

以機器學習建構一外銷銷售預測模型之研究

葉貞吟¹

國立屏東大學

jenyiny@gmail.com

蔡雅婷²

國立屏東大學商業自動化與管理學系

jennifer840727@gmail.com

摘要

為了在不斷變化的環境裡提升企業的競爭優勢，組織的管理階層必須依據資訊即時做出正確的決策，以因應瞬息萬變的市場趨勢。企業一切的採購、生產計劃源頭即為銷售預測(Kuo & Xue, 1999)，銷售預測在市場上的決策和規劃極為重要。銷售預測基於過去銷售數據，並考慮各種影響銷售的因素，各國整體性的發展與經濟環境變動影響各國人們消費，本研究以一家專事外銷文具用品的製造廠商為個案研究，運用該公司企業資源規劃系統(Enterprise resource planning, ERP)資料庫中過去外銷銷售歷史資料及外銷客戶國家整體經濟環境因素包括匯率、通膨指數、國內生產總值、購買力平價和外部因素的初等教育小學生年度學齡人口，以機器學習之方法，建構一預測未來銷售狀況之模型，並探討銷售量與經濟環境因素之關聯性。

關鍵字：機器學習、銷售預測、外部環境

A Machine Learning Approach of Sales Forecasting Model

JEN YIN YE^a YA TING TSAI^b

^a Department of Commerce Automation & Management, National Pingtung University

^b Department of Commerce Automation & Management, National Pingtung University

ABSTRACT

In the increasingly competitive environment, companies are constantly focused on adjusting the sales strategies to closely satisfy customers' requirements and preferences. Furthermore, accurate sales forecasting is crucial to assist in production planning and improving business. The sales forecast is usually based on past sales history data and considers various factors that possible affect sales through certain analytical methods to propose a predictive sales result (Heidi and Adamantios, 2003). This forecasting method assumes that past sales patterns will continue into the future. In the export-sales company, sales forecasting is challenging because the export-sales forecaster is plagued by a number of problems unknown in international market. Forecasting experts base their judgment often on the economic expectations, numerically represented in economic indicators. Macroeconomic information (*e.g.*, *Exchange Rates*, *Inflation Rate*, *Gross domestic product*, *Purchasing Power Parity*) has been used in improving forecasting. The other potential of external indicators is population needs to be included in export-sales forecasting model as well. However, to the best of our knowledge, limited work has been done on including those external indicators on sales forecasting that this paper aims to address.

Many forecasting techniques have been developed in recent years. The regression analysis method is usually used to reflect the relationship between sales and the independents. Machine learning model has the ability to be non-linear and able to find in a large amount of sample data therefore have good results of practical prediction in different field (Zhang et al., 1998; Hamzacebi et al., 2009; Singhal et al., 2011). This study presents a different use of this methodology in a real-world case of export-sales forecasting. The prediction models are compared and evaluated using four machine learning techniques

based on past sales history data and external indicators. The key insight of the study is a substantial improvement in prediction accuracy using machine learning technique. The result adds to the discussion of the continuing debate about superiority of computational methods over statistical techniques and provides a suitable sales forecasting method for enterprises.

Keywords: Machine Learning、Sales Forecasting、Export sales

1. 緒論

1.1 研究背景與動機

近年來隨著科技日新月異的進步與資訊的快速流通，資料量更是大量的蓬勃增加，人們漸漸發覺資料的重要性，開始蒐集大量的資料、利用統計方法並建立預測模型，以達到欲研究探討的目標。尤其，面臨大數據時代的衝擊，傳統預測系統難以應對市場需求的不確定性和多樣性，人類已經無法單純用大腦來進行計算和分析龐大的數據，而必須仰賴機器學習來處理資料和數據，加上現今所有企業皆正處於外部相互競爭日趨激烈，且客戶要求不斷提高以及技術發展越來越快的環境中，為了在不斷變化的環境裡提升企業的競爭優勢，組織的管理階層必須依現有的資訊即時做出正確的決策以因應瞬息萬變的市場趨勢，然而，不僅是科技的進步與發展外，經濟環境也日趨複雜，各種不確定因素增加，導致企業必須花更多的精力與時間維持正常營運和利潤外，如何較精確的去掌控未來影響銷售量之因素，便是對於企業成敗而言有著重大的影響力。

預測是企業長期的預算規劃與成本控制的基礎，有能力做出精確的預測對公司在做重大決策時是很重要的依據。故銷售容易受許多因素的影響，因此，找出重要的銷售額預測變數提供企業進行決策時之參考，可提升銷售管理效能。王平(2013)預測的不準確，必然導致供應鏈長期處於被動應對的低效率運營狀態。銷售預測對企業利潤的影響是如此重大，因此許多企業都定期執行銷售預測，但預測出的結果卻往往不盡理想(Dalrymple, 1987)。企業一切的採購、生產計劃的源頭即為銷售預測(Kuo & Xue, 1999)，一旦銷售預測不準確，所影響的並非只是終端零售業者，而是從製造商、批發商等整條供應鏈上下游的各個生產或銷售計劃，牽一髮而動全身(Chopra & Meindl, 2007)。

1.2 研究目的

近年來，由於科技高度發展，在技術層面上滿足客戶需求已不成問題，此銷售的轉變著重資料蒐集、分析以及處理的能力，以滿足客戶所期望之實際需求，有價值之客戶乃指「一個未來為企業帶來的利潤大於企業花在其身上之成本之客戶」，因此企業應將行銷的重心放在有價值的客戶上建立長期的關係，使其成為企業帶來長久的利潤(Kotler & Armstrong, 1999)，如同，Hughes(1996)認為應將有價值的舊客戶視為企業的一項珍貴資產，因此瞭解客戶的行為狀態及生命週期，對於企業而言也將等同於瞭解企業資產的消長趨勢，其重要性不待言。探討需求變異高、生命週期短的產品對於幾種不同預測方法的適用性，除了以預測精準度作為評選標準外，同時還考量經濟測量準則為產品選擇最佳的預測模型(Flores et al. 1993)，未來需求訂單的不確定性高和客戶訂單變動不穩定，都是管理者在預測時必然會面臨的決策問題。因此市場需求面的不確定因素，亦成為影響企業銷售成敗關鍵之一，在產品導向之思維引領下，企業內部專注於產品研發、標準化以及如何降低總體製造成本，以推式生產的方式將商品提供給客戶，卻忽略消費者的消費習慣與可能受限的消費環境，且每一種商品的銷售狀況都不盡相同，這些銷售狀況的不同有時會反映在需求曲線上。而客戶導向之銷售模式，重心就是以客戶的角度為思考的出發點，滿足變異性高的客戶需求為其銷售重點，Reichheld(1996)指出若企業能減少 5%的客戶流失率，可提升 25%到 85%的利潤。

在全球總體經濟的各項因素不確定下，終端產品需求的不確定性，進而影響到供應鏈的需求。管理者如何準確的預測客戶的訂單量，已成為管理者所重視的議題。因此，評估一家公司的經營績效有許多的指標，銷售量一直都是一個重要的項目，因為銷售量就代表著公司在市場上的佔有率，也是公司營運的基礎，銷售量的成長代表了獲利的機會，反之銷售量降低則會帶來獲利緊縮。因此有效的銷售量預測對於評估一家企業的經營績效與營運前景是很重要的項目。若企業能有效掌握影響自身營運狀況的因素，進而做出正確的銷售額預測，最後，精準的預測出其未來可能變動的趨勢，找出影響最大的關鍵因素並加以控制，便能使公司有效率的經營而能實現公司的目標，助於

提升企業的營運效率與評估企業的經營績效和營運前景。

鑒於國家總體經濟環境變動對人們消費影響，本研究利用機器學習之方法，考慮國家總體經濟環境與外部之因素，如匯率(Exchange Rates)、通膨指數(Inflation Rate)、國內生產總值(Gross domestic product, GDP)、購買力平價(Purchasing power parity)和初等教育小學生年度學齡人口，藉由經濟環境的改變與過去銷售歷史資料，提出一預測未來銷售狀況之模型，並探討銷售狀況與經濟環境因素影響之關聯性。

2.文獻回顧與探討

本研究主要目的為找出影響每筆客戶銷售量的重要因素，期望利用建立預測模型，提供企業較為準確之預測模型，以提升銷售額的預測趨勢，並了解機器學習內涵及國內外銷售預測相關之研究，故本章將逐一說明相關文獻。

2.1 銷售預測相關研究

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 & Cao, 2001; Carbonneau et al. 2007; Efendigil et al. 2009; Pedregal & Young, 2008; Tsai et al. 2009)。銷售預測對企業利潤的影響是如此重大，因此許多企業都定期執行銷售預測，但預測出的結果卻往往不盡理想(Dalrymple, 1987)；劉向上(1984)不論預測所使用之方法為何，或多或少還是會受到一些因素影響而產生誤差，而這些影響因素通常無法事先得知，因而使得預測結果與實際銷售量出現差距。銷售預測(Sales Forecasting)在商業領域中，佔有相當重要之角色，也是非常重要的議題(Luis & Richard, 2007; Chang et al. 2009)。

Still(1988)認為銷售預測為在一特定行銷計畫或所認定之外在因素下，個別廠商對未來某段期間內之銷售量或銷售額所做之推測。銷售預測有三個主要的目的：第一、有助長期規劃與目標訂定，第二、便於短期之作業性決策，第三、提供評估銷售績效之基礎。常見的銷售預測方法有主管意見評審會、銷售共同預測、相關因素法、客戶意見調查法、德菲法等非計量性的銷售預測方法，並說明在選擇特定預測方法時必須考慮之因素。銷售預測是企業在營運過程中於制定營運政策基礎，藉由銷售預測使各階段的決策能增進營運效率並進一步達到提升獲利和降低成本之目的(Philip et al. 2006)。目前已有許多針對銷售預測進行的研究，如(Chen et al. 2011)利用 ELM 建構零售產業的銷售預測模式，及商業領域中，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)。

銷售預測是以過去的歷史資料，考慮各種的影響因素後，結合企業的銷售數據，透過一定的分析方法提出未來的可能銷售(Heidi & Adamantios, 2003)，意指分析銷售歷史性資料，以取得銷售本身內在資訊方式來做預測；或同時加入影響銷售的外在資訊變數做預測。目前有許多預測方法已廣泛的被應用於銷售預測，如：移動平均法、天真預測法及迴歸分析等(Bianchi et al. 1998; Snyder et al. 2002; Holt, 2004)。徐守德等人(1994)時間序列分析法、單純預測法、主觀判斷法等均為可行之銷售預測方法。銷售量是一種時間序列(time-series)的資料，統計學家將時間序列分成四個部分：(1)長期趨勢(trend)、(2)週期性(cyclical effect)、(3)季節性(seasonality)、(4)其它無法預測的變動(irregularity) (Black, 2011; Taylor, 2004)。由於企業對於未來營運環境的不明朗且產品生命週期日益縮短，企業無法確定所提供之產品與服務在下一季是否還能受到客戶追逐與青睞，若能提前預知產品或服務的銷售趨勢，將能幫助企業提前擬訂有利之策略(Chang et al. 2005)。為使未來的銷售預測更正確，可改進輸入資料、降低時間序列資料的差異與供應鏈中資訊共享等(Geurts & Whitlark, 1999; Schrieber, 2005)。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 & McAleer, 2002; Goh & 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，這也大大提升了預測準確率，且被應用在各個領域的銷售預測。

表 1 銷售預測相關研究之整理

篇名	作者	年份	研究方法	研究變數	研究目的	研究結論
Exploring the use of deep neural networks for sales forecasting in fashion retail	A.L.D. Loureiro, V.L. Miguéis and Lucas F.M. da Silva	2018	Decision trees, Random forest, Support vector regression, Artificial neural networks, Liner regression	家庭、顏色、流行、商店類型、價格、規模、期望水平、銷售量	探討利用深度學習方法預測時裝業的銷售狀況，預測未來幾季新產品的銷售情況。	比較深度學習方法包含決策樹，隨機森林，支持向量迴歸，人工神經網絡和線性迴歸。結果發現深度學習在預測時尚零售市場的銷售方面具有良好的表現考慮到評估指標的一部分，其中表現績效最佳，即隨機森林。
Application of artificial intelligence (AI) for effective and adaptive sales forecasting	Abhang Mehendale and Nadheera Sherin H R	2018	ARIMA, Time-series, Delphi Method	銷售額、GDP、GNI、通貨膨脹、失業率、利率、總資產、折舊、總支出、股票變動、競爭者：利率、總資產、總銷售額、折舊、股票變動	應用人工智能(AI)術來識別和預測複雜的銷售模式，並將結果與傳統的預測模型進行比較。	當大規模使用時，神經網絡可以作為一個更好的替代傳統的預測方法。它的適應性使其優於傳統模型，但線性模型不適應變化，相較於神經網絡精確度表現較低。神經網絡考慮到影響的各種因素，因此有助於它更好地預測並防止企業虧損。
Explaining machine learning models in sales predictions	Marko Bohanec, Mirjana Kljajić Borštnar and Marko Robnik-Šikonja	2017	ML model(RF, NB, DT, NN, SVM), IME model	產品、公司規模、競爭者、預算、FRI、FRP、成長值、資源、客戶、交叉銷售、交易狀態	演示機器學習(ML)一般解釋方法於新應用，以解決B2B銷售預測的複雜現實商業問題。	在 B2B 銷售中，預測滑點（交易延遲，即交易未完成）經常發生。給予個案公司一個表現最佳的 ML 預測模型，使用所提出的解釋來揭示見解。
Sales Forecasting of New Entertainment Media Products	Christina Hofmann-Stöltinga, Michel Clementb, Steven Wuc and Sönke Albersd	2017	Exponential, Gen. Gamma, Weibull, Bass	產品、生產預算、國家、廣告、競爭者、價格	顯示 Weibull model 之優點，並與管理規劃數據相比，可以提供更好的銷售預測。	結果顯示，Weibull model 特定的擴散形狀優於所有模型在三個行業中的娛樂產品。基於結果，使用此模型對所有三個行業的估計參數進行進一步分析，測試預測性能是否受相應擴散模型選擇的影響，但找不到實質性的差異。
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；蘇子庭	2018	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)	每季之移動平均、(t-2)期之銷售金額、(t-1)期之銷售金額	使用四種不同方法(BPN、SVR、CMAC 與 MARS)，進而發展出兩種預測架構。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，而後再使用 BPN、SVR 與 CMAC 建立第二種預測模型，希望藉由 MARS 優	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。

					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。	
基於多種變數選擇技術與極限學習機於飯店業銷售預測	李天行；呂奇傑；鄭婷方	2017	逐步迴歸分析法、基因演算法、貪婪演算法、極限學習機(ELM)	經濟面指標：景氣對策分數、景氣同時指標綜合指數、景氣領先指標綜合指數、貨幣總計數(M1B)、消費者物價基本分類指數(CPI)、躉售物價基本分類指數(WPI)、失業率、入境人數、新臺幣兌美元匯率、股價指數(月)、進口物價指數、出口物價指數、進口貿易總值、出口貿易總值；歷史銷售資料之技術指標：去年同期銷售額、前二個月銷售額移動平均、前三個月銷售額移動平均、前一個月銷售額前二個月銷售額前三個月銷售額、前二個月銷售額之相對差距比率前三個月銷售額之相對差距比率、前二個月銷售額之乖離率、前三個月銷售額之乖離率。	整合變數選擇技術(variable selection)與極限學習機(extreme learning machine, ELM)預測技術於飯店業的銷售預測。透過對這些重要預測變數的討論與瞭解可達到重點管理及提升銷售管理效能的目的。	根據各飯店最適混合預測模式的變數篩選結果，對晶華飯店與亞都飯店而言，最重要的預測變數為「去年同期銷售額」，代表這兩家飯店的銷售具有穩定性，季節性的影響明顯。而國賓飯店的重要預測變數為「MA(3)」，代表其銷售額可能不太受總體經濟狀況的影響，較受短期的因素影響。

2.2 總體經濟環境

2.2.1 匯率的定義

一國貨幣兌換另一國貨幣的比率，是以一種貨幣表示另一種貨幣的價格。由於世界各國貨幣的名稱不同，幣值不一，所以一國貨幣對其他國家的貨幣要規定一個兌換率，即匯率(Exchange Rates)。匯率可以表示為一段時間內的平均匯率或期末的匯率，由國際貨幣基金組織分為三大類，反映了當局在確定一個國家匯率的多樣性，包含市場匯率，其中匯率「浮動」並且主要由市場力量決定；官方匯率，其中匯率由一個國家的當局「固定」；和兩者之間的安排，其中利率對另一種貨幣或一種貨幣組合具有穩定的價值。

2.2.2 通膨指數的定義

通膨指數(Inflation Rate)是貨幣超發與實際需要的貨幣量之比，用以反映通貨膨脹、貨幣貶值的程度。經濟學上，通貨膨脹率是指一般價格總水平在一定時期(通常是一年)內的上漲率。通貨膨脹衡量生活水平的下降，其中

包括食品，能源和不包括食品和能源的總匯率變動不穩定，透過消費者價格指數(Consumer price index, CPI)衡量的通貨膨脹定義為通常由特定家庭群體購買的商品和服務的價格制度。

2.2.3 國內生產總值的定義

國內生產總值(Gross Domestic Product, GDP)指的是一年內在某國所生產的最終財貨和勞務的市場價值。GDP係指一國境內所有的生產者(企業)在一定期間內「新創造」之產品與服務的總計，並非國內各產業產值的加總。衡量國內生產總值在一國境內，相較於其他包含境外生產水準的經濟指標更足以揭露出國內經濟活動的熱絡程度(范秉航，2006)。

2.2.4 購買力平價的定義

購買力平價(Purchasing Power Parity, PPP)是透過消除國家之間價格水平差異來平衡不同貨幣購買力的貨幣轉換率。最簡單的形式是，購買力平價顯示不同國家同一商品或服務的本國貨幣價格比率。同時也是針對產品組和每種不同的聚合水平(包括 GDP)計算的。定價的商品和服務是所有那些作為最終支出一部分的樣本如家庭消費、政府服務、資本形成和淨出口，由 GDP 覆蓋。

2.3 總體經濟環境對銷售之影響相關研究

研究指出檢測匯率的波動和出口在東亞五個國家之間的經濟關係，採用向量自我迴歸(VAR)、誤差修正模型(ECM)、變異數分解(VD)為實證方法，樣本國家為印尼、日本、韓國、新加坡和泰國，整合結果說明出口和匯率間具有長期的穩定均衡，統計結果顯示在多數國家，匯率的波動對實質出口有負向影響(Poon et al. 2005)。

Burney(1996)本文藉著橫斷面的數據資料估計國內生產總值之經濟成長方程式的增加，分析出口和經濟成長的關係，且使用最小平方方法(OLS)以及隨機係數法(RC)研究實證資料，樣本範圍分成二個時期，從 1965 年至 1980 年與 1980 年至 1990 年，變數有實質國內生產總值、實質出口、資本、勞動和能源等其他的解釋變數，實證說明出口與經濟成長存有正向關係的時期在 1980 年至 1990 年，而國內生產總值的成長，出口成長扮演其重要角色。Gharty(1993)以台灣、日本和美國為實證對象，探討出口和經濟成長的因果關係，實測方法為因果檢定法，美國與台灣的研究變數採國內生產總值與出口，日本變數採用國內生產總值、出口、資本存量和貿易條件。實證結果說明台灣的出口成長會導引經濟成長；美國的經濟成長導引出口成長，則日本為貿易條件導引經濟增長，且日本出口成長與經濟成長兩者之間互有回饋的因果關係。本研究預測結果支持大幅的 GDP 變化波動對出口的需求有相當大的影響，結果顯示，GDP 的波動對預測模型的研究有次要的影響。

表 2 總體經濟環境對銷售之影響相關研究之整理

篇名	作者	年份	研究方法	研究變數	研究目的	研究結論
台灣汽車銷售量與經濟因素關聯性之研究	吳炎崑； 洪煙平	2011	案例式推演算法、迴歸分析	汽車銷售量(千輛)、國民儲蓄率(%)、消費性利率(%)、消費物價指數(95年=100)、失業率(%)	探討經濟因素對台灣汽車銷售量的影響，以案例式推理及迴歸分析二種方法，經由模型的建立，研究經濟因素與汽車銷售量之間的關聯性，進而預測未來台灣地區汽車銷售量之趨勢。	結果說明：案例式推演以過去的歷史資料可以預測未來的汽車銷售達 74%至 77%的精確度；迴歸分析以過去的歷史資料可以預測未來的汽車銷售達約 82%精確度。因此，利用經濟環境歷史資料所做預測迴歸分析優於案例式推演。

2.4 RFM 模型

近年來，客戶關係管理已成為各企業愈來愈重視的議題之一，如何找到有價值的客戶對企業而言是非常重要的，因此需藉由 RFM 模型中的三項衡量指標，便可容易且清楚地得知客戶的消費行為，以達到提供企業進行客戶消費行為分析並規劃客戶關係經營策略之目標。RFM 模型不僅協助企業掌握重要客戶，更是幫助企業有效分配行銷資源的一項實用分析工具，是產業界公認為最有效的客戶價值分析方法。

1994 年 Hughes 提出了 RFM 模型，其為利用客戶的過去之三項資料做為變數，即最近購買日(R:Recency of the last purchase)，是客戶距離現在的日期的最後一筆交易，當客戶最後一次購買的日期距離現在愈近，則此客戶再度消費的機會愈高，因而在其最近購買日的指標分數也會較高；購買頻率(F:Frequency of the purchases)，是客戶在一定期間內的交易次數，當客戶的購買次數愈多時，則企業將認為其忠誠度與客戶價值較高；最後為購買金額(M:Monetary value of the purchase)，在一段時間內客戶購買該企業產品的總金額，一般而言，當客戶購買金額愈高時，則企業將認為其客戶價值較高。Hughes 指出，企業可從資料庫的分析當中，將客戶的交易資料轉算出客戶未來可能貢獻給企業的終身價值，也可透過客戶價值分析來預測未來的銷售狀況，並可依客戶的價值來決定如何分配行銷資源。 Arthur Hughes (1994) 對 RFM 模型指標進行之定義，本研究歸納整理如表 3。

表 3 RFM 指標定義及評量內容

指標項目	定義	評量內容
最近購買日 (Recency)	客戶距離現在的日期分析時點的最後一筆交易。	客戶再購率 客戶回應率
購買頻率 (Frequency)	在一定期間內，客戶購買企業產品或服務的交易次數。	客戶忠誠度 客戶價值
購買金額 (Monetary)	指一段期間內，客戶購買該企業產品的總交易金額。	客戶貢獻度 客戶價值

Kahan(1998)提及 RFM 模型是應用非常廣泛的行為分析技術，利用 RFM 模型可以更簡單、更快速來分析公司的客戶以獲得三項利益:包括增加客戶回應率、降低訂購成本及獲得利潤；具有以下之優點才得以被廣泛應用，第一、RFM 模型在獲取重要的客戶行為是具成本效益且易於客戶行為量化，第二、在預測未來客戶反應是有用的，也能在短時間內提高獲利，第三、透過少量的三項變數即可評估和預測消費者購買行為是相當有效的，第四、RFM 模型的資料是根據消費者歷史資料而來，不包含人口統計變數資料，故 RFM 模型針對特定的目標客群是更有意義的(Wei et al. 2010)。Goodman (1992)認為 RFM 分析可以使企業將獲利性高的客戶區隔出來，以便於將行銷資源作更有效的分配。

2.5 機器學習

機器學習(Machine Learning, ML)已經成為重要的資訊技術，其原理是指電腦模擬人類的學習行為，使獲得新的資訊或知識。讓機器可以自動學習、從巨量資料中找到規則，進而有能力做出預測，機器學習屬於人工智能的一門學科，其包括各種數學與計算機理論所發展的多項領域技術，人工智能的研究者研究開發一些演算法，想讓機器能像人一般思考，而資料探勘採用這些演算法(Jiawei et al. 2010；Stuart et al. 2010)。然而，前述的預測方法常無法得到良好的預測結果，主要原因在於這些方法都是假設資料為線性結構，但實務上的銷售資料很多是存在非線性的關係。機器學習是由五個基本元件所組成，分別為輸入資料、輸出資料、未知目標函數、訓練資料集與假說。機器學習的目標是要模擬連結輸入資料與輸出資料，所產生未知且複雜的目標函數，藉由觀察訓練資料集，使用數學演算法從一群候選函數中選出最合適的一個或多個函數，最後產出一假說，此假說將非常類似未知的目標函數(林軒田，2015)。因此，近年來，在機器學習方法中，由於擁有非線性及能在大量的樣本資料中找出隱含模式的能力，使其能在許多不同領域的實務預測問題上有良好的預測結果(Zhang et al. 1998；Hamzacebi et al. 2009；Singhal et al. 2011)。以上證明了迴歸演算法是一種預測實值結果的機器學習並是一個在預測相關領域中被廣泛使用的方法之一。

(一) 隨機森林(Random Forest)

隨機森林的核心精神為建立多個決策樹分類器，再由個別的決策樹分類器得到的分類結果取眾數決定。其作法主要將先隨機挑選資料集中部分或全部的特徵屬性，並隨機挑選部分或全部的資料建立決策樹分類器。其中，每個決策樹分類器將由不同的特徵屬性組合，不同的資料子集組合所建立，故每個決策樹分類器將可從不同的資料面向來進行分類和預測，藉以提升分類的正確率(Ho, 1998)，結合 Breimans 「Bootstrap aggregating 方法」

(Breimans, 2001)，於 2001 年所開發出的資料採礦與分類演算法則，屬於近年新興且複雜的程式語言。RF 其運算原理乃是選取 m 個子集合來預測決策樹的成長，然而根據訓練組的拔靴樣本決定每棵樹的成長，子集合 m 較所有自變數小得多，其功用是被用來分離結點。到目前為止，RF 已廣泛成功應用在各個不同的領域方面，包括生物資訊、預測績效、市場行銷研究、客戶關係管理等面向，皆獲得良好分類及預測成效，故本研究嘗試使用隨機森林，針對影響未來銷售狀況預測的變數進行探討，期望能建立一個精確的預測未來銷售狀況之模型。

(二) 邏輯式迴歸(Logistic regression)

邏輯式迴歸是用來分析依變項與自變項間之關係，主要使用於二元性資料，例如是或否。目的是為了要找出類別型態的反應變數和一連串的解釋變數之間的關係，邏輯式迴歸在分類只有兩類或少數幾類時，是為最標準的分析方法。邏輯式迴歸分析是一種分析二分類依變數資料受到一種或多種解釋變數影響之統計模型方法，而其分析特質在於類別型態之反應變數，最常見為二元反應變數(Walker, 2002)。

(三) 類神經網路(Neural Network)

類神經網路乃由大量人工神經元互相連結而組成，試圖模擬生物腦細胞訊號處理體系的資訊系統，其運算單元由類似生物腦神經元的簡單處理器構成(葉怡成, 2002)。要使得類神經網路能正確的運作，則必須透過訓練(training)的方式，使類神經網路反覆的學習，主要將建立輸入層、隱藏層、輸出層的人工神經網路，並將每個輸入的資料特徵屬性作為一個神經元，且每個輸出的資料作為一個神經元。在一個或多個隱藏層中建立多個神經元，再把每個神經元間進行連結，為每個連結建立權重值。最後，依資料集合來調整每個權重值，以讓輸入值經由激勵函數(activation function)和加權計算後，可以得到預期的輸出值。而當與輸出值不同時可以再回饋調整權重值，不斷讓類神經網路進行機器學習和調整，直到對於每個輸入都能正確對應到所需要的輸出(LeCun et al., 2015)。

3.研究方法及步驟

3.1 資料蒐集

本研究以一家專事外銷文具用品的製造廠商為個案研究，利用機器學習方法去建立一個精確、有效的預測未來銷售狀況之模型。其資料來源分為兩部份：一為外銷銷售歷史資料；另一為外銷客戶國家整體經濟環境因素包括匯率、通膨指數、國內生產總值、購買力平價和外部因素的年度學齡人口，為使本研究資料取得具官方及專業之正確性，銷售歷史資料是運用個案公司企業資源規劃系統(Enterprise resource planning, ERP)資料庫中過去2012年至2017年之外銷銷售歷史資料，共262筆；41個外銷客戶國家之總體經濟環境因素資料，分別為匯率、通膨指數、國內生產總值、購買力平價和外部因素的初等教育小學生年度學齡人口，其資料來源為World Bank Open Data。以下表格敘述影響未來銷售狀況預測模型之21項變數。

表 4 變數名稱與定義

變數	定義
目標變數	
預測年度與前一年度銷售量比較 (PS1)	上升=1；下降=0。
前一年度與前兩年度銷售量比較(PS2)	上升=1；下降=0。
前兩年度與前三年度銷售量比較(PS3)	上升=1；下降=0。
預測年度與前一年度匯率之比較(C1)	上升=1；下降=0。
前一年度與前兩年度匯率之比較(C2)	上升=1；下降=0。
前兩年度與前三年度匯率之比較(C3)	上升=1；下降=0。
預測年度與前一年度通膨指數之比較(I1)	上升=1；下降=0。
前一年度與前兩年度通膨指數之比較(I2)	上升=1；下降=0。
前兩年度與前三年度通膨指數之比較(I3)	上升=1；下降=0。

預測年度與前一年度國內生產毛額之比較(G1)	上升=1；下降=0。
前一年度與前兩年度國內生產毛額之比較(G2)	上升=1；下降=0。
前兩年度與前三年度國內生產毛額之比較(G3)	上升=1；下降=0。
預測年度與前一年度購買力平價之比較(P1)	上升=1；下降=0。
前一年度與前兩年度購買力平價之比較(P2)	上升=1；下降=0。
前兩年度與前三年度購買力平價之比較(P3)	上升=1；下降=0。
預測年度與前一年度學齡人口之比較(ST1)	每筆客戶國家的年度初等教育小學生人口總數。上升=1；下降=0。
前一年度與前兩年度學齡人口之比較(ST2)	每筆客戶國家的年度初等教育小學生人口總數。上升=1；下降=0。
前兩年度與前三年度學齡人口之比較(ST3)	每筆客戶國家的年度初等教育小學生人口總數。上升=1；下降=0。
最近一次購買金額(Recency)	每筆客戶從 2012 年至 2017 年，距離現在時間的最後一筆交易的消費總金額。
購買頻率(Frequency)	每筆客戶從 2012 年至 2017 年，累計五年的交易次數。
購買總額(Monetary)	每筆客戶從 2012 年至 2017 年，累計五年的消費總金額。

3.2 演算模型

過去已有許多研究利用機器學習來建構一預測模型，因此本研究參考先前的研究模型(Li, Chen, Zhang & Hai, 2018)研究模型，並採用機器學習的三種方法，依序分別為隨機森林(Random Forest)、邏輯式迴歸(Logistic regression)和類神經網路(Neural Network)，以建構一個精確、有效的外銷銷售狀況之預測模型。

3.3 評估模型

在機器學習方法中，分類是針對離散的資料建立一個分類器模型，透過分類器模型將數據依序分配到指定的類別中，常見分類器模型主要的評價指標有準確度(Accuracy)、精確率(Precision)、召回率(Recall)、F1 分數(F1-Score)、ROC 曲線以及 AUC 等。

(一) ROC 曲線

ROC 曲線的全名為 Receiver Operating Characteristic Curve (ROC Curve)，中文可譯為「接收器運作指標曲線」或「接收器操作特性曲線」。ROC 曲線是一條凸向左上方的曲線，在 ROC 曲線中，縱軸代表真正類率(True Positive Rate, TPR)；橫軸則代表負正類率(False Positive Rate, FPR)。ROC 曲線是一種二元分類模型，常用來評估二元變數中的分類表現，ROC 曲線下區域大小可用來評估其預測準度，(Mossman and Peng, 2014)。

(二) AUC

AUC(Area Under roc Curve)是由敏感度(x 軸)與 1-特異度(y 軸)所畫出的 ROC 曲線下的面積，用來判定模型分類(discrimination)能力的好壞，AUC 值愈大表示模型的分類能力愈好，AUC 值與模型分類能力為表 5 所呈現(張語恬等人，2007)。AUC 是目前常見的標準測量分類算法(Cortes & Mohri, 2004)。

表 5 AUC 值與模型分類能力

AUC 值	0.9~1.0	0.8~0.9	0.7~0.8	0.6~0.7	0.5~0.6
模型分類能力	很好	好	普通	差	很差

4.研究結果與分析討論

4.1 敘述性統計

在本研究中變數之統計分析若為 1，代表預測銷售之狀況為上升；若為 0 則代表預測銷售之狀況為下降。表 6 可見銷售量下降的客戶個數大於上升的客戶數。各變數在最小值方面除了消費總額，其他皆為 0。從消費總額最小值 158 到最大值 136000000，說明每筆客戶之消費能力相差甚大。

表 6 敘述性統計分析之結果

變數	最小值	最大值	平均數	標準差
預測年度與前一年度銷售量是否上升下降(PS1)	0	1	0.355	0.479
前一年度與前兩年度銷售量是否上升下降(PS2)	0	1	0.385	0.488
前兩年度與前三年度銷售量是否上升下降(PS3)	0	1	0.389	0.489
預測年度與前一年度匯率之比較(C1)	0	1	0.290	0.455
前一年度與前兩年度匯率之比較(C2)	0	1	0.679	0.468
前兩年度與前三年度匯率之比較(C3)	0	1	0.695	0.461
預測年度與前一年度通膨指數之比較(I1)	0	1	0.599	0.491
前一年度與前兩年度通膨指數之比較(I2)	0	1	0.469	0.500
前兩年度與前三年度通膨指數之比較(I3)	0	1	0.385	0.488
預測年度與前一年度國內生產毛額之比較(G1)	0	1	0.809	0.394
前一年度與前兩年度國內生產毛額之比較(G2)	0	1	0.733	0.443
前兩年度與前三年度國內生產毛額之比較(G3)	0	1	0.309	0.463
預測年度與前一年度購買力平價之比較(P1)	0	1	0.557	0.498
前一年度與前兩年度購買力平價之比較(P2)	0	1	0.366	0.483
前兩年度與前三年度購買力平價之比較(P3)	0	1	0.137	0.345
預測年度與前一年度學齡人口之比較(ST1)	0	1	0.590	0.493
前一年度與前兩年度學齡人口之比較(ST2)	0	1	0.477	0.500
前兩年度與前三年度學齡人口之比較(ST3)	0	1	0.553	0.498
最近一次購買(Recency)	0	3238410	61314.37	318053.9
購買頻率(Frequency)	0	252	11.340	29.339
購買總額(Monetary)	158	136000000	3909318	20398648

4.2 評估模式

(一) 預測性能指標

本研究利用 Microsoft Azure Machine Learning Studio 做出預測模型得出表 7，並列出預測性能指標，分別為準確率(Accuracy)及 F1 Score，透過機器學習方法其中三種預測模型，藉而求出最佳預期之結果。從表 7 可見隨機森林預測模型不論是準確率還是 F1 Score 的表現都非常優越，代表此預測模型準確度極高，在邏輯式迴歸和類神經網路預測模型下的結果，表現都不及隨機森林。

表 7 三種預測模型之各預測指標之結果

對策式比率	預測指標	邏輯式迴歸	類神經網路	隨機森林
0.75/0.25	準確率	0.662	0.689	0.949
	Score	0.000	0.371	0.924

(二) ROC 曲線

本研究利用三種分類性能指標及 ROC 曲線和曲線下的 AUC 數值判斷預測結果的優劣，以下為隨機森林、邏輯式迴歸以及類神經網路，分別將資料利用對策式比率 0.75/0.25 所得出的 ROC 曲線，如下圖 1、圖 2 和圖 3。由 ROC 曲線也可看出隨機森林預測模型最佳。

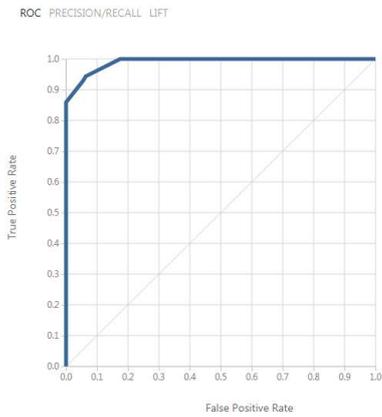


圖 1 0.75/0.25隨機森林之ROC曲線

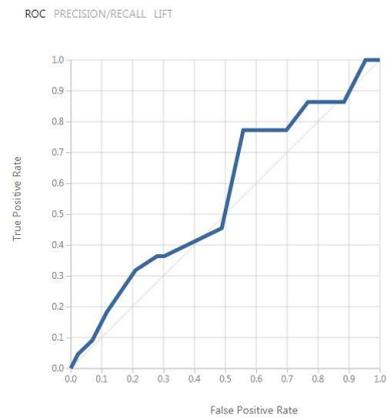


圖 2 0.75/0.25邏輯式迴歸之ROC曲線

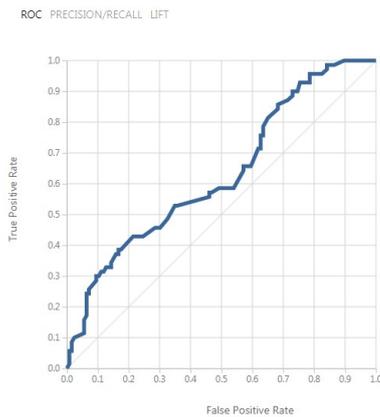


圖 3 0.75/0.25 類神經網絡之 ROC 曲線

(三) AUC

本研究中，AUC預測之結果整個值表現不錯，尤其預測模型以隨機森林的結果顯示最佳，準確度高達94.9%，代表本研究選擇的模型參數，提供一良好的預測模式。以下為隨機森林、邏輯式迴歸以及類神經網路，分別將資料利用對策式比率0.75/0.25所得出的AUC之結果，如表8。

表 8 三種預測模型之 AUC 之結果

對策式比率	預測指標	邏輯式迴歸	類神經網路	隨機森林
0.75/0.25	AUC	0.561	0.631	0.990

(三) 相關性分析

為了探討多項變數對銷售狀況之預測的相關性，本研究使用Kendall等級相關係數分析來決定兩變數之間的關係強度。Kendall等級相關係數又被稱作肯德爾等級相關係數(Kendall tau rank correlation coefficient)，是一種無母數分析(non-parametric analysis)，用以反映兩組變數之間關係密切程度的統計指標。

表 9 相關性分析結果比較

邏輯式迴歸相關性分析結果排名	Kendall等級相關係數結果排名
預測年度與前一年度學齡人口之比較(ST1)	1. 前一年度與前兩年度銷售量是否上升下降(PS2)
預測年度與前一年度購買力平價之比較(P1)	2. 預測年度與前一年度學齡人口之比較(ST1)
預測年度與前一年度國內生產毛額之比較(G1)	3. 最近一次購買(Recency)
預測年度與前一年度通膨指數之比較(I1)	4. 預測年度與前一年度國內生產毛額之比較(G1)
前一年度與前兩年度銷售量是否上升下降(PS2)	5. 預測年度與前一年度通膨指數之比較(I1)

由邏輯式迴歸及 Kendall 等級相關結果發現年度學齡人口增減為影響未來銷售狀況最重要因子，換句話說在文具產業中年度初等教育小學生人數的增加，可帶來銷售量的提升。其中本研究外銷國家以東南亞的緬甸及敘利亞為年度學齡人口最多之國家，至去年為止仍持續大幅度增加。本研究加入的外部總體經濟影響因素除了匯率以外，國內生產總值(Gross Domestic Product, GDP)包含通膨指數(Inflation Rate)和購買力平價(Purchasing Power Parity, PPP)，皆為影響預測未來銷售狀況之相關重要角色。Kendall 等級相關係數說明前一年度與前兩年度銷售量是升降(PS2)，與預測年度與前一年度銷售量升降最為相關。另若客戶最近一次購買金額與預測年度與前一年度銷售量升降相關亦高。

5. 結論與建議

商業模式中，行銷通路是銷售的重要管道，但影響客戶消費的因素諸多，為了增加獲利與優勢，如何讓企業能準確地掌握客戶需求、提高銷售量與收益，避免資源浪費在成交率低的客戶身上。本研究以銷售結果為主題，研究對象為一家專事外銷文具用品的製造廠商。研究結果可歸納出以下結論：本研究運用該公司企業資源規劃系統資料庫中過去外銷銷售歷史資料，並結合由公開網站所提供之外部總體經濟環境因素之資料，使用機器學習方法做預測，建構一預測未來銷售狀況之模型；整體而言，預測模型以隨機森林的結果顯示最佳，準確度高達 94.9%，代表本研究選擇的模型參數，提供一良好的預測模式。此預測模型亦可用於其他銷售行業，再加入其產業特別的影響因素，便可得出一幾乎可信的預測未來銷售狀況之結果，可供企業決策之判斷參考依據。

5.1 研究貢獻

本研究以一家導入企業資源規劃系統十年的專事銷售文具用品製造廠商為主要研究對象，結合歷史銷售資料及外部總體經濟因素，加上 RFM 模型中的三項衡量指標及產業相關之學齡人口變數，得出預測影響未來每筆客戶銷售量之重要因子，並建構出良好的預測能力之預測模型。此預測模型可供未來相關研究參考。

5.2 研究限制與未來研究之建議

由於本研究在統計樣本上資料之蒐集仍有些缺乏，未來若有更充足的樣本與影響變數，或是加入三種以上機器學習方法以做比較，可以更加提升預測模型之準確性。

6. 文獻參考

- 范秉航(2006)，金磚四國股市波動對全球股市之衝擊，臺灣經濟研究月刊，29(8)，36-46。
- 林軒田(2015)，機器學習技術，coursera。
- 劉向上(1984)，「如何從事銷售預測」，行銷月刊，第 86-91 頁。
- 徐守德，李鎮旗(1994)，「企業銷售預測之方法與實證研究—以台電公司為例」，管理評論，第十三卷，第一期，第 23-56 頁。
- 張語恬、朱基銘、簡戊鑑、周雨青、楊燦、盧瑜芬、白健佑、白璐、Thomas Wetter、孫建安、羅慶徽(2007)，台灣家庭醫學雜誌，17(4)，222-238。
- 葉怡成(2002)，類神經網路模式應用與實作，台北：儒林圖書。
- 王平(2013)，企業庫存管理中心-需求預測的主要問題及對策，市場行銷，(02):13-15。
- Aburto, L., & Weber, R. (2007). Improved supply chain management based on hybrid demand forecasts. *Applied Soft Computing*, 7(1), 136-144.
- Atkinson, A., Barnett, S., Gorte, R. J., Irvine, J. T. S., McEvoy, A. J., Mogensen, M., ... & Vohs, J. (2011). Advanced anodes for high-temperature fuel cells. In *Materials For Sustainable Energy: A Collection of Peer-Reviewed Research and Review Articles from Nature Publishing Group* (pp. 213-223).
- Barandiaran, I. (1998). The random subspace method for constructing decision forests. *IEEE transactions on pattern analysis and machine intelligence*, 20(8).

- Black, K. (2011). *Business statistics: for contemporary decision making*. John Wiley & Sons.
- Breiman, L. (2001). Random forests. *Machine learning*, 45(1), 5-32.
- 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(5), 344-355.
- Chang, T., & Caudill*, S. B. (2005). Financial development and economic growth: the case of Taiwan. *Applied economics*, 37(12), 1329-1335.
- Chopra, S., & Meindl, P. (2007). Supply chain management. Strategy, planning & operation. In *Das summa summarum des management* (pp. 265-275). Gabler.
- Cortes, C., & Mohri, M. (2004). AUC optimization vs. error rate minimization. In *Advances in neural information processing systems* (pp. 313-320).
- Dalrymple, D. J. (1987). Sales forecasting practices: Results from a United States survey. *International Journal of Forecasting*, 3(3-4), 379-391.
- Derynck, R., Zhang, Y., & Feng, X. H. (1998). Transcriptional activators of TGF- β responses: Smads. *Cell*, 95(6), 737-740.
- Efendigil, T., Önüt, S., & 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(3), 6697-6707.
- Geurts, M. D., & Whitlark, D. B. (1999). Six ways to make sales forecasts more accurate. *The Journal of Business Forecasting*, 18(4), 21.
- Doganis, P., Alexandridis, A., Patrinos, P., & Sarimveis, H. (2006). Time series sales forecasting for short shelf-life food products based on artificial neural networks and evolutionary computing. *Journal of Food Engineering*, 75(2), 196-204.
- Ghartey, E. E. (1993). Causal relationship between exports and economic growth: some empirical evidence in Taiwan, Japan and the US. *Applied Economics*, 25(9), 1145-1152.
- Goh, C., & Law, R. (2002). Modeling and forecasting tourism demand for arrivals with stochastic nonstationary seasonality and intervention. *Tourism management*, 23(5), 499-510.
- Steckler, A., McLeroy, K. R., Goodman, R. M., Bird, S. T., & McCormick, L. (1992). Toward integrating qualitative and quantitative methods: an introduction.
- Hamzaçebi, C., Akay, D., & Kutay, F. (2009). Comparison of direct and iterative artificial neural network forecast approaches in multi-periodic time series forecasting. *Expert Systems with Applications*, 36(2), 3839-3844.
- Holt, C. C. (2004). Forecasting seasonals and trends by exponentially weighted moving averages. *International journal of forecasting*, 20(1), 5-10.
- Kahan, R. (1998). Using database marketing techniques to enhance your one-to-one marketing initiatives. *Journal of Consumer Marketing*, 15(5), 491-493.
- Kshemkalyani, A. D., & Singhal, M. (2011). *Distributed computing: principles, algorithms, and systems*. Cambridge University Press.
- Kuo, R. J., & Xue, K. C. (1999). Fuzzy neural networks with application to sales forecasting. *Fuzzy Sets and Systems*, 108(2), 123-143.
- Kotler, P., & Armstrong, G. (1999). *Principles of marketing*.
- LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *nature*, 521(7553), 436.
- LeVee, G. S. (1993). The key to understanding the forecasting process, *Journal of Business Forecasting*, 11 (4), 12-16.
- Medina-Mora, R., Winograd, T., Flores, R., & Flores, F. (1993). The action workflow approach to workflow management technology. *The Information Society*, 9(4), 391-404.
- Pedregal, D. J., & Young, P. C. (2008). Development of improved adaptive approaches to electricity demand

- forecasting. *Journal of the Operational Research Society*, 59(8), 1066-1076.
- Lim, C., & McAleer, M. (2002). Time series forecasts of international travel demand for Australia. *Tourism Management*, 23(4), 389-396.
- Mossman, D., & Peng, H. (2014). Constructing "proper" ROCs from ordinal response data using weighted power functions. *Medical Decision Making*, 34(4), 523-535.
- Reichheld, F. F., & Teal, T. (1996). *The Loyalty Effect: The Hidden Force Behind Growth, Profits, and Lasting Value*.
- Schrieber, J. (2005). Demand visibility improves demand forecasts. *The Journal of Business Forecasting*, 24(3), 32.
- Still, Richard R., W.C. Edward & A.P., (1988). *Govoni Norman, Sales Management*, Prentice Hall.
- Stuart, M. A. C., Huck, W. T., Genzer, J., Müller, M., Ober, C., Stamm, M., ... & Winnik, F. (2010). Emerging applications of stimuli-responsive polymer materials. *Nature materials*, 9(2), 101.
- Tay, F. E., & Cao, L. (2001). Application of support vector machines in financial time series forecasting. *omega*, 29(4), 309-317.
- Taylor, F. W. (2004). *Scientific management*. Routledge.
- Tsai, T. H., Lee, C. K., & Wei, C. H. (2009). Neural network based temporal feature models for short-term railway passenger demand forecasting. *Expert Systems with Applications*, 36(2), 3728-3736.
- Vandevivere, P. C., Bianchi, R., & Verstraete, W. (1998). Treatment and reuse of wastewater from the textile wet-processing industry: Review of emerging technologies. *Journal of Chemical Technology & Biotechnology: International Research in Process, Environmental AND Clean Technology*, 72(4), 289-302.
- Wenger, E., McDermott, R. A., & Snyder, W. (2002). *Cultivating communities of practice: A guide to managing knowledge*. Harvard Business Press.
- Winklhofer, H., & Diamantopoulos, A. (2003). A model of export sales forecasting behavior and performance: development and testing. *International Journal of Forecasting*, 19(2), 271-285.
- Zhang, G., Patuwo, B. E., & Hu, M. Y. (1998). Forecasting with artificial neural networks:: The state of the art. *International journal of forecasting*, 14(1), 35-62.