

應用與比較多種資料探勘預測技術於 電腦代理商銷售預測之研究

呂奇傑^{1*}、蘇子庭²

摘要

銷售預測在供應鏈管理中佔有相當重要之角色，且資訊產業相較於其他產業而言，變動更加快速。因此若能建構一個精確有用的銷售預測模式，便可增進企業在供應鏈以及銷售管理的能力。本研究透過多種資料探勘(data mining)預測技術，包括倒傳遞類神經網路(back-propagation network, BPN)、支援向量迴歸(support vector regression, SVR)、多元適應性雲形迴歸(multivariate adaptive regression splines, MARS)以及小腦系統控制器(cerebellar model articulation controller, CMAC)等方法對資訊產業做銷售預測及應用。本研究應用某電腦品牌代理商之每月銷售資料為實證資料，評估資料探勘預測技術於電腦代理商銷售預測的有效性。實證結果顯示，在單一模式下，MARS 與 SVR 的預測準確率較佳，BPN 與 CMAC 的預測績效較差；而在混合模式下，則以整合 MARS 與 SVR 的預測模式較佳。

關鍵詞：銷售預測、資料探勘、多元適應性雲形迴歸、支援向量迴歸、類神經網路

1 健行科技大學 工業管理系

1 健行科技大學 工業管理系

* 聯絡作者: jerry.lu@uch.edu.tw; chijie.lu@gmail.com

Forecasting sales for computer reseller:

A Comparison of BPN, SVR, MARS and CMAC models

Chi-Jie Lu^{1*} Zi-Ting Su²

Abstract

Sales forecasting is one of the most important issues in managing computer reseller since computer products are characterized by product variety, rapid specification changes and rapid price declines. In this study, we use four different data mining prediction methods including back-propagation network (BPN), support vector regression (SVR), multivariate adaptive regression splines (MARS) and cerebellar model articulation controller (CMAC) to forecast sales for computer reseller and compare their forecasting performance. Experimental results from a real sales data of computer reseller show that the prediction accuracy of single MARS and single SVR models are the better than that of single BPN and CMAC. The integrated MARS-SVR model outperforms all other four single and two integrated forecasting models and is a promising model for forecast sales for computer reseller.

Keywords: Sales forecasting, data mining, multivariate adaptive regression splines, support vector regression, neural network.

1 Department of Industrial Management, Chien Hsin University and Science and Technology.

2 Department of Industrial Management, Chien Hsin University and Science and Technology.

*Corresponding author: jerry.lu@uch.edu.tw; chijie.lu@gmail.com

1. 前言

近年來，由於科技高度發展，經濟環境也日趨複雜，各種不確定因素增加，導致了企業必須花更多的精力與時間維持正常營運和利潤。因企業必須兼顧本身與上、下游之間的利益關係，使得整個供應鏈成為生命共同體，也就是供應鏈管理(supply chain management)之概念。在供應鏈管理流程中，企業必須向供應商採購相關之原料，先經由工廠加工為成品後，再透過公司之銷售部門(團隊)，將產品銷售至顧客手中。因此從上游之供應商至下游的顧客的關係是環環相扣，而銷售預測(sales forecasting)在供應鏈管理中佔有相當重要之角色，也是非常重要的議題(Luis and Richard 2007; Chang et al. 2009)。電腦資訊產業相較於其他產業而言，變動更加快速，而電腦資訊產品更是不在話下。若能精確掌握產業環境動態，便可為企業提高利潤，降低不必要之開銷。因此，建立一個適合的銷售預測模式對於企業的會是一個重要的議題。

而隨著資訊科技的發展，自動化資料收集設備，如銷售時點系統(point of sale, POS)，的大量使用，使得收集銷售資料是一件簡單的事情，較困難的是要如何從大量資料中擷取有用資訊。以前企業只能使用一些簡單的敘述性統計或者是須符合一般統計假設的資料(例如：誤差項須符合常態分配)。但在實務中，我們所蒐集到的資訊幾乎都是呈現非線性關係，且在此種情況下我們很難去發覺到隱藏其中的重要資訊。因此我們若能以更有效率的方式去發掘有用的資訊以提供銷售預測使用，相信能對於企業的銷售決策有莫大幫助。

資料探勘(data mining)便是針對如何在龐大的資料海中擷取有用資訊而發展出來的方法。一般而言，資料探勘是從資料倉儲(data warehouse)或是資料庫中，發現有意義之規則或是模式，進而透過自動或是半自動化之工具對大量資料進行探索與分析之過程(Michael and Gordon 2004)。資料探勘並非是一種技術或為一套軟體分析工具，而是一種整合專業統計及資訊科技之應用。資料探勘相關技術之應用可以概分為預測模式(prediction)、分類問題(classification)、資料摘要(summarization)、相依模式(dependency modeling)以及變動與偏差偵測(change and deviation detection)等五大類型(Fayyad et al. 1996)。其中預測模式是資料探勘中最常被探討議題之一(Berry and Linoff 1997)。預測模式主要是用現有的數值來預測未來的數值。亦即，根據被檢驗對象的歷史性觀察值來推估該屬性未來的可能結果，其所分析的數值會與時間有關，相關議題如銷售預測、股價預測及需求量預測等(Tay and Cao 2001; Carbonneau et al. 2008; Efendigil et al. 2009; Pedregal and Young 2008; Tsai et al. 2009)。但由於建構預測模式的過程常需考量許多複雜與多樣化因素的影響，如經濟環境、政府政策、產業景氣、消費者心理等，建立有效的預測模型極為不易，因此如何建構能產生高預測準確度與低預測誤差的預測模式是一項具挑戰性的任務。

目前已有許多資料探勘技術被發展出來，但卻很少應用在銷售預測之領域。且應用範圍也都以金融服務業及製造業等為主要對象(Carbonneau et al. 2008; Efendigil et al. 2009; Pedregal and Young 2008; Tsai et al. 2009)，鮮少看到將其應用於資訊服務業的文獻。因此本研究想透過多種人工智慧技術，如倒傳遞類神經網路(back-propagation network, BPN)、支援向量迴歸(support vector regression, SVR)、多元適應性雲形迴歸(multivariate adaptive regression

splines, MARS)以及小腦系統控制器(cerebellar model articulation controller, CMAC)等方法對資訊產業做銷售預測及應用。並且透過 MARS 能選取重要變數之能力，進一步探討這些變數在管理上的意涵。

在本研究中所使用之方法，倒傳遞類神經網路(BPN)具有學習準確度高、回想速度快、可處理具雜訊之資料等優點，是使用最廣泛的人工智慧工具，但卻也擁有學習速度較為緩慢與易陷入局部最小值(local minimum)等缺點(Lee and Chen, 2005)。而支援向量迴歸(SVR)是一個新興的預測工具，其擁有全域最佳解(global optimum)及考慮結構風險(structural risk)等優點，使 SVR 所建構預測模型之一般化能力(generalization)較佳(Tay and Cao 2001)。而在 MARS 部分，由於其具有篩選重要變數的能力，可以藉此分析各預測變數對預測結果的影響程度(Lee & Chen, 2005)。這對企業決策者而言，無疑可以提供更有用的資訊以做為決策之參考依據。最後，CMAC 雖然傳統上常用於分類與控制領域，但近來有一些新穎的 CMAC 演算法被提出來用於預測的議題 (Lu and Wu 2009)，因此本研究想再加入 CMAC 來探討其應用於銷售預測之模式建構。

本研究使用四種不同方法(BPN、SVR、CMAC 與 MARS)，進而發展出兩種預測架構。在此為了展現出篩選變數之重要性，故將模型分為有無經過 MARS 篩選之模型。第一種單純使用此五種方法建立預測模型。第二種則先以 MARS 作為變數篩選之工具，爾後再使用 BPN、SVR 與 CMAC 建立第二種預測模型，希望藉由 MARS 優異的篩選變數之能力，看是否可篩選出對資訊產業較為重要之變數，並發展其管理意涵，以便公司高層做更好、更快速之決策。

2.文獻回顧與探討

2.1、銷售預測

商業領域中，透過良好的銷售預測可以更有效率地幫助決策者計算生產或是原料成本，甚至是銷售價格(LeVee, 1993)。所以企業能做出正確的銷售預測，便能得到相當重要的利潤，像是擴大現金流量、了解買賣時機、了解顧客所需、規劃產能與了解銷售趨勢等(Chang et al. 2009)。

傳統上，許多研究以時間序列模式進行銷售預測，如 ARIMA、SARIMA 等(Lim and McAleer 2002; Goh and Law 2002)。而人工智慧技術蓬勃發展後，許多研究都以結合傳統與人工智慧技術來做銷售預測，這也大大提昇了預測準確率，且被應用在各個領域的銷售預測。Kuo 與 Xue(1998) 使用了 ANN、ARIMA 與結合模糊理論的 ANN-FNN(fuzzy neural network)，並加入 if-then 篩選原則來對某家木瓜牛奶公司做銷售預測。而其結果顯示，FNN 之預測能力的確優於 ANN 與 ARIMA。Kuo 與 Xue(1999) 結合了模糊理論與類神經網路對一家知名連鎖超商做銷售預測，發現其結合模型的預測能力的確優於單一使用類神經網路的預測。

Luis and Richard (2007) 結合了 ARIMA 與 ANN 模式對 Chilean 超級市場做銷售預測，其結果顯示此預測模型可以幫助公司高層做更正確的決策。Hyunchul et al. (2007)先透過 ICA(independent component analysis)篩選變數後再用 ANN 來對韓國一家知名購物商城做預測。其預測能力的確優於先透過 PCA(principal component analysis)做變數篩選後再使用 ANN 之方法。Chang and Wang(2006) 使用了逐步迴歸分析(stepwise regression analysis, SRA)、模糊

德爾菲法(fuzzy Delphi method, FDM)、R/S 分析(rescaled range analysis)與結合模糊理論的倒傳遞類神經網路(fuzzy back-propagation, FBPNN)等四種方法對台灣電路板產業進行銷售預測。其結果顯示出, FBPNN 的預測能力的確優於其他三種方法。Chang et al. (2007) 使用灰色相關分析(grey relation analysis, GRA)、進化模糊權重神經網路(weighted evolving fuzzy neural network, WEFuNN)、進化模糊神經網路(evolutionary fuzzy neural network, EFuNN)、SRA 與 BPN 等五種方法對台灣電路板產業做銷售預測。其結果顯示, WEFuNN 的預測能力優於其他四種方法。

銷售預測雖已經應用在各個領域, 但卻沒有看到應用在高科技產品之下游, 最靠近消費者之銷售預測, 故本研究將以電腦品牌代理商為例, 對其做銷售預測。

2.2 資料探勘技術於銷售預測

本研究將使用四種資料探勘預測技術(SVR、MARS、SVR 與 CMAC), 以下分別簡要的回顧其用於預測的文獻。

BPN 是一個廣為人知與應用的預測技術, 廣泛的應用於各種預測議題, 其中 BPN 用於銷售預測的有, Ilan et al. (2001)使用倒傳遞類神經網路(BPN)、溫特斯指數平滑法(Winters exponential smoothing)、ARIMA 以及多元迴歸(multivariate regression)模型預測美國零售業之月營業額, 預測期間從 1978 年 1 月至 1995 年 4 月。以前一期(one-step)或前數期(multi-step)資料作為輸入變數。其結果顯示, 在不穩定的經濟條件下, 以前數期資料建構之 BPN 模型預測表現最佳。Celia et al. (2003) 使用了 ANN 與 ARIMA 建構模型來對美國兩家女性服飾專賣店做銷售預測, 而其結果顯示 ANN 相對於 ARIMA 有更佳的預測準確率。Chang et al. (2006) 結合了 SOM(self organization map)、neural network、GA(genetic algorithms)與 FRB(fuzzy rule base), 以及使用多元迴歸與 BPN 等方法對電路板(printed circuit board)產業做銷售預測。其結果顯示, 混和模型的預測能力的確優於多元迴歸與 BPN。

SVR 雖已廣泛的用於各種預測議題, 但最常被用於股價預測。Huang et al. (2005) 以支援向量機預測 Nikkei 225 指數每日漲跌方向, 與隨機漫步模型(random walk model)、線性鑑別分析(linear discriminant analysis)、二項式鑑別分析(quadratic discriminant analysis)與倒傳遞類神經網路比較, 實證結果, 支援向量機預測結果皆較其他方法準確。Pai and Lin (2005a) 混和了 ARIMA 與 SVMs 模式對某家公司之股價做預測。其結果也顯示, 此混和模型的預測能力的確優於單一使用 SVMs 與 ARIMA。Kim(2003) 使用 SVR、BPN 與 CBR(case-based reasoning)等方法, 並以報酬率作為技術指標對韓國每日綜合股價指數收盤價漲跌做預測。其結果顯示, SVR 具有較佳預測能力。而 SVR 應用於銷售/需求預測的文獻就比較少, Pai and Lin(2005b) 使用 SVR、SARIMA 與廣義迴歸神經網路(federal regression neural networks, GRNN)來對具季節性之工業機械總值做預測。而其結果顯示, SVR 預測準確率的确優於 SARIMA 與 GRNN。Chen and Wang(2007) 其所提出結合基因演算法(genetic algorithms)與支援向量迴歸(support vector regression)來對中國大陸旅遊業做預測, 其結果顯示 GA-SVR 的預測準確率的确優於 ARIMA 與 BPN。

在 MARS 部分, 主要被用於資料探勘的分類(classification)問題, 較少被用於預測相關議題。Zareipour et al. (2006)將 MARS 應用在安大略湖每小時能源價格(hourly Ontario energy price, HOEP)的預測上, 研究解果顯示 MARS 提供了準確的預測。Zhou and Leung (2007)使用 MARS

進行物件導向軟體維護性做預測(object-oriented software maintainability)，並且比較多元線性迴歸模型(multivariate linear regression models)、ANN、迴歸樹模型(regression tree models)、支援向量機(support vector models)的預測結果，研究發現 MARS 的預測準確度是最佳的。Kiran and Ravi (2008)藉由使用如多元線性迴歸和 MARS 以及人工智慧如 BPN、TreeNet 等方法，建構模型去準確預測軟體的可靠度，結果顯示無論是 MARS 或是 BPN 這類的工具都有相當好的表現。

最後在 CMAC 部分，其主要用於控制、診斷與分類，但鮮少應用在預測領域(Lu and Wu 2009)。Kuo(2009)使用 CAMC 以及 neuro-fuzzy CMAC，想藉此比較兩種預測模式之優劣。透過模擬資料之形態，且將其資料分為訓練資料與測試資料。而其結果顯示，neuro-fuzzy CMAC 具有較好之預測績效。Qiaolin et al. (2005) 將 CMAC 用在電力價格預測。

3.研究方法

3.1 BPN

類神經網路是目前發展極為迅速的一門學科，其最大的優點是除可應用於建構非線性之模式外，且能彌補傳統統計方法建構模式時須設立許多假設條件的缺點 (Zhang et al. 1998)。而類神經網路的原始想法及基本構造皆和神經生物學中的神經元 (neuron) 構造相似。一般來說，類神經網路中神經元的作用和真實的神經元相仿，是將外界的輸入值依相對重要性的不同給予權數 (weights)，加總後再經由神經元中的激發函數 (activation function) 轉換，而得到一輸出值。

類神經網路的網路型態有許多類，而其中又以倒傳遞類神經網路(back propagation neural networks, BPN)為最具代表性、應用最廣的模式之一。倒傳遞類神經網路乃隸屬於監督式學習網路模式的一種，其資料是以順向 (forward) 之方式向前傳遞。其網路結構一般而言包含三層神經元：輸入層 (input layer)、隱藏層 (hidden layer) 及輸出層 (output layer)。一般而言，輸入層神經元的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。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，計算與實際值的差距後，再回傳給隱藏層去修正連接鍵的權數，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。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，研究者須以文獻探討、專家意見判斷或經由統計方法處理，來選取輸入層之輸入變數。而在輸入層之變數決定後，對於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，並沒有一定理可資遵循 (Zhang et al. 1998)。

3.2、SVR

為說明 SVR 模式，首先考慮一個典型的迴歸模式，其目的是估計一個未知方程式 $r(x)$ 來預測一個未知的值 q ，兩者的關係為， $q = r(x) + \delta$ ，其中 δ 是隨機獨立並且平均數為零的隨機誤差 (random error)， x 是多變數的輸入變數， q 是單一 (scalar) 的輸出值。這個方程式可以用有限個數的訓練樣本來建立： $(x_i, q_i), (i = 1, \dots, n)$ 。在迴歸問題中可分為線性迴歸與非線性迴歸問題，線性迴歸問題是較易處理，並且已有相當多的技術被提出來與成功的應

用；相對而言，非線性迴歸則較不容易處理。SVR 主要就是被發展來處理非線性迴歸的議題的工具，其主要概念是將一個在低維度輸入空間（input space）中非線性迴歸之問題，轉換成在高維度特徵空間（feature space）中的線性迴歸問題。因此在 SVR 中，輸入變數 \mathbf{x} 首先被轉換（map）到一個高維度的特徵空間（F）中，在這特徵空間中，輸入變數可以被修正為與輸出變數有線性的關係。以數學模式表示，SVR 模式可以表示為（Vapnik 2000）

$$f(\mathbf{x}) = (\omega \cdot \Phi(\mathbf{x})) + b \quad (1)$$

其中， ω 為權重向量， b 是為偏差值（bias）， $\Phi(\mathbf{x})$ 為一轉換函數，用以將輸入變數 \mathbf{x} 以非線性的方式轉換到高維度的特徵空間中， $(\omega \cdot \Phi(\mathbf{x}))$ 則用來描述在特徵空間 F 中的值。其中任何被用來當核心函數之函數並需滿足 Mercer's 條件(Vapnik 2000)。常用的轉換函數有多項式（polynomial）函數及輻射基底函數（radial basis function, RBF）(Vapnik 2000)。

3.3、MARS

MARS 基本概念為運用數段解釋方程式(spline basis function, BF)加總，以組合較有彈性的預測模型解決多元且複雜的資料問題(Friedman 1991)。模型中藉由資料本身參數間之交互關係來決定解釋方程式個數，另外，經過評估每個解釋方程式的缺適性(loss of fit, LOF)，以決定包含之影響變數個數。篩選的過程採用前推式與後推式(forward and backward stepwise procedure) 的演算法：其中在前推式階段，主要是藉由使用遞迴分割的概念，透過增加 BF 的個數(包含了新的主要效果、節點或交互作用等)，模型經由不斷成長成為過度配適模型(overfit model)，一直到基本方程式的數目達到研究者所指定的極大值，同時也建構出一完整模型，所以在此階段模型中將會包含了眾多的 BF。

而在後推階段，主要是根據 LOF 決定 BF 個數時，考量每個 BF 在加入主要模式後是否具有貢獻性。換言之其概念為，以 LOF 高低做為判斷標準，並且採用 GCV 準則(generalized cross-validation criterion)(Friedman 1991)，將對模式貢獻度較小的 BF 一一刪除，最終產生最適的 MARS 模型，以期在可接受的範圍內降低模型的複雜度，讓資料處理速度與判斷更加快速。

以 MARS 與統計分析之迴歸模型建構過程相比，其對於資料限制較寬鬆，所以當資料結構較為複雜時(如資料為非線性結構)，迴歸分析往往受限於變數間可能存在的高度自我相關或是交互作用，而難以建構出解釋能力較佳的模式。而 MARS 則可克服這些缺點，依照不同資料屬性，預測出間斷或連續的應變數，保留資料中的重要訊息。

3.4 CMAC

Marr 在 1969 年提出小腦皮質模型之概念，而在 1975 年時，Albus 根據這個模型發展出小腦皮質模型的數學模式，並將此模型命名為小腦模型控制器(cerebellar model articulation controller, CMAC)。小腦利用這三層(分子層、細胞層、粒狀層)儲存各部份傳入纖維所來的訊息，然後由小腦的傳出纖維傳出訊息，而小腦模型控制器主要為一種模仿人類小腦皮質分層儲存訊息的架構而來。CMAC 通常用於學習複雜且非線性之函數，其中包含五種記憶單位(cells)：輸入空間(input space)、感應單位(sensory cell)、連結單位(association cell)、實體記憶單位(physical memory)與輸出單位(output cell)。基本上小腦模型控制器(CMAC)是類似於人類小腦的一種學習結構。在此技術中，每一變數都被量化，所以整個問題的狀態空間被分成一

塊塊的離散狀態。對於一輸入離散狀態，可將此狀態對映一組量化的輸入向量（視為記憶體位址），而從記憶體中獲得其輸出。傳統上 CMAC 通常都使用在分類上，而本研究嘗試將其用於預測議題上，並使用 Lu and Wu(2009)所提出的 CMAC 預測技術，有關這演算法的詳細內容，請參閱 Lu and Wu(2009)。

4.實驗結果

4.1 資料蒐集與績效衡量準則

本研究以台灣某家電腦代理商之月銷售金額為預測對象，研究期間從 1996 年 1 月至 2009 年 2 月，共 158 個月，趨勢圖如圖 1 所示。所採用之變數有每季之移動平均(X1)、每半年之移動平均(X2)、前一期之銷售金額(X3)、前兩期之銷售金額(X4)、前三期之銷售金額(X5)、前期與後期之變動率(X6)、3 個月之相對強弱指標(X7)、6 個月之相對強弱指標(X8)、12 個月之相對強弱指標(X9)、3 個月之乖離率(X10)、6 個月之乖離率(X11)與 12 個月之乖離率(X12)等 12 個變數，其中 X6 至 X12 是參考在股票市場常用的技術指標。訓練與測試比例則是採用 8：2，即訓練比數為 126 筆(期間從 1996 年 1 月至 2006 年 6 月)、測試筆數為 32 筆(2006 年 7 月至 2009 年 2 月)。

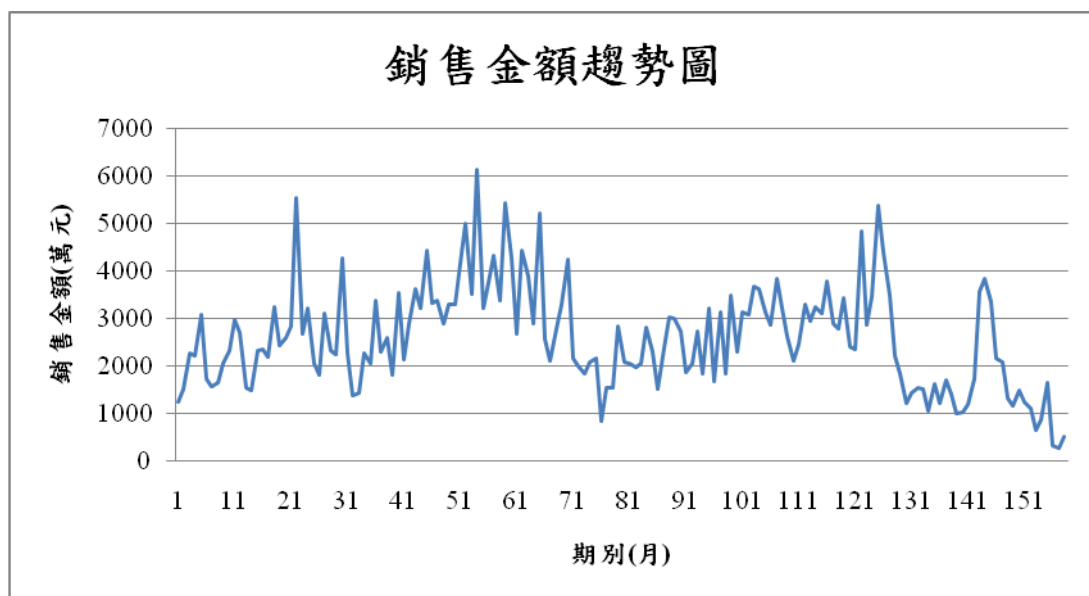


圖 1、銷售趨勢圖

在預測績效衡量評估準則方面，本研究將藉均方根誤差(root mean square error, RMSE)、平均絕對百分比誤差(mean absolute percentage error, MAPE)、平均絕對離差(mean absolute difference, MAD)與百分比均方根誤差(root mean square percentage error, RMSPE)等 4 個常用指標評估預測模式的預測結果準確度。RMSE、MAD、RMSPE 與 MAPE 為衡量預測值與實際值之間的誤差(deviation)，此四個指標皆為愈小表示預測值與實際值愈接近。

4.2、預測結果-單一模式

為比較模式之結果，本節先行敘述單一預測模式之結果。本研究使用四種預測技術，BPN、SVR、CMAC 與 MARS 建構預測模式。

首先在利用 MARS 部分，銷售金額資料經過單位轉換之前處理後，經由 MARS 分析，利

用篩選基礎方程式(basis function, BF)時，找出顯著 BF，並根據包含在其中的變數重要性，決定所需變數之個數，進而得到較具貢獻性的輸入變數。由表 1 為 MARS 篩選變數之結果。由表 1 可知，原本 12 個變數經由 MARS 分析過後篩選出三個變數，分別為 X1「每季之移動平均」、X4「(t-2)期之銷售金額」與 X3「(t-1)期之銷售金額」，此三個變數對於影響預測銷售金額之重要程度，其中以「每季之移動平均」之相對重要性達 100%為最高。推論因銷售趨勢並無劇烈之變動，且沒有受到季節因素之影響，故成為單純預測模型中最重要之變數。

表 1 MARS 變數篩選結果

變數代碼	變數名稱	重要性(%)
X1	每季之移動平均	100.000
X4	(t-2)期之銷售金額	78.939
X3	(t-1)期之銷售金額	46.694

在表 2 中，MARS 共建構出四條較具顯著性基本方程式，並發展成 MARS 最適模型方程式。從最適模型方程式中可知，使用 MARS 預測時，變數 X1「每季之移動平均」與 X4「(t-2)期之銷售金額」有正向關係，且變數 X3「(t-1)期之銷售金額」與 X4「(t-2)期之銷售金額」則具有反向關係。

表 2 MARS 模式基本方程式及最適模式方程式

基本方程式：

$$BF2 = \max(0, X1_MA(3) - 381.335);$$

$$BF4 = \max(0, X4_ (t-2) - 1226.768);$$

$$BF5 = \max(0, 1226.768 - X4_ (t-2));$$

$$BF6 = \max(0, X3_ (t-1) - 282.490);$$

MARS 最適模型方程式：

$$Y = -365.252 + 3.000 * BF2 - 1.000 * BF4 + 1.000 * BF5 - 1.000 * BF6;$$

在單純 BPN 部分，輸入層神經元方面，在此採用與 MARS 相同之 12 個變數做為輸入變數，而隱藏層中神經元之數目則採用 $2n$ 、 $2n+1$ 、 $2n+2$ 等組合進行試誤(trial and error)，其中 n 為輸入層神經元數目，故以 22、23、24、25 與 26 等五種組合進行測試；最後在網路輸出層的部分，則為一個神經元。而測試過程中，當學習率超過 0.2 時，學習效果會呈現無法收斂之狀況，因此須不斷往下修正試誤，故學習率選擇上主要以測試 0.001、0.005、0.01、0.015、0.02、0.025、0.03、0.05 與 0.1 等九種組合為網路模式的選擇範圍。經試誤結果，當網路結構為{12-24-1-0.1}：即輸入層節點數為 12，隱藏層節點數為 24，輸出層節點數為 1，且學習率為 0.1 時，有最小測試資料 RMSE 值 1000.5580，為最佳模式。

在單純 CMAC 部分，本研究是以在不同學習率下之 RMSE 之比較結果，選擇最小之 RMSE 作為單純預測模型之參考技術指標。而所選擇出不同學習率，分別為 0.001、0.005、0.010、0.015、0.020、0.025、0.030 與 0.1 等八種，在這八種不同學習率下可發現在學習率為 0.1 時會有最小測試資料 RMSE 值 545.3490。

在單純 SVR 部分，本研究所選擇之參數範圍為 $\gamma=2^{-7}$ 至 2^9 而 $C=2^{-7}$ 至 2^9 。在此範圍之下做參數搜尋，以找出最佳參數組合。經過試誤之結果我們發現，當 $\gamma=2^1$ 且 $C=2^{-3}$ 時有最小測試資料 RMSE 值 187.6825，此為單純 SVR 之參數最佳組合。

本研究建構之四種單一模式之預測結果如表 3 所示。其中 MARS 模式的 MAD 與 RMSE 分別為 23.04 與 100.98，而 SVR 之 MAPE 與 RMSPE 分別為 0.20%與 0.23%，此兩種方法在

技術指標上各有領先。由此可知，MARS 與 SVR 之銷售預測準確率皆優於 CMAC 與 BPN。

表 3 各模式於總銷售金額預測誤差結果彙總表

	MAD	RMSE	MAPE	RMSPE
MARS	23.04	100.98	4.25%	24.02%
BPN	854.85	1000.56	1.50%	2.27%
SVR	305.28	187.68	0.20%	0.23%
CMAC	419.46	545.35	0.26%	0.33%

4.3、預測結果-混合模式

由於 MARS 擁有優異之篩選與解釋變數之能力，故本節將先以 MARS 做為變數篩選之工具後，再以所挑選之變數做為 BPN、SVR 與 CMAC 之輸入變數建構預測模式，建立本研究第二種混合預測模型(MARS-BPN、MARS-SVR 與 MARS-CMAC)。如前所述，MARS 所挑選出來的重要預測變數為每季之移動平均、(t-1)期與(t-2)期之銷售金額，因此將用來建構兩階段預測模式。

在 MARS-BPN 部分，經由試誤可知，當網路結構為{3-6-1-0.005}時有最小的測試 RMSE 值，為 MARS-BPN 之最佳模式。而同樣經過試誤結果，在 MARS-CMAC 部分，學習率為 0.1 是最佳網路結構；而在 MARS-SVR 部分則是 $\gamma=23$ 而 $C=2-6$ 時有最小測試資料 RMSE。

本研究建構之 MARS-BPN、MARS-SVR 與 MARS-CMAC 之預測結果如表 5 所示，其中 MARS-SVR 之 MAD、RMSE、MAPE 與 RMSPE 分別為 141.68、49.21、0.11%與 0.16%，其表現皆優於 MARS-BPN 與 MARS-CMAC。

表 4 混合模型預測結果彙總表

	MAD	RMSE	MAPE	RMSPE
MARS-SVR	141.68	49.21	0.11%	0.16%
MARS-CMAC	538.13	639.53	0.28%	0.32%
MARS-BPN	605.93	748.95	0.53%	0.59%

為了比較單一模型與混合模型之結果，我們將單一模式與混合模式的結果統整於表 5。由表 5 可知，使用 MARS 或是 MARS-SVR 可達最佳預測效果；但整合 MARS 與 SVR 的預測模式可以提供較佳的預測績效，且能夠找出重要的預測變數，是建構銷售預測模式的首選。

表 5 單純與混合模型預測誤差結果彙總

方法	MAD	RMSE	MAPE	RMSPE
MARS	23.04	100.98	4.25%	24.02%
BPN	854.85	1000.56	1.50%	2.27%
SVR	305.28	187.68	0.20%	0.23%
CMAC	419.46	545.35	0.26%	0.33%
MARS-SVR	141.68	49.21	0.11%	0.16%
MARS-CMAC	538.13	639.53	0.28%	0.32%
MARS-BPN	605.93	748.95	0.53%	0.59%

5.結論

本研究使用 MARS、SVR、BPN 與 CAMC 之建構兩種銷售預測架構：單一模式與混合模式，其中混合模式是先以 MARS 篩選出變數後再結合 SVR、CMAC 與 BPN 建構預測模式，

本研究並針對台灣某大型電腦通路商之月銷售金額進行預測。實證結果顯示，在單一預測模式下，以 SVR 與 MARS 之預測準確度較高且預測誤差較低；而在兩階段混合模式中，以 MARS-SVR 之預測準確率最佳且誤差也最小。整體而言，預測模式以 MARS-SVR 的績效最佳，代表先經由 MARS 篩選出重要變數後再建構預測模式的方法的確能提高預測準確率。最後，MARS 所篩選出來的變數為 X1「銷售額之每季移動平均」、X3「前一期之銷售金額」與 X4「前二期之銷售金額」，代表這三個變數對銷售結果有重要的影響，廠商可以根據這三個變數對其銷售狀況進行有效管理。

參考文獻

1. Berry, M. J. A., & Linoff, G. (1997). *Data mining technique: For marketing, sales, and customer support*. NY: John Wiley & Sons.
2. Carbonneau, R., Vahidov, R. and Laframboise, K. (2007). Machine learning-based demand forecasting in supply chains. *International Journal of Intelligent Information Technologie*, 3, 40-57.
3. Celia, F., Ashish, G., Amar, R., & Les, S. (2003). Forecasting women's apparel sales using mathematical modeling. *International Journal of Clothing Science and Technology*, 15(2), 107-125.
4. Chang, P. C., Liu, C. H., & Fan, C. Y. (2009). Data clustering and fuzzy neural network for sales forecasting: A case study in printed circuit board industry. *Knowledge Based Systems*, 22, 344-355.
5. Chang, P. C., & Wang, Y. W. (2006). Fuzzy Delphi and back-propagation model for sales forecasting in PCB industry. *Expert Systems with Applications*, 30, 715-726.
6. Chang, P. C., Liu, C. H., & Wang, Y. W. (2006). A hybrid model by clustering and evolving fuzzy rules for sales decision supports in printed circuit board industry. *Decision Support Systems*, 42, 1254-1269.
7. Chang, P. C., & Wang, Y. W., & Liu, C. H. (2007). The development of a weighted evolving fuzzy neural network for PCB sales forecasting. *Expert Systems with applications*, 32, 86-96.
8. Chen, K. Y., & Wang, C. H. (2007). Support vector regression with genetic algorithms in forecasting tourism demand. *Tourism Management*, 28, 215-226.
9. Efendigil, T., Öñüt, S. and Kahraman, C. (2009), A decision support system for demand forecasting with artificial neural networks and neuro-fuzzy models: A comparative analysis. *Expert Systems with Applications* 36, 6697-707.
10. Fayyad, U., Piatetsky-Shapiro, G., & Smyth, P. (1996). The KDD process for extracting useful knowledge from volumes of data. *Communications of the ACM*, 39(11), 27-34.
11. Friedman, J. H. (1991). Multivariate Adaptive Regression Splines (with discussion). *The Annals of Statistics*, 19, 1-141.
12. Goh, C., & Law, R. (2002). Modeling and forecasting tourism demand for arrivals with stochastic nonstationary seasonality and intervention. *Tourism Management*, 23, 499-510.
13. Huang, W., Nakamori, Y., & Wang, S. Y. (2005). Forecasting stock market movement direction with support vector machine. *Computer and Operations Research*, 32, 2513-2522.
14. Hyunchul, A., Eunsup, C., & Ingoo, H. (2007). Extracting underlying meaningful features and canceling

- noise using independent component analysis for direct marketing. *Expert Systems with Applications*, 33, 181-191.
15. Ilan, A., Min, Q., Sadowski, R. J. (2001). Forecasting aggregate retail sales: a comparison of artificial neural networks and traditional methods. *Journal of Retailing and Consumer Services*, 8, 147-156.
 16. Kim, K. J. (2003). Financial time series forecasting using support vector machines. *Neurocomputing*, 55, 307-319.
 17. Kiran, N. J., & Ravi, V. (2008). Software reliability prediction by soft computing techniques. *The Journal of Systems and Software*, 81, 576-583.
 18. Kuo, H. C. (2009). CMAC-based neuro-fuzzy approach for complex system modeling. *Neurocomputing*, 72, 1763-1774.
 19. Kuo, R. J., & Xue, K. C. (1998). A decision support system for sales forecasting through fuzzy neural networks with asymmetric fuzzy weights. *Decision Support Systems*, 24, 105-126.
 20. Kuo, R. J., & Xue, K. C. (1999). Fuzzy neural networks with application to sales forecasting. *Fuzzy Sets and Systems*, 108, 123-143.
 21. Lee, T. S., Chiu, C. C., Chou, Y. C., & Lu, C. J. (2006). Mining the customer credit using classification and regression tree and multivariate adaptive regression splines. *Computational Statistics & Data Analysis*, 50, 1113-1130.
 22. Lee, T. S., & Chen, I. F. (2005). A two-stage hybrid credit scoring model using artificial neural networks and multivariate adaptive regression splines. *Expert Systems with Applications*, 28, 743-752.
 23. LeVe, G. S. (1993). The key to understanding the forecasting process. *Journal of Business Forecasting*, 11 (4), 12-16.
 24. Lim, C., & McAleer, M. (2002). Times series forecasts of international travel demand for Australia. *Tourism Management*, 23, 389-396.
 25. Luis, A., & Richard, W. (2007). Improved supply chain management based on hybrid demand forecasts. *Applied Soft Computing*, 7, 136-144.
 26. Lu, C.-J. and Wu, J.-Y. (2009), Forecasting financial time series via an efficient CMAC neural network, *Lecture Notes in Electrical Engineering*, accepted and further coming.
 27. Michael, J. A. B., & Gordon, S. (2004). *Data Mining Techniques for Marketing, Sales, and Customer Relationship Management* (2ed.). NY: John Wiley & Sons, Inc.
 28. Pai, F. P., & Lin, C. S. (2005a). A hybrid ARIMA and support vector machines model in stock price forecasting. *Omega*, 33, 497-505.
 29. Pai, P. F. & Lin, C. S. (2005b). Using support vector machines in forecasting production values of machinery industry in Taiwan. *International Journal of Advanced Manufacturing Technology*, 27(1), 205-210.
 30. Pedregal, D. J., and Young, P. C. (2008), Development of improved adaptive approaches to electricity demand forecasting. *Journal of the Operational Research Society*, 59, 1066-76.
 31. Qiaolin, D., Jing, T., & Jianxin, L. (2005). Application of new FCMAC neural network in power system marginal price forecasting. *The 7th International Power Engineering Conference*, 1-57.

-
32. Tay, F. E. H. and Cao, L. (2001), Application of support vector machines in financial time series forecasting, *Omega*, 29, pp. 309-317.
 33. Tsai, T. -H, Lee, C. -K and Wei, C. -H (2009), Neural network based temporal feature models for short-term railway passenger demand forecasting. *Expert Systems with Applications* 36, 3728-36.
 34. Vapnik, V. N., 2000, *The Nature of Statistical Learning Theory*, Springer, New York, NY.
 35. Zareipour, H., Bhattacharya, K., & Canizares, C. A. (2006). Forecasting the hourly Ontario energy price by multivariate adaptive regression splines. *Power Engineering Society General Meeting*, 1-7.
 36. Zhang, G., Patuwo, B. E., & Hu, M. Y. (1998). Forecasting with artificial neural network: The state of the art. *International Journal of Forecasting*, 14, 35-62.
 37. Zhou, Y., & Leung, H. (2007). Predicting objected-oriented software maintainability using multivariate adaptive regression splines. *The Journal of Systems and Software*, 80, 1349-1361.