理需求，進而鞏固具有此特徵之既有會員，並將分析的結果轉為顧客維繫方案的對象，俾就運動行銷上得到最大之利潤與最小之成本。而透過對這些顧客的鞏固活動所得到更進一步的資料，再經過流失管理系統的執行與測試，其所衍生出來的新資料回饋至原來的顧客流失分析，如此地循環應用，才





是最理想的做法。

## （二）對後續研究者的建議

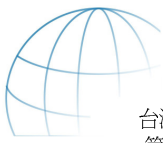
由於本研究僅針對判別率的問題加以探討，事實上在顧客流失分析中，還有其他重要的因素也會影響顧客的存續狀況，例如企業的組織團隊、顧客的生活型態、顧客的消費特性和滿意度等變數。由於研究中未能取得相關的資訊，無法將此概念轉化成一個有效的函數，因此會形成整個衡量過程中的不完整，若能將上述相關的變數加入判別模式的評估判斷上，相信更能提高模式在評估顧客流失上的實用性。

資料採礦技術在健康休閒俱樂部產業中仍存在相當多的應用，例如顧客忠誠度或貢獻度分析，運用資料採礦技術分析那些顧客對企業具有較高的忠誠度或較高的利潤貢獻度，如此有助於企業進行顧客的保留工作。因此，建議後續研究者也能朝此方向發展。

## 參考文獻

- Stotlar, D. K. (1999, 8 月)。關係行銷法在體適能產業的運用。論文發表於行政院體育委員會主辦之「1999 健康體能與有氧運動大會」，臺北。
- 王信勝 (2001)。整合分析層級程序與類神經網路之信用評分模型。輔仁大學資訊管理學系碩士班碩士論文。
- 江世傑 (2001)。模糊類神經網路在消費性貸款之應用。國立成功大學工業管理學系碩士班碩士論文。
- 呂奇傑 (2001)。演化式類神經網路分類技術於資料探勘上之應用。輔仁大學應用統計學研究所碩士論文。
- 吳國禎 (2000)。資料探索在醫學資料庫之應用。中原大學醫學工程學系碩士班碩士論文。
- 邱志洲、李天行、周宇超、呂奇傑 (2002)。整合鑑別分析與類神經網路在資料探勘上之應用。工業工程學刊，19 (2)，9-22。
- 周慶華 (2001)。整合基因演算法及類神經網路於現貨開盤指數之預測—以新加坡交易所摩根臺股指數期貨為例。輔仁大學金融研究所碩士論文。





- 施致平（1999）。從邏輯斯諦迴歸（Logistic Regression）論運動參與之預測。**體育學報**，27，21-30。
- 夏載（2001）。剖析資料採礦在顧客關係管理中的應用。**電子化企業經理人報告 eBusiness Executive Report**，20，71-75。
- 陳麒文、陳鴻雁（2001，12月）。Data Mining 於運動行銷上之應用。論文發表於中華民國運動資訊與傳播協會主辦之「2001 年體育科技與資訊」研討會，臺北。
- 許峻源（2001）。類神經網路與 MARS 於資料探勘分類模式之應用。輔仁大學應用統計研究所碩士論文。
- 張政一（2001）。類神經網路於有價證券預測股價及漲跌之研究。中國文化大學國際企業管理研究所碩士論文。
- 彭慧雯（2001）。建構信用卡資料挖礦架構及其實證研究。國立臺北科技大學生產系統工程與管理研究所碩士論文。
- 黃雅蘭（2001）。臺灣股價指數期貨套利之研究—類神經網路與灰色理論之應用。國立臺灣科技大學資訊管理系碩士班碩士論文。
- 葉怡成（1992）。類神經網路模式應用與實作。臺北：儒林書局。
- 劉嘉鴻（2001）。整合灰預測及類神經網路模型研究股市盤後期貨價格之資訊內涵：以摩根臺股指數及日經 225 指數為例。輔仁大學金融研究所碩士論文。
- 魯振華（2001）。人壽保險廣告類型有效性的預測模式—應用類神經網路。逢甲大學保險學系碩士班碩士論文。
- 蔡裕春（2001）。臺灣地區營造工程物價指數預測之研究—以類神經網路與 ARIMA 模式。輔仁大學應用統計學研究所碩士論文。
- 數博網（2001）。歹年冬資料採礦度小月。**動腦雜誌**，304，44-46。
- 謝邦昌（2001）。資料採礦入門及應用—從統計技術看資料採礦。臺北：資商訊息顧問股份有限公司。
- 賴欣儀（2000）。消費者風險態度與資金運用行為對銀行信用卡業務影響之研究。國立中山大學企業管理學系研究所碩士論文。
- 羅家德（2001）。網際網路關係行銷。臺北：聯經出版社。

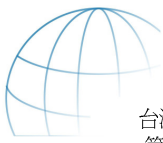




- 蘇文娟 (2000)。臺灣上市企業財務危機預測之實證研究。國立東華大學國際經濟研究所碩士論文。
- 龔昶元 (1998)。Logistic Regression 模式應用於信用卡信用風險審核之研究—以國內某銀行信用卡中心為例。 **臺北銀行月刊**, 28 (9), 35-49。
- Berry, M. J. A., & Linoff, G. S. (1997). *Data mining techniques: For marketing, sales, and customer support*. New York: Wiley Computer.
- Berson, A., Smith, S., & Thearling, K. (2000). *Building data mining applications for CRM*. New York: McGraw-Hill.
- Browne, M. J., & Hoyt, R. E. (1995). Economic and market predictors of insolvencies in the property-liability insurance industry. *Journal of Risk and Insurance*, 62, 309-327.
- Cabena, P., Hadjinaian, P., Stadler, R., Verhees, J. & Zanasi, A. (1997). *Discovering data mining from concept to implementation*. New Jersey: Prentice Hall PTR.
- Cios, K., Pedrycz W., & Swiniarski R. (1998). *Data mining methods for knowledge discovery*. Boston: Kluwer Academic Publishers.
- Cooper, D. R., & Emory, C. W. (1998). *Business research method*. Chicago: Irwin.
- Dahya, J., Lonie, A. A., & Power, D. M. (1998). Ownership structure, firm performance and top executive change: an analysis of UK firms. *Journal of Business Finance & Accounting*, 25, 1089-1118.
- Desai, V. S., Crook, J. N., & Overstreet, Jr. G. A. (1996). A comparison of neural networks and linear scoring models in the credit union environment. *European Journal of Operational Research*, 95, 24-37.
- Efron, B. (1975). The efficiency of logistic regression compared to normal discriminant function analysis. *Journal of the American Statistical Association*, 70, 892-898.
- Fayyad, U. M., Piatetsky-Shapiro, G., & Smyth, P. (1996a). From data mining to knowledge discovery in databases. *AI Mag.*, 17, 37-54.
- Fish, K. E., Barnes, J. H., & Aiken, M. W. (1995). Artificial neural networks: A new methodology for industrial market segmentation. *Industrial Marketing*







- Management*, 24, 431-438.
- Fisher, R. A. (1936). The Use of Multiple Measurements in Taxonomic Problems. *Annals of Eugenics*, 7, 179-188.
- Freeman, J. A., & Skapura, D. M. (1992). *Neural networks algorithms, applications, and programming techniques*. MA: Addison-Wesley Publishing Company.
- Friedman, J. H. (1988). Multivariate Adaptive Regression Splines. *Technical Report*, 102.
- Friedman, J. H. (1990). Multivariate Adaptive Regression Splines. *Technical Report 102 Rev.*
- Friedman, J. H. (1991). Multivariate adaptive regression splines. *The Annals of Statistics*, 19(1), 1-141.
- Griffin, W. L., Fisher, N. I., Friedman, J. H., & Ryan, C. G. (1997). Statistical Techniques for the Classification of Chromites in Diamond Exploration Samples. *Journal of Geochemical Exploration*, 59, 233-249.
- Guape, F. H., & Owrang, M. M. (1995). Database mining discovering new knowledge and cooperative advantage. *Information Systems Management*, 12, 26-31.
- Hosmer, D. W., & Lemeshow, S. (1989). *Applied logistic regression*. New York: John Wiley & Sons.
- Hwang, D. Y., Lee, C. F., & Liaw K. T. (1997). Forecasting bank failures and deposit insurance premium. *International Review of Economics and Finance*, 6(3), 317-334.
- Johnson, R. A., & Wichern, D. W. (1998). *Applied multivariate statistical analysis*. NJ: Prentice-Hall Inc.
- Khirallah, K. (1999). Should retail banks race toward the one-to-one future? *Bank Technology News*, 12(4), 41-43.
- Kim, J. C., Kim, D. H., Kim, J. J., Ye, J. S., & Lee, H. S. (2000). Segmenting the Korean housing market using multiple discriminant analysis. *Construction Management and Economics*, 18, 45-54.





- Lee, H., Jo, H., & Han, I. (1997). Bankruptcy prediction using case-based reasoning, neural networks, and discriminant analysis. *Expert Systems With Applications*, 13, 97-108.
- Matheus, C. J., Chan, P. K., & Piatetsky-Shapiro, G. (1993). Systems for knowledge discovery in databases. *IEEE trans. On Knowledge Discovery and Data Engineering*, 5(6), 903-913.
- Newell, F. (2000). *Loyalty.com: Customer relationship management in the new era of internet marketing*. New York: McGraw-Hill.
- Odum, E. P. (1971). *Fundamental of ecology*. Philadelphia: W.B. Saunders Co.
- Peppers, D., & Rogers, M. (1993). *The future of one to one: Building business relationships one customer at a time*. New York: Currency Doubleday.
- Piatetsky-Shapiro, G. (1993). An overview of knowledge discovery in database: Recent progress and challenges. *Rough Sets, Fuzzy Sets and Knowledge Discovery - Proceedings of the International Workshop on Rough Sets and Knowledge Discovery*, 93, 1-10.
- Smarajit, B. (1996). Classification Using Splines. *Computational Statistics and Data Analysis*, 22, 505-525.
- Salford System (2000). *MARS V2.0-for windows 95/98/NT*. San Diego, CA: Salford System.
- Sung, T. K., Chang, N., & Lee, G. (1999). Dynamics of modeling in data mining: interpretive approach to bankruptcy prediction. *Journal of Management Information Systems*, 16, 63-85.
- Trevino, L. J., & Daniels, J. D. (1995). FDI theory and foreign direct investment in the United States: a comparison of investors and non-investors. *International Business Review*, 4, 177-194.
- Vesta Services (1998). *Qnet97-neural network modeling for windows 95/98/NT*. Winnetka, IL: Vesta Services, Inc.
- Vellido, A., Lisboa, P. J. G., & Vaughan, J. (1999). Neural networks in business: A





- survey of applications (1992-1998). *Expert Systems With Applications*, 17, 51-70.
- Werbos, P. J. (1974). *Beyond regression: new tools for prediction and analysis in the behavioral sciences*. Unpublished doctoral dissertation, Harvard University, MA.
- Walters, S. (1994). *Customer service: A how-to-do-it manual for librarians*. New York: Neal-Schuman Publishing.
- Zhang, G., Patuwo, B. E., & Hu, M. Y. (1998). Forecasting with artificial neural networks: The state of the art. *International Journal of Forecasting*, 14, 35-62.
- The Technology Review Ten (2001). *MIT Technology Review*, 2001 January/February. Retrieved December 11, 2001, from [http://www.technologyreview.com/magazine/jan01/tr10\\_toc.asp](http://www.technologyreview.com/magazine/jan01/tr10_toc.asp)

