

## 研究論文

# 網路算命使用者 行為與特徵分析： 資料探勘技術之應用\*

帥嘉珍\*\* 陳杏枝\*\*\* 陳耀斌\*\*\*\*

---

### 摘要

命理在華人社會中一直有很大的影響力，近年來隨著使用網際網路的人口快速增長，在台灣命理相關網站日漸風行。本研究使用資料探勘工具，針對台灣民眾使用網路算命行為，分析網路算命使用者的個人特質、使用行為以及使用者動機。本研究以約略集合理論（Rough Set Theory）、決策樹（Decision Tree）與貝氏網路（Bayesian Net-

---

\* 本文使用資料部分係採自行政院國科會資助之「台灣社會變遷基本調查計畫：第四期第五次」計畫（NSC 93-2420-H-001-002-B1）、「台灣社會變遷基本調查計畫：第五期第五次」計畫（NSC 97-2420-H-001-001-B1）、「台灣社會變遷基本調查計畫：第六期第五次」計畫（NSC 102-2420-H-001-007-SS2）。上述計畫係由中央研究院社會學研究所執行，計畫主持人為傅仰止、章英華、杜素豪及廖培珊四位教授。該資料由中央研究院人文社會科學研究中心調查研究專題中心學術調查研究資料庫釋出。作者感謝上述機構及人員提供資料協助，然本文內容由作者自行負責。

\*\* 通訊作者。明新科技大學資訊管理系副教授，E-mail: jjshuai@must.edu.tw，新竹縣新豐鄉新興路1號，03-5593142 分機 3440。

\*\*\* 淡江大學通識教育中心教授，E-mail: hsinchih@mail.tku.edu.tw。

\*\*\*\* 明新科技大學資訊管理研究所碩士，E-mail: k10639@yahoo.com.tw。

work) 的資料探勘方法來進行分析, 希望能區隔民眾使用網路算命之特徵、偏好, 以瞭解台灣民眾使用網路算命的趨勢。研究發現: 網路算命使用者與其個人宗教信仰以及網路宗教的使用密切關係, 有朝向上班族及高學歷兩種趨勢, 本研究結果將有助於增進對網路算命研究課題之瞭解, 同時也可以提供網路業者做為經營參考。

關鍵詞: 網路算命、資料探勘、約略集合理論、決策樹、貝氏網路

## **Characteristics and Behavior Analysis of Online Fortune Telling Users—A Data Mining Approach**

**Jia-Jane Shuai\*, Hsinchih Chen\*\*, Yao-bin Chen\*\*\***

### **ABSTRACT**

Fortune telling has always been influential in Chinese society. In recent years, internet users have increased considerably in number. This has led to attention to internet-related behavior and characteristics. This study aims to understand the motivation and behavior pattern for Taiwanese users of on-line fortune telling. This paper proposes a method that links the rough set theory, Bayesian network and decision tree modeling for causal analysis. According to our research, usage of on-line fortune telling is related to users' religion and online religion experience. On the other hand, there are a growing number of well-educated people and full-time workers who use online fortune telling. Data mining results for targeting marketing to on-line users can offer Internet enterprises information for marketing decision making. Based on the findings of this study, conclusions and implications for management are discussed.

---

\* Associate Professor, Department of Information Management, Minghsin University of Science and Technology.

\*\* Professor, Center for General Education and Core Curriculum, Tamkang University.

\*\*\* Master of Information Management, Department of Information Management, Minghsin University of Science and Technology.

**Keywords:** online fortune telling, Data Mining, Rough Set Theory, Bayesian network, Decision Tree

---

## 一、緒論

命理在華人社會中一直有很大的影響力，算命看相的行為在台灣社會非常普遍，根據 Yes123 求職網 2015 年「羊年職場命理調查」調查 1,348 名上班族結果顯示，78.6% 的上班族表示，自己曾經算過命，其中有 11.3% 的上班族透露，算命的次數在 10 次以上，整體平均值則是 4.2 次。調查也顯示，有 63.5% 的上班族表示，在農曆年前後，會到廟裏進行點光明燈與安太歲的儀式。而上班族算命或吸收命理資訊的來源以「網路」最高，占 63.64%。（Yes123 求職網 2015）1111 人力銀行 2007 年進行的網路調查「企業求才命理大調查」結果顯示，1,227 名企業主或人資主管中，34.7% 面試或過濾履歷時，會在基本資料中找尋命理線索，最常參考的依據是星座（81.99%）、面相（55.63%）、血型（21.13%）（1111 人力銀行 2007）。

根據台灣網路資訊中心（Taiwan Network Information Center, TWNIC）公布的「2014 年臺灣寬頻網路使用狀況調查」指出，台灣地區曾經上網比例有逐年增加的趨勢，到 2014 年 5 月分為止，台灣上網人口超過 1,700 萬。其中 12 歲以上的民眾曾經上網比例已經超過了七成七。使用行動及無線區域網路上網者皆有大幅度成長，從 2012 年的 25.91% 成長為 2014 年的 47.27%。隨著網路科技之普及，算命開始數位化，線上算命服務發展迅速，不只專業算命網站紛紛成立，許多入口網站、內容網站及電信業者更是陸續推出相關服務，也帶動網

路算命的風潮。

本研究試圖瞭解是哪些人在使用網路算命？使用哪一種傳統算命方式的人比較容易轉向網路算命？以及使用算命的動機和網路算命的關係為何？

中央研究院於 1983 年開始推動「台灣社會變遷基本調查」資料庫的建置，以間隔五年為原則，從事實時性之調查，是台灣地區社會科學研究領域具代表性的資料庫。由於提供良好可靠且豐富的資料，吸引許多研究者使用。然而大多數的研究論文之統計分析方法大多採用一般線性模型，如迴歸分析、結構方程模式等來進行資料分析，但由於線性模型必須符合許多假定，又由於資料庫中的變項，包括連續與間斷變項，其間的關係有可能是非線性關係，或是變項之間有嚴重的共線性問題，如果使用一般線性模型，就無法精確解釋變項間複雜的交互作用關係。

本研究使用 2004 年、2009 年及 2014 年「台灣地區社會變遷基本調查」的資料（章英華、傅仰止 2004；傅仰止、杜素豪 2009；傅仰止等 2015），透過資料探勘（data mining）的方式，探討不同時間點底下相關因素的變化情形。先以約略集合理論（Rough Set Theory）做屬性篩選，再利用貝氏網路（Bayesian Network）來分析算命行為、算命動機的分類與預測，最後以決策樹（Decision Tree）來分析網路算命使用者的特徵規則。不需大量的樣本資料、也不需進行假設統計即可以對資料進行分析，從真實資料中，藉由機器學習，歸納分類規則以發掘網路算命使用者的行為模式，進而獲得知識性之規則，使用的彈性較大。

本文之內容安排為：第二節為網路算命之文獻探討，第三節研究方法中說明約略集合理論、決策樹與貝氏網路之理論基礎，第四節研

究發現說明資料來源以及樣本基本資料分析，針對個人背景資料、算命行為與算命動機以決策樹來分析出核心屬性與決策規則，第五節為本研究之結論與討論。

## 二、文獻回顧

所謂的網路算命（Online Fortune-Telling），是將傳統的算命方式加入了網路元素，而台灣網際網路的發達也是近二十年來的事，連帶地，網路算命是比較晚近的現象。有關網路算命的議題，還未得到太多研究者的關注，目前可蒐得的研究論文，大部分是碩士學位論文（張國偉 2003；薄懷武 2009；林彥錡 2009；魏劭楷 2010；劉素君 2012）。

首先，本文把所蒐得的網路算命相關論文之研究主題和研究方法，做一簡要的說明。郭貞（2004）針對兩岸三地的 348 個算命網站做內容分析，以求瞭解網站的收費狀況和提供的服務種類；訪談 10 名網站的經營者來瞭解網站經營模式；以及在 4 個算命網站上招募了 6,088 位網友填答問卷。這項問卷調查的項目包括人口變項、人格特質、網路算命動機、瀏覽算命網站次數、滿意度、對網路即時算命的相信程度、對大師親算的相信程度、自認算命結果對行為的影響程度。這份問卷調查的研究成果，另有專文發表（郭貞 2005；Kuo 2009）。陳建榮（2004）想要初步瞭解學生使用網路算命的情形，但其研究並沒有依據一般問卷調查的要求來執行，只是隨意調查臺灣師範大學和崇右企專學生 549 名。張國偉（2003）邀請 20 名受訪者（其中 10 名為學生），先到某算命網站註冊，由大師親算後，再訪談他們，以求瞭解網站要如何經營才能得到顧客的信任感。薄懷武（2009）則以台

灣的 2 個算命網站做內容分析，並對 9 名有網路算命經驗者做深度訪談，來瞭解網路算命者如何經由一連串的心理機制而產生可信的這一過程。林聶錡（2009）關注的是哪些因素影響民眾的網路命理行為（包含算命、占卜、點燈、求籤、祭拜等）之參與意願，而其所謂的「行為」也包括在網路上搜尋相關訊息。該研究在網站、部落格、電子布告欄上發放網路問卷，取得有效問卷 406 份。魏劭楷（2010）則是想要瞭解一般民眾是如何看待網路算命和傳統算命，以及為何會想要使用網路算命。該研究是在算命網站的討論區發出訪談邀請函，挑出 14 名經常使用網路算命的網友，以半結構的線上訪談來蒐集資料。劉素君（2012）研究西洋占星／星座算命網站的使用者行為，試圖瞭解使用動機、網站信任度、付費意願等因素對使用者滿意度的影響。她在建構問卷之前，挑選台灣和中國大陸有網路算命經驗者各 10 名，以初步瞭解台灣和中國大陸的消費者在網路算命的現況。然後，設計線上問卷，在 Facebook、RC 語音和新浪微博上發布，收得有效問卷 518 份。

接下來，本文針對這些論文的研究發現和本文的研究旨趣有關聯的部分，做一綜合說明。一般而言，網路算命可分為兩種：一是使用線上算命程式，命主輸入個人的基本資料，例如：姓名、性別、生日等，命主會立即收到算命的結果；另一種是有真人算命師的服務，這和傳統的算命類似，只是把網路當成媒介，命主不需親自去算命館算（魏劭楷 2010: 33-34；郭貞 2004: 1）。一般由大師親算的服務需要支付費用。許多算命網站的運作，是同時結合這兩種網路算命的形式（郭貞 2004: 1-2；劉素君 2012: 40-42）。

魏劭楷（2010）從訪談研究中得出這樣的結論：相信算命會準是因為對算命師產生信任感，而信任感是來自命主和算命師一對一的互動，尤其是面對面的互動；因此現階段的網路算命無法取代傳統算

命，真正遇到困難時，還是尋求傳統的算命。他認為網路算命是「能夠滿足某部分人的需求，其中包含想要尋求消遣的使用者」（魏劭楷 2010: 112），言下之意，使用線上算命程式的人，大概是消遣娛樂的成分居多。確實有些人在有些時候是把使用線上算命程式，當成消遣、滿足好奇心的活動，一位網路算命使用者是如此表示：「連妳無聊的時候妳都還會連上電腦，好奇看看我今天的運勢如何之類的」（顏任儀 2012: 110）。

雖然網路算命有些時候有其消遣的成分在內，但大部分的人在大部分的時候不見得是以消遣、好玩的動機在使用網路算命。郭貞（2005）的大樣本網友調查結果顯示，影響網友相信線上程式即時算命的各種因素中，動機因素的影響力最為明顯，其重要性依序為：解決困境、增進瞭解、網路優勢。好奇、好玩的動機則呈現出負面的效果，也就是抱持好奇、好玩動機的人，比較不相信線上程式即時算命。綜合來看，網友使用並相信線上程式即時算命是有其嚴肅的動機，而這些嚴肅動機的強度和重要性也同樣可見於相信網站大師親算的模式中（郭貞 2005: 166-168）。林聶錡的網路問卷中（2009: 50）有一道題目：「基於何種因素而考慮參與命理行為」，前三名依序為愛情或婚姻（占全體的 70.9%）、事業／工作（66.3%）、瞭解自己個性（61.1%）。林聶錡是以百分比來呈現，我們無法得知問卷填答者是如何同時勾選兩項以上的動機。「愛情或婚姻」和「事業／工作」，在上述郭貞的研究中則歸屬於「解決困境」一類，可能也是擔心在使用迴歸線性統計分析的「共線性」問題，而不再區分「愛情或婚姻」和「事業／工作」等細項。

再者，人們使用網路算命時會挑選何種算命項目或形式？陳建榮在兩所大學學生的調查顯示，學生最常使用的網路算命項目，百分比



最高的前五名依序爲：星座 (91.9%)、姓名學 (68.5%)、生肖 (60.0%)、八字 (37.5%)、占星／星座 (36.4%) (陳建榮 2004)。薄懷武則是計算雅虎奇摩「知識+」的算命類別之五子項 (星座血型、塔羅占卜、命理風水、紫微斗數、其他) 的算命使用筆數。在 2006 年 9 月一個月的筆數爲 2,085，其中命理風水占全體的 37.5%、星座血型 35.2%、紫微斗數 19.8%、塔羅占卜 4.4%、其他 3.1% (2009: 52-53)。這筆資料和陳建榮對學生的調查相比較，除了星座、占星／星座可確定是網路算命中較爲流行的項目之外，我們無法得知其他的哪一個算命方式也是網路算命的流行項目。與此問題相關的另一問題是：使用哪一種傳統算命方式的人，比較容易轉向網路算命？這些議題都值得研究者關注。

最後，本文關心的是什麼樣的人會使用網路算命？這些人具有什麼樣的人口變項特徵。郭貞的六千多名的網友背景之研究，女性占樣本全體的 69.1%。年齡分布以 20-29 歲最多 (57.0%)、其次分別是 30-39 歲 (24.1%)、13-19 歲 (12.9%)、40-49 歲 (4.7%)。職業分布，以學生最多 (34.0%)、其次是電信通訊業者 (10.2%) 和工商服務業者 (10.1%)。教育程度方面以大專畢業和肄業者居多 (60.4%)、其次是高中畢業和肄業者 (25.7%)。婚姻狀況以單身未婚者最多 (78.8%)、其次是已婚有子女者 (12.5%) (郭貞 2004: 3)。

林聿錡回收的有效網路問卷 406 份，其中女性占全體樣本的 67.2%。年齡分布則以 20-24 歲居多占 45.3%，其次是 25-29 歲 (30.0%)，再來是 15-19 歲 (15.5%)。教育程度以大專、大學最多 (70.0%)，其次是研究所以上 (19.0%)。職業分布，學生占 59.1%，其次是一般工、商、服務業 (16.5%)。有 63.8% 的樣本每月可支配的費用在 5,000 元以下 (林聿錡 2009: 47-48)。劉素君也是在網路上發布問



卷，有效問卷為 518 份。男性占 53.3%。未婚者占 89.4%。年齡方面，以 15-24 歲居多（58.1%），其次是 25-34 歲（24.3%），再來是 14 歲以下（9.9%）。教育程度以大學居多（39.6%），其次是高中職（34.6%），再來是國中（14.9%）。職業以學生身分最多（61.0%），其次是餐飲服務（6.8%）。每月可支配所得，5,000 元以下占 53.3%，其次是 20,001 元以上（19.31%），再來是 5,001-10,000 元（15.3%）（劉素君 2012: 46-47）。

劉素君的網友樣本，是以男性為多，這與郭貞以及林聶錡的樣本有很大的出入。這樣的結果，不知是否是因使用西洋算命網站的人真的是以男性居多？還是因為其他問卷發放的缺失所導致。其次，這三個研究樣本都是直接在網路上招募，其很難避免的研究缺失之一是，參與填答的人會偏向學生族群，因為他們有較多的空閒，也比較會被參與填答即可參加抽獎、贈送點券的邀請說詞所吸引。要避免這問題，大概只能藉助符合嚴謹抽樣程序的全國性的調查。

瞿海源曾使用 1985 年、1990 年和 1995 年的「台灣社會變遷基本調查」資料來研究台灣民眾的術數行為，其中一項術數行為即是算命。雖然其研究的依變項是傳統的算命行為，但因為使用的資料庫也是本研究所使用的，值得我們參考。該研究使用邏輯迴歸分析，統計結果顯示，女性、信仰非佛非道的民間信仰、教育程度較高者算命的機率較大；職業為國防和警察，以及學生算命發生的機率較小（瞿海源 1999: 19）。和前面的網路算命文獻相比較，我們可看到不論是網路或傳統的算命，女性和教育程度較高者都較容易受到吸引。其次，宗教信仰也是影響算命行為的重要變數，之前的網路算命文獻並未關注到此變數。還有，以前的學生族群排斥傳統算命，現在的學生族群卻是接受網路算命。或許是因現在的學生從小就使用網路，他們大部

分是經由網路接觸宗教信仰相關訊息。正如 Lövheim (2008) 在瑞典的社群網站 Lunar Stotm 上的問卷調查所發現，瑞典青少年較常從網路上接觸宗教活動與訊息，而非從社區教會。

最後，本文特別著墨選取適當統計方法的重要性。郭貞在檢視人口變項對網路算命行為的影響時，原本打算所有的變項是以多元迴歸的統計方法來處理。所以她一開始先把婚姻狀況，和性別、年齡、教育程度、收入一起納入多元迴歸分析中，卻無顯著效應。她覺得這樣的統計結果有違其研究假設，因此只好挑出婚姻狀況一項，以單因子變異量分析來檢視婚姻狀況對網路算命的相關行為變項之影響。其他的人口變項（性別、年齡、教育程度、收入）才放入多元迴歸分析中（郭貞 2005: 159-162）。爲了避免變數之間的相互干擾，而把某些變數做單獨分析，這樣的切割處理是較為謹慎，但也使得讀者容易迷失在多張的統計表格中，而無法很快抓住全貌。再者，在使用一般線性模型時，又同時想要觀察多個變項的影響力，需要相當大的樣本，若在調查研究經費不足的情況下，所謂初探性的研究也是很難展開。

針對上述所提的一些限制，本文嘗試從資料探勘的角度切入，採用資料探勘中的約略集合理論、決策樹與貝氏網路做為分類的工具，此三種分類方法已應用在許多領域上，例如教育（江羿臻、林正昌 2014；鄒小蘭、王珣棻 2015）、運輸（吳怡瑾等 2014）及管理（Kuzey et al. 2014；Fu et al. 2015；Wu et al. 2010；帥嘉珍等 2009；Shuai and Li 2005）方面，能夠同時處理連續與類別變項的資料，並提供預測變項的重要性。其優勢在於不需要統計假設與大量的資料，可直接分析原始的資料，從而挖掘出隱藏在資料背後的重要意涵，並產生易懂而直觀的決策規則。

### 三、研究方法

本研究使用中央研究院所主持的「台灣地區社會變遷基本調查」資料，並且爲了比較不同時間點的差異，選取了 2004 年（第四期第五次）、2009 年（第五期第五次）以及 2014 年（第六期第五次）三個年度。分別有 1,730、1,903、1,934 筆資料，其中有使用過網路算命且資料完整者分別有 236、307、299 筆資料。

首先對於不同年度的網路算命使用者進行交叉表描述統計，以約略集合理論進行屬性篩選，進一步則以貝氏網路來做算命行為與算命動機的因果關係與機率分析，最後以決策樹針對個人背景資料屬性產生萃取規則。

#### （一）變項選取

2004 年的變遷調查有個人背景資料、算命行為、算命動機等三種類型資料；而 2009 年則有二種類型資料，爲個人背景資料、算命行為；2014 年則只有個人背景資料，如表 1 所示。

表 1 網路算命屬性資料年度表

屬 性 名 稱	2004	2009	2014
性別、年齡、職業狀況、月收入、婚姻狀況、教育程度、宗教信仰	✓	✓	✓
最近一年有沒有主動找人算八字	✓	✓	
最近一年有沒有主動找人算紫微斗數	✓	✓	
最近一年有沒有主動找人算手相、面相	✓	✓	

表 1 網路算命屬性資料年度表 (續)

屬 性 名 稱	2004	2009	2014
最近一年有沒有主動找人摸骨	✓	✓	
最近一年有沒有主動找人占星／星座算命	✓	✓	
算命是爲了子女管教問題	✓		
算命是爲了求學問題	✓		
算命是爲了財運問題	✓		
算命是爲了事業問題	✓		
算命是爲了婚姻感情問題	✓		
算命是爲了健康問題	✓		
算命是爲了住的問題	✓		
算命是爲了預測自己的運氣	✓		
算命是爲了瞭解自己的命運	✓		
透過網路搜尋及閱讀宗教／信仰相關的訊息			✓
有沒有在網路上算過命	✓	✓	✓

1. 被預測變項：「有沒有在網路上算過命」定義爲被預測變項。

2. 預測變項

(1) 算命動機：只有 2004 年的變遷調查中有算命動機，包含子女管教、求學問題、財運問題、事業問題、婚姻感情問題、健康問題、住的問題、預測自己的運氣、瞭解自己的命運等 9 項動機。

(2) 算命行爲：2004 及 2009 年問卷中的算命行爲包含了八字、紫微斗數、手面相、摸骨、占星／星座等 5 個行爲。

(3) 個人背景資料：在文獻回顧中有提到，以往網路算命相關的研究皆是以基本的人口變數爲主，因此，個人背景資料我們選取了性別、年齡、職業狀況、月收入、婚姻狀況、教育程度與宗教信仰等 7 個背景屬性，其類型定義如表 2。此外在 2014 年的問卷中，新增一題，詢問透過網路搜尋及閱讀宗教／信仰相關的訊息頻率，爲了瞭解網路宗教與網路算命的相關性，我們亦將此變項重新編碼列入情境屬性。

上述的變項，除了個人背景屬性之外，其餘變項皆是「有」和「沒有」的二元編碼。

表 2 個人背景屬性選取表

個人背景資料	
屬性名稱	類 型 定 義
1. 性別	(1)男 (2)女
2. 年齡	(1) 25 歲以下 (2) 26-45 歲 (3) 46-65 歲 (4) 66 歲以上
3. 職業狀況	(1)已就業（專業、半專業與其餘有工作等） (2)未就業（家庭主婦、無業等） (3)學生
4. 月收入	(1) 3 萬元以下 (2) 3-6 萬元 (3) 6 萬元以上
5. 婚姻狀況	(1)已婚 (2)未婚 (3)已離婚 (4)配偶已去世
6. 教育程度	(1)國中以下 (2)高職或高中 (3)專科 (4)大學以上
7. 宗教信仰	(1)沒有 (2)民間、佛、道教 (3)天主、基督 (4)其他

## （二）資料分析方法

本研究使用的約略集合理論、決策樹與貝氏網路在此做一介紹。

### 1. 約略集合理論

約略集合理論(Rough Set Theory, RST)在1982年由波蘭的Zdzisław Pawlak教授所提出(Pawlak 1982)。約略集合演算法具備強大的資料庫分析能力，可輕易進行移除重複紀錄、簡化屬性、發掘屬性關係、推演法則等操作。從模糊或不完整的大量資料中挖掘出資料屬性彼此之間的規則，並能從資料中擷取重要的核心屬性。跟傳統的統計方法不同的是，在進行約略集合理論分析時，不需做任何的假設，以及受到樣本資料量的限制。因此，經過多年的發展，RST已經被各個領域所成功的應用，例如決策分析(Decision Analysis)、資料庫知識發掘

(Knowledge Discovery from Databases)、專家系統 (Expert Systems) 等 (Pawlak 2002)。對於任何模糊不清的資訊，都可藉由近似 (Approximation) 的基本概念來進行辨別，是一種處理資料的分類方法。因此，本研究利用約略集合理論，希望得知有使用網路算命者的核心屬性與特徵。以下說明約略集合理論整體資料的處理程序。

(1) 資訊表

對於所要處理的資料，通常會以一個資訊表來呈現，資訊表的每一行必須要放置被觀察或被描述的屬性值 (attribute)，並以每一列的表格做為一個事件 (event)。屬性分為兩種：被觀察或被描述的為條件屬性 (condition Attributes)，用來決定事件的屬性為決策屬性 (decision Attributes)。

(2) 上下界近似關係與分類

在資料表中，RST 是建構在兩個基本集合上，如圖 1 所示，分別為下界近似集合 (lower approximation sets) 與上界近似集合 (upper approximation sets)。下界近似集合表示該集合的元素一定是屬於在該

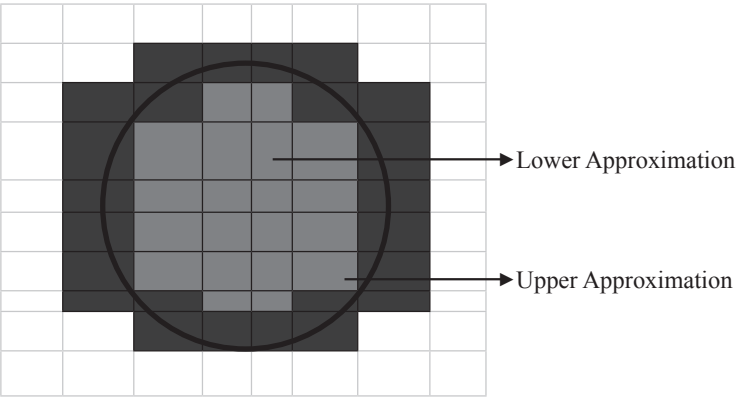


圖 1 上界與下界近似關係 (Walczak and Massart 1999)

集合裏，而上界近似集合表示該集合的元素有可能屬於在該集合裏。另外，無法透過上、下限近似集合來加以定義的區域，我們稱之為邊界（Walczak and Massart 1999）。

### (3) 屬性的核心與折減

在資訊表中，往往會有很多的條件屬性，處理起來很複雜，因此可以透過屬性的折減（reduct），去除較不重要的條件屬性，刪減這些屬性並不會影響到整個的決策品質。如果能加以折減，不僅能降低系統的複雜度，更可以讓決策者瞭解具有影響的變數有哪些。而屬性的核心（core），是資訊表經過簡化後最重要的屬性。

## 2. 貝氏網路

貝氏網路（Bayesian Network, BN）是結合貝氏機率理論與圖形模式所定義的知識表示方式，特點為重視對於先驗資訊（Prior Information）的蒐集、挖掘與量化。貝氏網路可以透過機率理論與圖形模式來說明變數間彼此的因果關係，貝氏網路最大的特徵是它的學習能力，可以利用統計的方法來做分類，其中貝氏機率理論是一種適應性的學習（adaptation），即為條件機率分配。貝氏網路能在不確定的情況下，發現到變數彼此之間關係的結構。因此，本研究利用貝氏網路，希望得知使用網路算命網友的機率分配與屬性跟屬性之間的因果關係。

貝氏定理的描述如下：假設  $X_1, X_2, X_3, \dots, X_n$  為樣本空間（Sample Space） $S$  的一個分割，且有一事件（Event） $A$ ，則在  $A$  發生的情況下， $X_i$  會發生的機率可以以下列公式表示：

$$P(X_i|A) = \frac{P(X_i)P(A|X_i)}{P(A)} = \frac{P(X_i)P(A|X_i)}{\sum P(X_i)P(A|X_i)}$$

其中： $P(X_i)$ ：事前機率（Prior Probability）



$P(A|X_i)$ ：樣本機率 (Sample Probability)

$P(X_i|A)$ ：事後機率 (Posterior Probability)

由上式可以看出，貝氏定理最大的優點即是結合事前機率與樣本機率，只要利用以前的經驗（即事前機率），不需要太多的資料即可進行分析，比起一般的統計方法，貝氏定理能更有效地運用樣本的資訊。

貝氏網路可以結合過去的統計資料來說明節點與節點的關係，能夠使用在不確定性的資料中。本研究貝氏網路所使用的分類方法為樹狀貝氏分類器 (Tree Augmented Naive Bayes, TAN)，是為簡易貝氏分類器 (Naive Bayes) 的擴展。與簡易貝氏分類器不同的是，簡易貝氏分類器為屬性間彼此條件獨立，而樹狀貝氏分類器加入了相依的屬性。換言之，即允許更多屬性之間的依賴關係。樹狀貝氏分類器具有下列優點 (Friedman et al. 1997)：

(i) 能產生因果關係圖除了決策屬性，在圖中的條件屬性中，越上層的屬性越是重要 (圖 2)。

(ii) 能允許每個屬性能再有一個其他節點做為父節點。

(iii) 能找出屬性之間的依賴關係。移除簡易貝氏分類器一開始假設所有屬性都是獨立的狀態，因為這在現實當中是不切實際的，樹狀貝氏分類器就是為了彌補這個缺點而開發出來的方法。

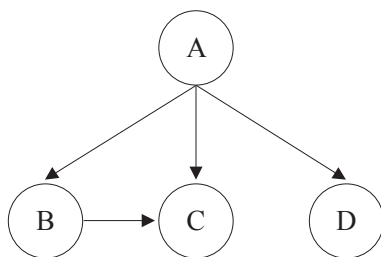


圖 2 樹狀貝氏分類器結構圖

### 3. 決策樹

決策樹 (Decision Tree) 是一項常用的資料探勘技術，是以樹狀結構為基礎的分類分析方法。在機器學習中，決策樹是一個預測模型，所產生的分類可以得知事件發生的關係與規則 (Gürbüz and Yapici 2009)，並可以處理連續與類別性變項。每棵決策樹皆為一種樹形結構，能將欲分類的屬性做解釋與描述，把大量混雜的資料做有效的分類，並且根據決策樹的層級，可以決定變數的強弱，越上層的層級越是重要。如圖 3 所示，樹中的每個節點皆表示為某個屬性，每條分叉路徑則代表某個可能的屬性方向，最底層的葉子點就代表一個最終的決策。決策樹可將大量的資料做區分，且圖形淺顯易懂。本研究採用的 C4.5 是 Quinlan (1993) 改善他自己發展出來的 ID3 (Iterative Dichotomiser 3) 演算法中不能處理連續型數值的問題，使其可以處理連續與類別性變項。由於本研究需處理多元屬性的資料，所以 C4.5 是一個較合適的決策樹演算法。決策樹理論，其優點在於不需要統計假設與大量的資料，即可直接分析原始的資料，並產生易懂而直觀的決策

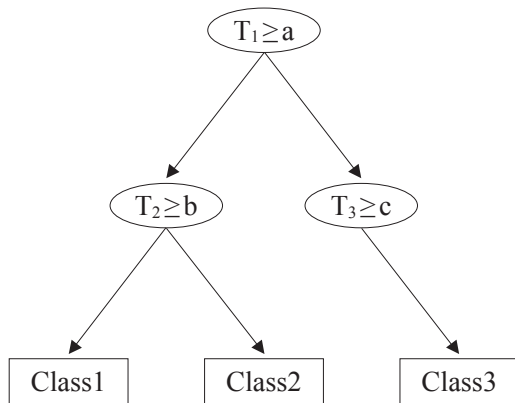


圖 3 決策樹結構圖

規則。例如在圖 3 例子中，一條規則可以呈現如下：

$$\text{IF } T_1 \geq a \ \& \ T_2 \geq b \ \text{THEN Class1.}$$

一般衡量決策規則表現最常見的度量就是「準確率」(accuracy)與「涵蓋率」(coverage)。例如我們要知道決策樹規則 R 的相關指標，可以計算如下：

$$\text{Num.covers}=\text{規則 R 所涵蓋值組的個數}$$

$$\text{Num.correct}=\text{規則 R 能正確判別值組的個數}$$

$$\text{「涵蓋率」Coverage(R)=Num.covers/|D|} \quad /*D:\text{訓練資料集*/}$$

$$\text{「準確率」Accuracy(R)=Num.correct/Num.covers}$$

介紹了三種不同資料探勘的技術方法理論基礎後，本研究在此針對約略集合理論、貝氏網路與決策樹三種不同資料探勘技術的特點加以整理，如表 3 所示：

表 3 三種理論特色比較表

項目	約略集合理論	貝氏網路	決策樹
用途	資料分類、簡化屬性	資料分類、因果關係	資料分類、決策規則
理論特色	(1)不需任何前提假設或關於資料的多餘資訊。 (2)利用近似值的概念，可表達模糊或不確定性的資料。 (3)針對複雜的資訊做處理，以獲得簡化並有規則性的知識。 (4)可獲得關聯規則並用以推導其他資料，進而進行決策推導。	(1)將問題的知識以因果關係表示。 (2)透過貝氏統計的方法，將領域知識與資料之間做結合。	(1)產生易於瞭解的樹狀結構。 (2)提供清楚的指引，告訴我們在進行預測和分類時哪一個變數最重要，及其先後順序。 (3)亦適合處理大量的資料。即使有很多變數被載入模型，決策樹依然可以在短時間內被建構。

#### （四）使用工具與驗證方式

##### 1. ROSE：

在本研究中所用來執行約略集合理論的軟體為 ROSE2，全名為 Rough Sets Data Explorer（Predki et al. 1998; Predki and Wilk 1999），為一套基於約略集合理論的規則發現軟體。可進行約略集合理論的近似度分析、屬性化簡、決策規則的產生，並可將所獲得之分類規則與原始資料來進行驗證。

##### 2. WEKA：

WEKA 的全名 Waikato Environment for Knowledge Analysis（1993），是一個功能強大且可以在 JAVA 的環境下做機器學習（Machine Learning）以及資料探勘（Data Mining）的軟體。WEKA 集合了能夠處理大量資料的各式機器學習演算法，其功能包含資料的預先處理、分類、群集化、關聯分析、以及在新的互動式介面上的視覺化，是目前資料探勘使用非常廣泛的工具。本研究在貝氏網路、決策樹等分類方法，均使用 WEKA v3.72 版本。另外，WEKA 中決策樹所使用的 J48 元件就是利用 C4.5 的演算法所產生。

##### 3. 交互驗證方法：

本文中的貝氏網路與決策樹均利用十摺交互驗證（Ten-fold Cross-validation）方式來做訓練的驗證。此驗證法會先將資料集先分為十等份，每次任取其九為訓練資料集，餘一為測試資料集，反覆將各資料集視為測試性的資料，總共十次執行結果，最後將正確率加總取算術平均值即為最後的實驗結果。

## 四、研究結果

首先我們針對網路算命的行為與動機，進行約略集合理論中的核心與折減，使用 ROSE2 軟體來運算，從眾多變數與屬性中，找出核心屬性，以利進行接下來的貝氏網路因果關係。

約略集合理論的第一步驟，即建立決策表，因此，必須將所收集的樣本一一以資訊表來呈現。約略集合理論還提供了屬性的品質損失表，屬性的品質損失代表如果某屬性自核心（core）中移除，將導致分類準確性的損失值。品質損失值越高，代表此屬性越重要。為挑選較重要屬性，我們可以將損失表中的數值依照下列公式加以標準化，讓數值呈現常態分配，並將高於平均值，也就是標準化值為正值的屬性選取出來，做為最小的折減集合參考（Maciocha and Kisielnicki 2009）。

$$\frac{X - \text{損失值平均數}}{\text{損失值標準差}}$$

算命行為的屬性中，2004 年和 2009 年的品質損失表之結果類似，在此只呈現 2004 年的結果，如表 4 所示。表 4 中八字算命、紫微斗數算命、占星／星座算命這三個屬性比起其他的屬性來說，高於平均值，也就是標準化值為正的，是比較重要的屬性。同樣的，2004 年算命動機之品質損失表（表 5）顯示：財運問題、事業問題、婚姻感情問題、健康問題、瞭解自己的命運是比較重要的屬性。

最後我們將標準化值為正的屬性選取出來如表 6 所示，這些屬性將提供給後面章節的貝氏網路與決策樹方法來做使用，以達到較佳的分析效果。

表 4 2004 年不同方式的算命行為之品質損失表

算 命 行 為		
屬性名稱	品質損失	品質損失標準化
八字算命	0.006	1.143726
紫微斗數算命	0.004	0.103975
手、面相算命	0.001	-1.45565
摸骨算命	0.003	-0.4159
占星／星座算命	0.005	0.62385

表 5 2004 年算命動機之品質損失表

算 命 動 機		
屬性名稱	品質損失	品質損失標準化
子女管教問題	0.014	-0.32925
求學問題	0.012	-0.64118
財運問題	0.018	0.294596
事業問題	0.025	1.386334
婚姻感情問題	0.021	0.762484
健康問題	0.021	0.762484
居住問題	0.003	-2.04484
預測自己的運氣	0.014	-0.32925
瞭解自己的命運	0.017	0.138633

表 6 不同方式的算命行為與算命動機之屬性折減表

項目	選 取 屬 性	屬性數
算命行為 (2004/2009 年)	{ 八字算命、紫微斗數算命、占星／星座算命 }	3
算命動機 (2004 年)	{ 財運問題、事業問題、婚姻感情問題、健康問題、瞭解自己的命運 }	5

第二階段為貝氏網路的方法分析，我們使用 WEKA 軟體來做運算，以貝氏網路來分析各個屬性之間的因果關係圖與機率分配，分類法為樹狀分類器，並且使用了十摺交互驗證法來做分析。在算命行為（2004/2009 年）與算命動機（2004 年）方面，因為這兩種類型的問題只有是與否的分類，較不適合做規則分類，故將著重在因果關係的分析。因此，我們讓貝氏網路來做算命行為與算命動機的因果關係與機率分析。

### （一）算命行為

將 2004、2009 年不同方式的算命行為（可複選）與網路算命的使用做交叉分析，藉此瞭解不同方式的算命行為與網路算命的關係。交叉分析的結果見表 7 及表 8。

2004 年的資料經過交叉分析之後，可以看到民眾在使用不同的方法算命時，不是以網路算命的形式進行。換言之，也就是日常的算命是採用非網路的傳統算命形式。此外，不論是 2004 年或 2009 年，占星／星座的使用者在網路算命的使用比例中都是最高的，顯示使用占星／星座方式算命的人比較容易轉向網路算命。其次，是紫微斗數的使用者也有較高的比例是有使用網路算命。



表 7 2004 年不同方式的算命行為與有使用網路算命之交叉統計表

算 命 行 為	使用網路算命 (百分比 %)		總計
	有	無	
有使用過八字算命	29(28.7%)	72(71.3%)	101
有使用過紫微斗數算命	<b>30(47.6%)</b>	33(52.4%)	63
有使用過手、面相算命	12(30.8%)	27(69.2%)	39
有使用過摸骨算命	2(40.0%)	3(60.0%)	5
有使用過占星／星座算命	<b>40(93.0%)</b>	3( 7.0%)	43

表 8 2009 年不同方式的算命行為與有使用網路算命之交叉統計表

算 命 行 為	使用網路算命 (百分比 %)		總計
	有	無	
有使用過八字算命	52(33.3%)	104(66.7%)	156
有使用過紫微斗數算命	47(42.7%)	63(57.3%)	110
有使用過手、面相算命	25(41.0%)	36(59.0%)	61
有使用過摸骨算命	5(38.5%)	8(61.5%)	13
有使用過占星／星座算命	26(72.2%)	10(27.8%)	36

算命行為的貝氏網路結構如圖 4，可萃取出算命行為之間的相關性。由於貝氏網路僅會在圖中顯示變數間的關聯，不會顯示實際的機率計算，所以本文先以 2009 年算命行為的分析為例說明。TAN 貝氏網路如下，按照變數間的影響力排列，最底層的變數最具影響力。在問卷中有 36 位占星／星座使用者，這 36 位有 26 位曾使用網路算命，比例高達 72%（參見表 9 條件機率表），所以 TAN 模式認為占

星／星座與網路算命關聯最高，故占星／星座變數顯示在 TAN 最底層。在下圖左邊是相關的計算，2009 年的 307 位網路算命的使用者中，有 52 位做過八字算命。這 52 人中有 24 位也有算紫微斗數。而紫微斗數的 47 位使用者中，有 13 位使用占星／星座。

接下來以 TAN 的角度，從下往上描述變項關係。圖中顯示了占星／星座使用者往往會使用網路算命，而喜歡算占星／星座並且使用網路算命的人，往往也喜歡算紫微斗數。喜歡算紫微斗數並且使用網路算命的人，使用八字算命的可能性也較高。換句話說，如果有民眾曾使用占星／星座算命這個行為的話，那麼民眾還有可能會使用紫微斗數算命，而紫微斗數算命也會影響到八字算命的的使用。TAN 屬性間

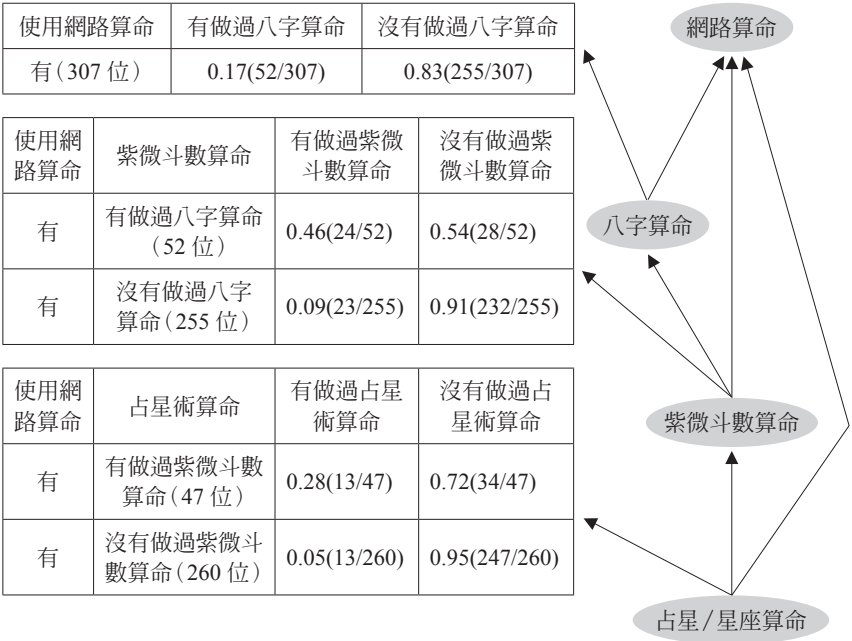


圖 4 2009 年算命行為之貝氏網路結構圖

隱藏的條件機率如表 9。因為篇幅有限，接下來的其他貝氏網路將不再說明其所有隱含機率計算。

接下來我們以決策樹來分析使用行為與網路算命之關係，使用 C4.5 演算法，挑選出來主要的分類變數也是「占星／星座算命」，以大幅縮減屬性種類。代表樣本中相信占星／星座算命的人，大多也是網路算命的使用者。可見一直以來占星／星座算命的使用者對網路算命的接受度最高，也就是說喜愛占星／星座算命的人，常常也是會使用網路算命的族群。

**表 9 2009 年算命行為屬性間之條件機率**

說明	條件機率
1	$P(\text{網路算命}   \text{占星星座使用者}) = 26/36 = 0.72$
2	$P(\text{紫微斗數}   \text{占星星座使用者 and 網路算命}) = 13/26 = 0.5$
3	$P(\text{八字算命}   \text{紫微斗數 and 網路算命}) = 24/47 = 0.51$
4	$P(\text{網路算命}   \text{紫微斗數 and 八字算命 and 占星星座使用者}) = 4/5 = 0.80$

說明 1：在 36 位占星／星座使用者中有 26 位使用網路算命（72%）。

說明 2：在 26 位使用網路算命及占星／星座的網友，有 13 位也使用紫微斗數。

說明 3：在 47 位使用紫微斗數及網路算命的網友中，有 24 位使用八字算命。

說明 4：三種命理方式（八字、紫微、星座）都使用的網友有 5 位，其中有 4 位使用網路算命。

## （二）算命動機

將 2004 年的算命動機（可複選）和網路算命的使用與否來做交叉分析，如表 10 所示。2004 年資料經過交叉分析之後，可以看到民眾的動機若為詢問子女管教問題與事業問題者比較不會使用網路算命。若是為了求學問題，有 71.9% 的人會使用網路算命，推測這部分的網路算命者應為學生族群。

2004 年算命動機的貝氏網路結構如圖 5 所示，可以看到算命動機

表 10 2004 年算命動機與有使用網路算命之交叉統計表

算 命 動 機	使用網路算命（百分比 %）		總計
	有	無	
算命是爲了子女管教	7(20.0%)	28(80.0%)	35
算命是爲了求學問題	46(71.9%)	18(28.1%)	64
算命是爲了財運問題	47(41.6%)	66(58.4%)	113
算命是爲了事業問題	53(30.1%)	123(69.9%)	176
算命是爲了婚姻感情問題	49(45.0%)	60(55.0%)	109
算命是爲了健康問題	39(35.5%)	71(64.5%)	110
算命是爲了住的問題	5(41.7%)	7(58.3%)	12
算命是爲了預測自己的運氣	33(35.1%)	61(64.9%)	94
算命是爲了了解自己的命運	70(38.9%)	110(61.1%)	180

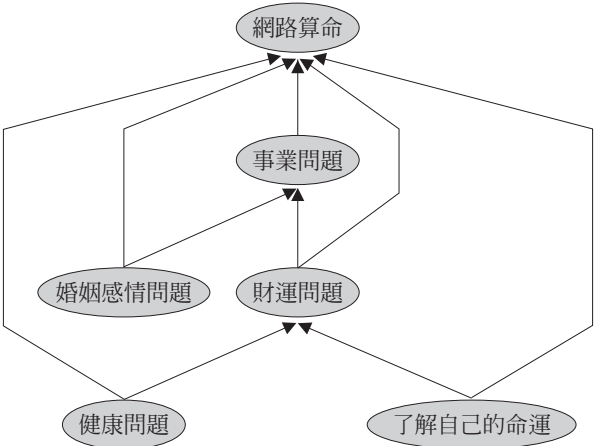


圖 5 2004 年算命動機之貝氏網路結構圖

之間的相關性，圖中顯示了健康問題、瞭解自己的命運與婚姻感情問題是算命動機的源頭。在屬性的因果關係方面，健康問題與瞭解自己的命運直接影響了財運問題。換句話說，如果民眾有健康問題與瞭解自己的命運這兩個因素的話，那麼也可能有動機去詢問財運問題，最後財運問題也常和事業問題有關。而單純只有婚姻感情問題的人，也常會詢問事業問題。

### （三）個人背景資料

個人背景資料與使用網路算命的交叉分析結果如表 11 所示，在性別方面，三個年度皆以女性有使用網路算命的情況較多，顯示曾使用網路算命這一族群，相較於男性，女性對網路算命有較高的熱忱。在年齡方面，三個年度皆以 26-45 歲族群為主，而且 2014 年（63.5%）與 2009 年（60.6%）相較於 2004 年（47.9%），使用網路算命的情況有大幅增加情形，46-65 歲熟年網友也有明顯的增加，網路算命使用者的年齡有逐漸提高的趨勢，不再集中於學生或是年輕族群。

在職業狀況方面，三個年度皆以已就業族群為主，而且 2014 年（72.9%）與 2009 年（73.6%）相較於 2004 年（66.1%），使用網路算命的情況有明顯的增加，可以看出上班族對於網路算命的接受度日漸提高，與文獻中 Yes123 求職網（2015）與 1111 人力銀行（2007）調查的情況相符合。在月收入方面，三個年度皆以 3 萬元以下的族群為主，但在 2014 年，3-6 萬元族群使用網路算命的百分比比 2004 年的數值有明顯的成長（增加 11.6%），顯示薪水在中等程度時，使用網路算命的情況也有越來越多的情況。在婚姻狀況方面，三個年度皆以未婚族群為主，占約六成左右，與文獻中郭貞（2005）的情況相符合。在教育程度方面，三個年度皆以大學以上族群為主，而且 2014 年的百分比

表 11 使用網路算命之個人背景資料統計表

	有使用過網路算命人數（百分比 %）		
	2004 年	2009 年	2014 年
<b>性別</b>			
男	102(43.2%)	139(45.3%)	134(44.8%)
女	134(56.8%)	168(54.7%)	165(55.2%)
<b>年齡</b>			
25 歲以下	110(46.6%)	86(28.0%)	71(23.7%)
26-45 歲	113(47.9%)	186(60.6%)	190(63.5%)
46-65 歲	12( 5.1%)	35(11.4%)	37(12.3%)
66 歲以上	1( 0.4%)	0( 0.0%)	1( 0.3%)
<b>職業狀況</b>			
已就業	156(66.1%)	226(73.6%)	218(72.9%)
未就業	20( 8.5%)	40(13.0%)	39(13.0%)
學生	60( 25.4%)	41(13.4%)	41(13.7%)
<b>月收入</b>			
3 萬元以下	155(65.7%)	187(60.9%)	133(44.5%)
3-6 萬元	68(28.8%)	104(33.9%)	121(40.4%)
6 萬元以上	13( 5.5%)	16( 5.2%)	37(12.4%)
<b>婚姻狀況</b>			
已婚	82(34.7%)	95(30.9%)	106(35.4%)
未婚	151(64.0%)	202(65.8%)	179(59.8%)
已離婚	3( 1.3%)	8( 2.6%)	10( 3.3%)
配偶已去世	0( 0.0%)	2( 0.7%)	0( 0.0%)
<b>教育程度</b>			
國中以下	8( 3.4%)	14( 4.6%)	6( 2.0%)
高職／高中	60(25.4%)	64(20.8%)	51(17.0%)
專科	58(24.6%)	45(14.7%)	46(15.3%)
大學以上	110(46.6%)	184(59.9%)	195(65.2%)
<b>請問您目前的宗教信仰</b>			
沒有	67(28.3%)	56(18.2%)	38(12.7%)
民間、佛、道教	153(64.8%)	222(72.3%)	231(77.2%)
天主、基督	9( 3.8%)	17( 5.5%)	22( 7.3%)
其他	7( 2.9%)	12( 3.9%)	8( 2.6%)
小計	236( 100%)	307( 100%)	299( 100%)

(65.2%) 相較於 2004 年的百分比 (46.6%)，使用網路算命的情況有明顯的增加。換言之，在 2014 年網路算命的使用者之中，高達六成五有大學畢業以上之高學歷。在個人背景資料 (2004 / 2009 / 2014 年) 的網路算命之貝氏網路結構圖與機率分配表。準確率方面如表 12 所示。宗教信仰方面，則以民間、佛、道教這類人較多使用網路算命。在台灣社會中，許多自認為沒有宗教信仰的人，其實也深受儒、釋、道三教雜糅思想的影響，因此也有相當比例的人會使用網路算命。

2004 年個人背景特性的貝氏網路結構如圖 6 所示，圖中顯示了教育程度與婚姻狀況是個人背景使用網路算命的源頭，可以說是影響使用網路算命最主要的個人特徵，其次則為年齡。這項發現與郭貞

表 12 個人背景之分類準確表 (貝氏網路)

項 目	年度	準確率 (%)
個人背景資料	2004	85.7%(1483/1730)
	2009	83.6%(1592/1903)
	2014	85.1%(1645/1934)

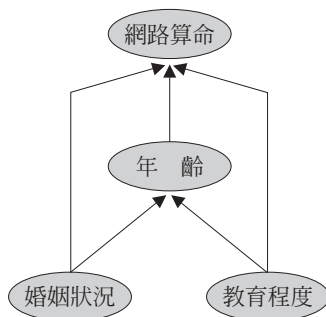


圖 6 2004 年個人背景之貝氏網路結構圖



(2004) 的結論相同，郭貞的研究發現未婚、年齡較輕者與教育程度較高者會使用算命網站。經由貝氏網路可以更進一步顯示所有屬性間彼此的關係，婚姻狀況與年齡有關係、年齡與教育程度有關係。2009 年使用網路算命者的個人背景，屬性篩選結果與 2004 年屬性相同，如圖 7 的貝氏網路結構圖所示。圖中顯示教育程度是 2009 年使用網路算命主要的影響因素。

2014 年資料，加入新的題項：「是否透過網路搜尋及閱讀宗教／信仰相關的訊息頻率。作者將此變選項納入後，發現有趣的現象：民眾透過網路收集宗教訊息為最具影響力的指標（見圖 8）。另一方面，未婚或是教育程度較高者，對網路算命的使用仍然為重要因素。在刻板印象中，人們往往以為教育程度低者比較會去算命，但是網路命理的使用者卻以大學以上學歷為主。

接下來採用決策樹的方法分析，使用 WEKA 軟體來運算，使用 C4.5 演算法，驗證法與貝氏網路一樣使用了十摺交互驗證法來做分析。準確率如表 13 所示。

接下來以決策樹找出 2004 年使用網路算命者個人特徵的隱含規

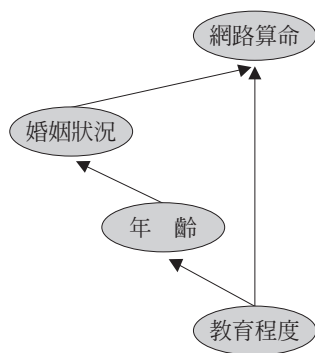


圖 7 2009 年個人背景之貝氏網路結構圖

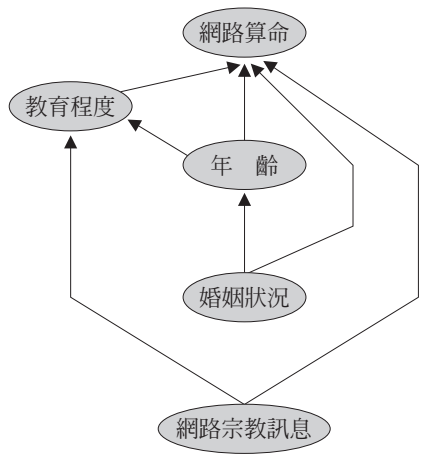


圖 8 2014 年個人背景之貝氏網路結構圖

表 13 個人背景之分類準確表 (決策樹)

項目	年度	準確率 (%)
個人背景資料	2004	86.0%(1489/1730)
	2009	84.5%(1609/1903)
	2014	83.4%(1613/1934)

則，圖 9 為其樹狀結構圖。

圖 9 中的每一個終端結點代表一條決策規則，根節點為年齡，表示此變項為 2004 年個人背景屬性中影響網路算命使用的重要變項。在分類的條件方面，我們將樹狀圖的規則製成表格，並將涵蓋率較高（涵蓋率高於 10%）的規則選出，彙總於表 14。在 2004 年可以選出二條使用網路算命的規則：(1)圖 9 中的節點 14 代表年齡為 25 歲以下、性別為女、月收入為 3 萬元以下、教育程度為大學以上者，有使用網

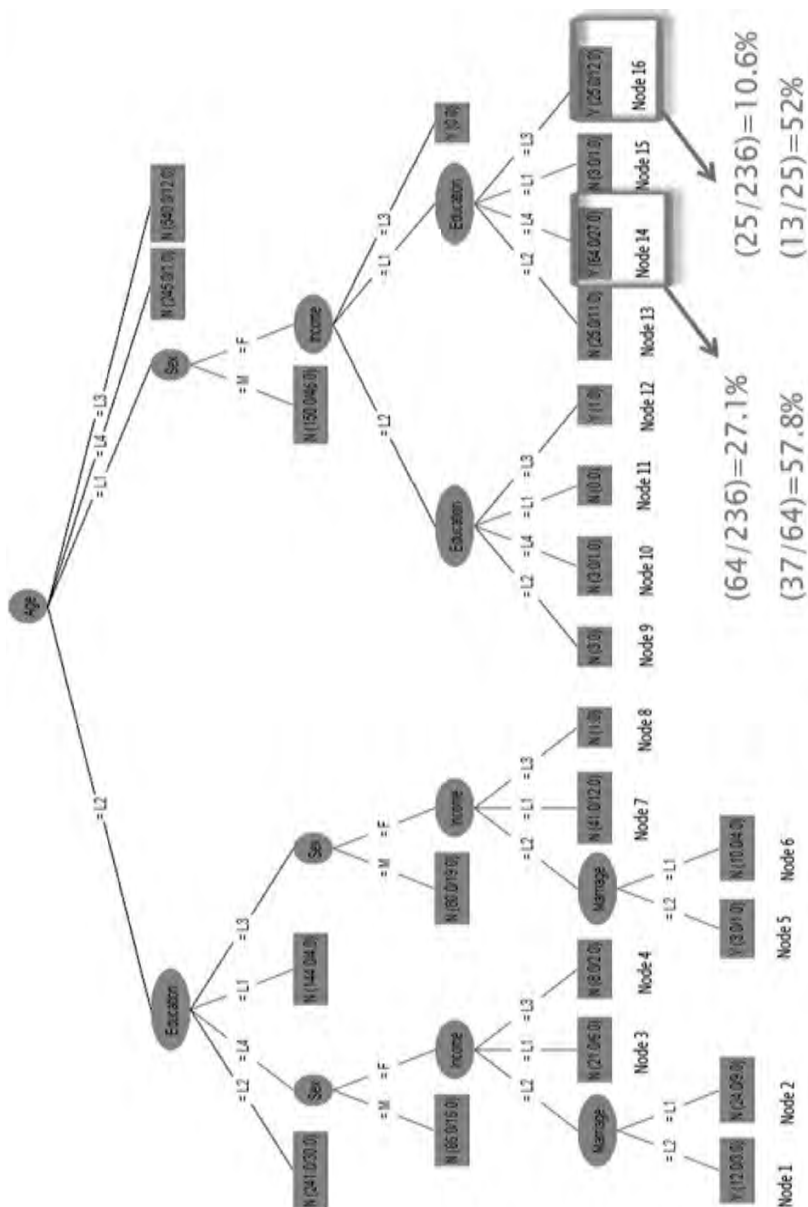


圖 9 2004 年個人背景之樹狀結構圖

路算命，準確率為 57.8%；(2)圖 9 中的節點 16，年齡為 25 歲以下、性別為女、月收入為 3 萬元以下、教育程度為專科者，有使用網路算命，準確率為 52%。2009 年的規則樹狀結構如圖 10 所示，根節點也是年齡，表示年齡此變數在 2009 年的個人背景屬性中也是影響網路算命使用最重要的變數。

在分類的條件方面，同樣的我們將樹狀圖的規則製成表格，並將較有意義的規則，以表 14 顯示。在 2009 年有二個可能發生使用過網路算命的終端節點：(1)節點 10，年齡為 25 歲以下、性別為女、教育程度為大學以上，準確率為 64.0%；(2)節點 14，年齡為 26-45 歲、教育程度為大學以上、婚姻狀況為未婚、性別為女性，準確率為 55.4%。

2014 年的決策樹呈現兩組資料，一組包含宗教信仰和網路搜尋宗教訊息，一組不含宗教信仰和網路搜尋宗教訊息，決策規則彙總於表 14。

2014 年不包含宗教信仰和網路搜尋宗教訊息的規則樹狀結構如圖 11 所示，根節點和 2009 年相同，也是年齡，表示在不考慮宗教信仰和網路搜尋宗教訊息的情況下，年齡為 2014 年個人背景屬性中影響網路算命使用的重要變數。在 2014 年則有三條較顯著的規則：(1)節點 18，年齡為 26-45 歲、教育程度為大學以上，婚姻狀況為未婚，月收入 3-6 萬元，有使用網路算命，準確率為 60.5%。(2)節點 13，年齡為 26-45 歲、教育程度為大學以上的已婚女性，有使用網路算命，準確率為 60%。(3)節點 1，年齡在 25 歲以下，月收入在 3 萬元以下的女性，有使用網路算命，準確率為 66.7%。

2014 年包含宗教信仰和網路搜尋宗教訊息的規則樹狀結構如圖 12 所示，根節點為「是否使用網路搜尋宗教資訊」，表示在考慮宗教信仰變數和網路搜尋宗教訊息的情形下，民眾是否有使用網路搜尋宗

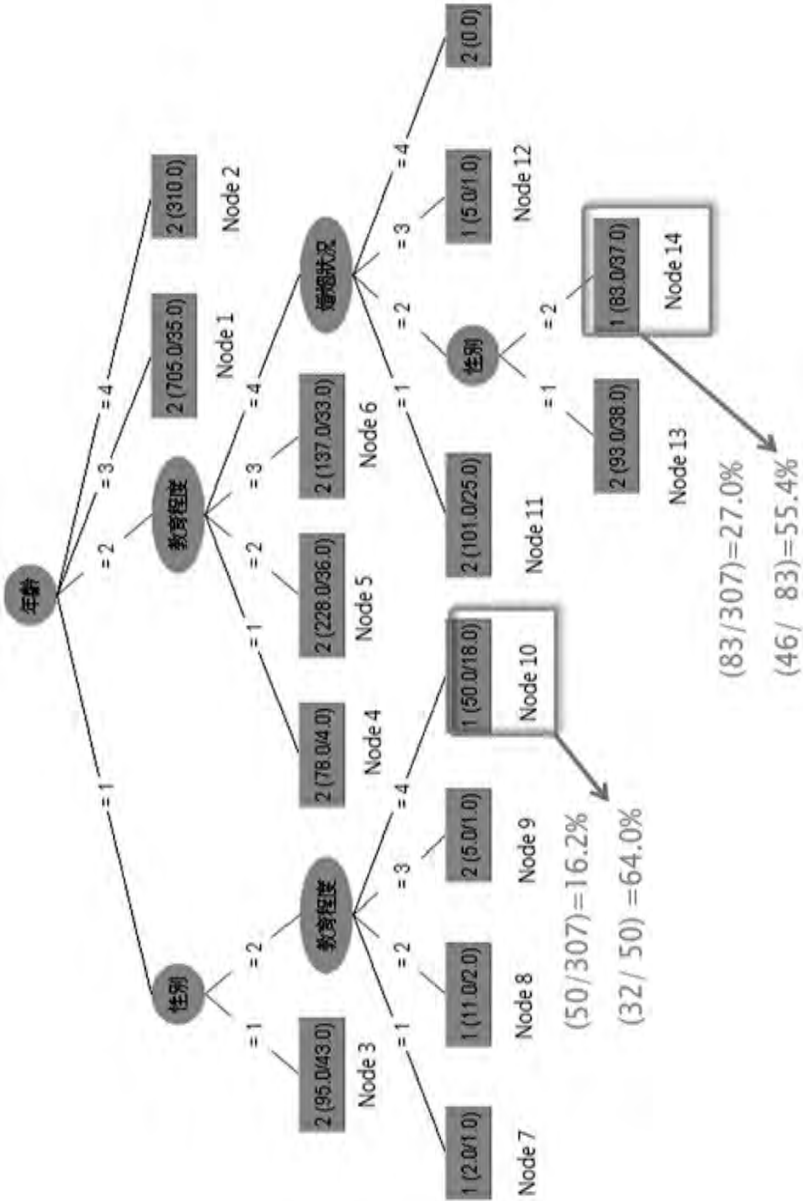


圖 10 2009 年個人背景之樹狀結構圖

表 14 個人使用網路算命之決策規則表

年度	節點	規則描述	涵蓋率	準確率
2004	14	IF (年齡=25 歲以下) & (性別=女) & (月收入=3 萬元以下) & (教育程度=大學以上) THEN 有使用過網路算命	27.1% (64/236)	57.8% (40/75)
	16	IF (年齡=25 歲以下) & (性別=女) & (月收入=3 萬元以下) & (教育程度=專科) THEN 有使用過網路算命	10.6% (25/236)	52% (13/25)
2009	14	IF (年齡=26-45 歲) & (教育程度=大學以上) & (婚姻狀況=未婚) & (性別=女) THEN 有使用過網路算命	27.0% (83/307)	55.4% (46/83)
	10	IF (年齡=25 歲以下) & (性別=女) & (教育程度=大學以上) THEN 有使用過網路算命	16.2% (50/307)	64.0% (32/50)
2014 (圖 11)	18	IF (年齡=26-45 歲) & (教育程度=大學以上) & (婚姻狀況=未婚) & (月收入=3-6 萬) THEN 有使用過網路算命	14.4% (43/299)	60.5% (26/43)
	13	IF (年齡=26-45 歲) & (教育程度=大學以上) & (婚姻狀況=已婚) & (性別=女) THEN 有使用過網路算命	8.4% (25/299)	60% (15/25)
	1	IF (年齡=25 歲以下) & (月收入=3 萬元以下) & (性別=女) THEN 有使用過網路算命	7% (21/299)	66.7% (14/21)
2014 (圖 12)	1	IF (網路搜尋宗教訊息=沒有) THEN 沒有使用過網路算命	92.2% (1507/ 1635)	90.9% (1370/ 1507)
	8	IF (網路搜尋宗教訊息=有) (年齡=26-45 歲) & (教育程度=大學以上) THEN 有使用過網路算命	47.5% (142/299)	54.2% (77/142)

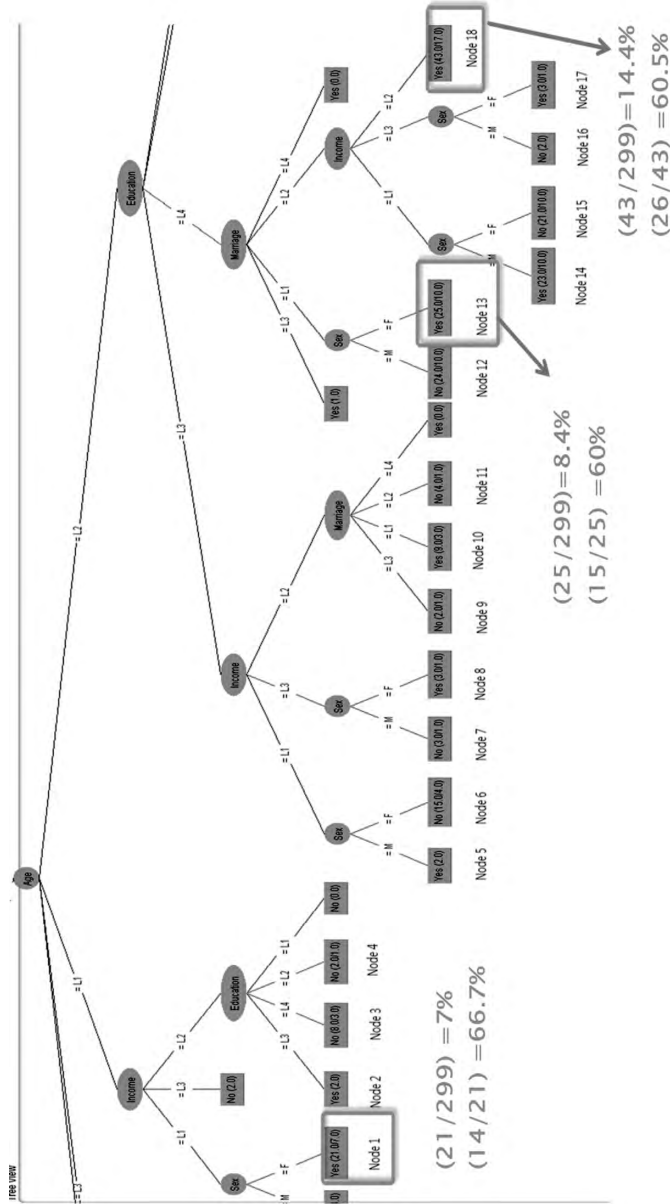


圖 11 2014 年個人背景之樹狀結構圖 (不含宗教信仰及網路搜尋宗教訊息)





教資訊此一行爲，是 2014 年個人背景屬性中影響網路算命使用的最重要變數。若包含宗教信仰變數和網路搜尋宗教訊息，則 2014 年有兩個可能發生使用過網路算命的終端節點：(1)節點 1，沒有使用網路搜尋宗教資訊者，則沒有使用網路算命，準確率爲 90.9%；(2)節點 8，有使用網路搜尋宗教資訊，年齡爲 26-45 歲以下、教育程度爲大學以上者，則有使用網路算命，準確率爲 54.2%。

本文作者也把 2004 年、2009 年和 2014 年的變遷調查，做了 logistic regression 模型。爲了不影響本文的敘述結構，logistic regression 模型的統計結果和其優缺點，另外寫於文末的附錄一。熟悉 logistic regression 的讀者，可同時參考。此外，本文作者也以變遷調查模擬了一個線性迴歸模型，來呈現變項之間的共線性問題，有興趣的讀者，可與作者聯繫。

## 五、結論與討論

網路的使用已成為現代人生活的一環，一些傳統的算命方式也逐漸移至網路上。針對網路算命此新興研究議題，本研究嘗試發現是哪些人在使用網路算命？使用哪一種傳統算命方式的人比較容易轉向網路算命？以及使用算命的動機和網路算命的關係爲何？本研究使用 2004 年、2009 年及 2014 年的「台灣地區社會變遷基本調查」資料，採用資料探勘的方式來分析，研究結果彙整如下：

這三個年度使用網路算命的民眾中，皆以女性、未婚族群以及大學以上的教育程度者爲主。這和以往台灣社會網路算命的研究發現相符合。除此之外，以往的研究文獻強調網路算命使用者集中在非常年輕的人（20 至 24 歲，或 20 至 29 歲）、學生族群和每月可支配所得很少的人，但在本研究中卻已看到變化的軌跡。

在年齡方面，2004 年 25 歲以下和 26-45 歲這兩個年齡層的網路算命使用者，比例相當，至 2009 年 26-45 歲年齡層的網路算命使用者則大幅增加，成為重要的分類條件。46-65 歲年齡層的網友也有明顯的增加。網路算命使用者的年齡已呈現往上提升的趨勢，不再局限於年輕族群。

就職業和收入來看，網路算命使用者以上班族群為主。學生使用網路算命的比例則大幅下降。使用網路算命者的月收入，在 2004 年和 2009 年是以 3 萬元以下居多，但是月收入 3-6 萬者的比例則快速增加，至 2014 年其比例已相當接近月收入 3 萬元以下的。網路算命使用者不再集中於學生族群和低收入者，這樣的研究發現和以往的文獻有很大的不同。

本文以貝氏網路來探勘這些人口變項，折減出這三個年度的重要屬性變項為教育程度、婚姻狀況和年齡。然後，本文藉助決策樹分析方法，挑選出重要的規則。2004 年為 25 歲以下、教育程度大學以上、月收入 3 萬元以下的女性，以及 25 歲以下、教育程度為專科、月收入 3 萬元以下的女性。2009 年為 26-45 歲、教育程度大學以上的未婚女性，以及 25 歲以下、教育程度大學以上的女性。2014 年為 26-45 歲、教育程度大學以上、月收入 3-6 萬的未婚者；26-45 歲、教育程度大學以上的已婚女性；以及 25 歲以下、月收入 3 萬元以下的女性。上述這些挑選出的重要規則，能呈現比較細緻的類別組合。也因為如此，我們可以在 2014 年的重要規則中同時看到月收入 3-6 萬，以及月收入 3 萬以下，這兩種情況下搭配不同的組合條件都有可能從事網路算命。在 2009 年的規則中，26-45 歲和 25 歲以下也是可同時並存的類別，只是搭配的組合條件不同。此外，比較這三年度的重要規則，可看出同樣是大學以上的教育程度，在 2004 年是 25 歲以下、月收入

3 萬以下的未婚女性，到了 2014 年則增加了 26-45 歲、月收入 3-6 萬的未婚者，以及 26-45 歲的已婚女性。也就是說，在 26-45 歲此年齡層，中間收入的未婚者和已婚女性是有較多的人使用網路算命。這個研究發現，值得未來持續的觀察。

2014 年加入「使用網路搜尋宗教資訊」變項後，發現網路搜尋宗教訊息和網路算命有密切關係。以決策樹方法挑選出的兩條重要規則為：沒有使用網路搜尋宗教資訊，就沒有使用過網路算命；有使用網路搜尋宗教資訊、26-45 歲、教育程度大學以上則有使用過網路算命。這是一個有趣的發現，高學歷者藉由網路科技更能優游於網路宗教與網路算命的場域裏，未來或許可以針對網路宗教與算命的議題進行深入探討。

至於不同算命方式的使用者和網路算命之間的關係。本文以貝氏網路來探勘資料，研究發現占星、星座算命的使用者對網路算命接受度最高，也就是說喜愛占星／星座算命的人，常常也是會使用網路算命的族群。使用紫微斗數算命和八字算命者也是有較高比例的人會從事網路算命。這部分的研究發現，相當有趣，和前面的文獻回顧所提及為網友常使用的網路算命項目相似。占星／星座、紫微斗數、八字這三種算命方式的共同點都是需要使用大量的文字，是否較容易吸引教育程度較高者從傳統算命移轉至網路算命？或是使用文字較方便讓業者架設相關的算命網站？還有，占星／星座和紫微斗數都涉及星座的星盤，不同的是西洋式星盤和中國式星盤的差異。為什麼星盤的算命方式和網路算命緊密關聯，這有待未來學者的研究。

本文以貝氏網路來探勘不同的算命動機，研究結果顯示，健康問題、瞭解自己的命運和婚姻感情問題是一切算命動機的源頭。而健康問題與瞭解自己的命運直接影響了財運問題；換言之，若民眾有健康

問題與瞭解自己的命運這兩個算命動機的話，那麼可能也有動機去詢問財運問題，而詢問婚姻感情及財運問題的人，很可能也會查詢事業問題。這樣的研究結果確實符合我們對於一般人會去算命的印象。健康問題若遲遲未得到解決，會去求助算命。若沒有健康問題，則可能是財運上出了問題，但卻會以「想要瞭解自己的命運」來呈現。若沒有健康、財運的急迫問題，則可能是爲了婚姻感情問題。當然，也可能是各種問題的夾雜，久了，只有「想要瞭解自己的命運」的模糊動機。

上述的這些和網路算命有關的研究成果，希望也可提供網路業者做爲一個參考。總之，本研究是一探索性的研究，從資料探勘的角度，來瞭解網路算命相關的因素。本文並未採用線性模型的統計分析方法，因爲線性模型的統計方法，除了考量變項之間是否呈現線性關係，且需符合一些多變項分析的基本假定，例如，變項必須符合常態性與獨立性。本研究關注的是網路算命的新興議題，除了有從事網路算命者的樣本數不大之外，使用的資料大多爲類別變數，並且一些人口變項（女性、未婚、較年輕的年齡層、高教育程度）可能高度相關，因此採用資料探勘方法來達成本研究之目標。

## 附錄一 Logistic Regression 分析模型

爲了對應本文的研究旨趣，我們也針對 2004 年、2009 年和 2014 年這三年的變遷調查分別做了迴歸模型，參見表 1 至表 3。迴歸模型的依變項爲「是否有網路算命」，是 binary variable，因此採用 logistic regression。再者，非常年長的人，比較不會從事網路算命，2004 年的變遷調查，65 歲以上的受訪者當中，只有一位 80 歲的女性，有網路算命的行爲。擔心迴歸模型的運算會受到此「偏差值」的影響，此 logistic 迴歸模型選取 65 歲以下的樣本。同理，2014 年的變遷調查，65 歲以上的受訪者中只有一位 69 歲的男性有網路算命的行爲，迴歸模型也是選取 65 歲以下的樣本。2009 年的資料則沒有 65 歲以上的人有網路算命的行爲，但爲了能和 2004 年和 2014 年的統計結果做比較，也同樣選取 65 歲以下的樣本。此外，表 1 至表 3 中的變項分類方式完全參照本文用來定義變項的分類方式。

表 1 是 2004 年的二元 logistic 迴歸之分析結果。表格中所列出的數字，是 logistic 迴歸分析後，每一類別和對照組的勝算比（odds ratios）。例如，最右邊的一欄，第一個數值是 2.088，意指會發生網路算命的機率比上不會發生網路算命的機率的勝算（odds），女性是男性的 2.088 倍。同理，會發生網路算命的機率比上不會發生網路算命的機率的勝算，25 歲以下的類別是 46-65 歲類別的 7.671 倍；26-45 歲類別是 46-65 歲的 3.928 倍。綜合來看，所有變項的模型中，女性、年齡較年輕、未婚、教育程度較高的人比較會從事網路算命。

我們同時也試了很多不同的模型，這些變項中，對總模型影響力較大的是年齡、教育程度，其次是婚姻狀況、職業，最後是性別、月

表 1 2004 年「有在網路上算過命」和相關變項之間的二元 logistic 迴歸分析的勝算比值

	模型(1)	模型(2)	模型(3)	模型(4)	模型(5)含 連續變數	所有變項
性別						
女性	2.398*** (.169)	2.132*** (.169)	1.928*** (.162)		2.116*** (.171)	2.088*** (.171)
年齡別 (對照組：46-65 歲)					.910*** (.013)	
25 歲以下	12.744*** (.386)		8.219*** (.390)	7.768*** (.398)		7.671*** (.402)
26 - 45 歲	6.278 (.321)		4.367*** (.326)	4.115*** (.328)		3.928*** (.329)
職業別 (對照組：已就業)						
未就業 (家庭主婦、無業等)	.773 (.284)	.805 (.290)	.726 (.277)	.920 (.296)	.907 (.305)	.777 (.297)
學生	1.890** (.248)	1.488 (.252)	1.111 (.254)	1.134 (.262)	1.025 (.264)	1.169 (.267)

表 1 2004 年「有在網路上算過命」和相關變項之間的二元 logistic 迴歸分析的勝算比值 (續)

	模型(1)	模型(2)	模型(3)	模型(4)	模型(5)含 連續變數	所有變項
月收入 (對照組：6 萬元以上)					1.060 (.036)	
3 萬元以下	.441* (.355)	1.205 (.359)		1.178 (.362)		.878 (.372)
3 - 6 萬	.824 (.348)	1.356 (.346)		1.214 (.348)		1.125 (.352)
婚姻狀況 (對照組：已婚)						
未婚	2.203*** (.215)	2.825*** (0.190)	1.581* (.218)	1.445 (.222)	1.003 (.238)	1.625* (.226)
已離婚或配偶已去世	.552 (.618)	.618 (0.623)	.734 (.632)	.907 (.634)	.961 (.633)	.785 (.641)
教育程度 (對照組：國中以下)					1.174*** (.028)	
高中(職)		7.620*** (.392)	4.008*** (.360)	5.004*** (.397)		4.729*** (.400)

表 1 2004 年「有在網路上算過命」和相關變項之間的二元 logistic 迴歸分析的勝算比值 (續)

	模型(1)	模型(2)	模型(3)	模型(4)	模型(5)含 連續變數	所有變項
教育程度 (對照組：國中以下)					1.174*** (.028)	
專科程度		12.111*** (.405)	6.361*** (.370)	7.858*** (.408)		7.345*** (.411)
大學以上		18.121*** (.410)	9.873*** (.369)	12.757*** (.412)		11.319*** (.416)
Constant	.028***	.008***	.006***	.006***	.644***	.006***
N	1485	1483	1613	1483	1483	1483
-2Log Likelihood	1044.810	1021.288	1049.258	1009.152	981.384	990.429
Ominibus test Chi-Square	259.018***	275.186***	318.039***	287.322***	315.089***	306.044***
Cox & Snell R <sup>2</sup>	.160	.169	.179	.176	.191	.186
Nagelkerke R <sup>2</sup>	.274	.291	.313	.302	.328	.320

註：括弧內的數字為標準差；\* $P < 0.05$ ；\*\* $P < 0.01$ ；\*\*\* $P < 0.001$ 。



收入。月收入對總模型貢獻效益最小，這可從模型(3)看出，移除月收入，對模型幾乎沒甚麼影響。值得注意的是，教育程度和職業別中的「學生」類別，有很高的關聯。若不同變數之間的類別並不是 complete (or quasi-complete) separation (Osborne 2015: 88)，是會影響 logistic 迴歸的運算。

若移除教育程度此變項，重新做一次分析，其數值列在模型(1)。模型(1)中的職業別學生類，以及月收入 3 萬以下，這兩類別的檢定則轉為顯著，P 值分別為 0.01 和 0.021。我們也做了移除年齡變項的模型(2)，並未改變顯著的變數，只是教育程度變項中高中以上的類別，對照國中以下類別，倍數提升非常多。主要是因為年齡和教育程度有很高的相關，相關係數是  $-0.560$ ， $P < .001$ 。

最後，本文作者把年齡、月收入和 education 此三變項，改回成連續變數，重做的 logistic 迴歸分析，列在模型(5)。連續變項的勝算比說明也類似類別變項，只是把連續變項想成 20 個或 100 個類別。每一單位的增加，例如年齡增加一歲，勝算值是對照組（最小年齡）的 0.91 倍。換句話說，勝算值是減少的。隨著年齡增加，會從事網路算命的機率比不會從事網路算命的機率的勝算值，呈現減少的情形。一般是以  $-2\text{Log Likelihood}$  來檢測 lack of fit，因此數值愈小，代表模型愈 fit。模型(5)的  $-2\text{Log Likelihood}$  數值比總模型的數值略為小一些。整體模型檢定 Omnibus Test 也是可幫助我們判斷模型是否愈適合。模型(5)的 Chi-square 值是比總模型的略大，代表模型較適合。不過，未婚的細項卻成了不顯著的類別。

值得一提的是，原本婚姻狀況變項中「配偶已去世」細格中的 cases 為 0，和已婚的勝算比數值為 0，而出現標準差非常大的奇怪數值。一般的做法是合併細格，因此表 1 中我們是把「配偶已去世」和

離婚類別合併成一新類別。2014 年的資料，也同樣出現「配偶已去世」的細格中之次數為 0，也是和離婚細格合併成一新類別。

表 2 是 2009 年變遷調查的 logistic 迴歸之分析結果。最右欄的「所有變項」模型中，呈現女性、年齡較年輕、未婚、教育程度較高的人，比較會有網路算命的行為。和 2004 年的資料相較，2009 年的顯著變項中，多了「未就業」（家庭主婦、無業等）這一細項。會發生網路算命的機率比不會發生網路算命的機率之勝算，「未就業」是「已就業」的 0.614 倍。也就是說，和「已就業」相比較，「未就業」的勝算是比較少的。2004 年「未就業」對「已就業」的勝算比值也是小於 1，不過這個勝算比值的檢定未達顯著水準。另一個和 2004 年分析結果不一樣的地方是，2009 年「學生」這一細項的勝算是「已就業」勝算的 0.793 倍。2004 年這個勝算比值卻是大於 1，為 1.169，雖然這兩個勝算比值之檢定皆未達顯著水準。

變項之間高度關聯，年齡和教育程度的相關係數為  $-0.461$ ， $P < .001$ 。教育程度和月收入的相關係數為  $0.356$ ， $P < .001$ 。性別和月收入的相關係數為  $-0.243$ ， $P < .001$ 。若移除年齡此變項，依模型(2)：「三萬元以下」和「3 至 6 萬」這兩細項的比值得檢定，則轉為顯著。P 值分別為 0.004 和 0.015。若移除性別，見模型(4)，則月收入的兩個細項也是轉為顯著，P 值分別為 0.015 和 0.027。2009 年的變項中對總模型貢獻效益較大的是年齡、教育程度、婚姻狀況此三個變項，月收入的貢獻效益最小。若移除月收入，依模型(3)所示：對總模型的解釋力幾乎沒甚麼影響。接下來，我們也把年齡、月收入和 education 改成連續變數，其統計結果列在模型(5)。模型(5)的  $-2\text{Log Likelihood}$  數值比總模型的數值小。Omnibus Test 的 Chi-square 值是比總模型的大，代表模型較適合，但是未就業和未婚這兩細項卻成了不顯著的類別。

表 2 2009 年「有在網路上算過命」和相關變項之間的二元 logistic 迴歸分析的勝算比值

	模型(1)	模型(2)	模型(3)	模型(4)	模型(5)含 連續變數	所有變項
性別						
女性	1.861*** (.147)	1.699*** (.147)	1.777*** (.144)		1.734*** (.149)	1.726*** (.149)
年齡別 (對照組：46-65 歲)					.934*** (.010)	
25 歲以下	9.476*** (.311)		6.381*** (.314)	5.687*** (.324)		5.891*** (.325)
26-45 歲	4.457*** (.241)		3.426*** (.247)	3.211*** (.249)		3.162*** (.250)
職業別 (對照組：已就業)						
未就業 (家庭主婦、無業等)	.665 (.213)	.496*** (.209)	.643* (.206)	.608* (.219)	.548 (.234)	.614* (.219)
學生	1.132 (.280)	.986 (.273)	.831 (.282)	.741 (.288)	.583 (.288)	.793 (.291)

表 2 2009 年「有在網路上算過命」和相關變項之間的二元 logistic 迴歸分析的勝算比值 (續)

	模型(1)	模型(2)	模型(3)	模型(4)	模型(5)含 連續變數	所有變項
月收入 (對照組：6 萬元以上)					.948 (.036)	
3 萬元以下	.902 (.311)	2.526** (.320)		2.193* (.322)		1.712 (.330)
3-6 萬	1.270 (.305)	2.115* (.307)		1.976* (.307)		1.744 (.310)
婚姻狀況 (對照組：已婚)						
未婚	2.644*** (.180)	3.582*** (0.164)	2.343*** (.182)	2.031*** (.184)	1.453 (.205)	2.195*** (.186)
已離婚	1.125 (.403)	1.213 (0.401)	1.354 (.406)	1.297 (.407)	1.297 (.404)	1.312 (.408)
配偶已去世	.944 (.760)	.618 (.750)	1.189 (.764)	1.319 (.763)	1.222 (.754)	1.126 (.765)

表 2 2009 年「有在網路上算過命」和相關變項之間的二元 logistic 迴歸分析的勝算比值 (續)

	模型(1)	模型(2)	模型(3)	模型(4)	模型(5)含 連續變數	所有變項
教育程度 (對照組：國中以下)					1.091*** (.015)	
高中(職)		3.116*** (.312)	2.193* (.319)	2.280** (.320)		2.280** (.320)
專科程度		6.747*** (.316)	4.001*** (.319)	4.433*** (.325)		4.282*** (.326)
大學以上		8.609*** (.322)	4.881*** (.316)	5.827*** (.330)		5.422*** (.332)
Constant	.032*** 1625	.012*** 1625	.015*** 1647	.011*** 1625	.934*** 1625	.010*** 1625
-2Log Likelihood	1271.065	1265.128	1246.997	1245.273	1220.743	1231.738
Ominibus test Chi-Square	304.073***	310.009***	343.163***	329.864***	354.394***	343.399***
Cox & Snell R <sup>2</sup>	.171	.174	.188	.184	.196	.190
Nagelkerke R <sup>2</sup>	.275	.280	.304	.296	.316	.307

註：括弧內的數字為標準差；\* $P < 0.05$ ；\*\* $P < 0.01$ ；\*\*\* $P < 0.001$ 。

表 3 是 2014 年資料的分析結果。最右手邊一欄的所有變項的模型之分析結果，大致上是類似表 1 和表 2。其差異有兩點：一是「未就業」細項的勝算是「已就業」的勝算之 1.125 倍，雖然此勝算比的檢定未達顯著水準。而 2004 年和 2009 年同一細項的勝算比值卻是小於 1。另一個差異處是「3 萬元以下」這一細格的勝算比值為 0.720，和表 1 的數值相近，但和表 2 的比值卻是差距很大。很難解釋不同時間點的起伏現象。然而，表 3 中的月收入，對總模型的貢獻效益是最小的，這點和表 1、表 2 類似。

不過，月收入改成連續變項之後，依模型(5)所示，月收入成了顯著的變項。勝算比值為 1.09，接近 1，意指隨著月收入的增加，會從事網路算命的機率比不會從事網路算命的機率之勝算比值，呈現稍微增加的情形。此外，這個包含連續變數的模型(5)，其  $-2\text{Log Likelihood}$  數值比總模型的數值大。Ominibus Test 的 Chi-square 值也比總模型的小，代表總模型是要比模型(5)適合，這樣的結果是和表 1、表 2 中的模型(5)是較適合的模型之研究節果，有所不同。

綜合比較這三年的數值，我們大概可得下列的結論：女性、年齡較年輕、未婚、教育程度較高的人，比較會有網路算命的行為。2004 年至 2014 年的變化，在於性別和年齡兩變項的勝算比值呈現縮小的情形。2004 年女性的勝算相對於男性的勝算，其比值為 2.088，到了 2014 年這個比值降為 1.660。「25 歲以下」相對於「46-65 歲」的勝算比值，則從 2004 年的 7.671，降至 2014 年的 4.446。亦即，會從事網路算命的機率比不會從事網路算命的機率的勝算，從女性是男性 2.088 倍略減為 1.660 倍；「25 歲以下」是「46-65 歲」的 7.671 倍則降至 4.446 倍。

整體來看，同樣的資料採用 logistic regression 的分析方式，其優點是可以看出一個粗略簡單的全面圖像，但卻無法呈現類別之間較細

表 3 2014 年「有在網路上算過命」和相關變項之間的二元 logistic 迴歸分析的勝算比值

	模型(1)	模型(2)	模型(3)	模型(4)	模型(5)含 連續變數	所有變項
性別						
女性	1.800*** (.146)	1.667*** (.144)	1.561** (.141)		1.778*** (.146)	1.660*** (.147)
年齡別 (對照組：46-65 歲)					.942*** (.010)	
25 歲以下	6.648*** (.298)		4.157*** (.292)	4.254*** (.306)		4.446*** (.308)
26 - 45 歲	4.882*** (.222)		3.438*** (.219)	3.464*** (.229)		3.364*** (.230)
職業別 (對照組：已就業)						
未就業 (家庭主婦、無業等)	1.141 (.227)	.858 (.219)	.904 (.210)	1.141 (.231)	1.137 (.227)	1.125 (.230)
學生	1.384 (.273)	1.188 (.256)	.899 (.261)	1.051 (.276)	.812 (.248)	1.085 (.279)

表 3 2014 年「有在網路上算過命」和相關變項之間的二元 logistic 迴歸分析的勝算比值 (續)

	模型(1)	模型(2)	模型(3)	模型(4)	模型(5)含 連續變數	所有變項
月收入 (對照組：6 萬元以上)					1.090*** (.026)	
3 萬元以下	.435*** (.249)	1.002 (.252)		.883 (.256)		.720 (.264)
3 - 6 萬	.792 (.228)	1.337 (.225)		1.168 (.229)		1.067 (.232)
婚姻狀況 (對照組：已婚)						
未婚	1.507 (.176)	2.401*** (0.162)	1.507* (.176)	1.531* (.182)	1.330 (.196)	1.605** (.182)
已離婚或配偶已去世	.974 (.361)	1.018 (0.362)	1.340 (.356)	1.340 (.371)	1.126 (.363)	1.296 (.371)
教育程度 (對照組：國中以下)					1.071*** (.015)	
高中(職)		7.364*** (.479)	5.410*** (.483)	4.903*** (.485)		4.982*** (.486)



表 3 2014 年「有在網路上算過命」和相關變項之間的二元 logistic 迴歸分析的勝算比值 (續)

	模型(1)	模型(2)	模型(3)	模型(4)	模型(5)含 連續變數	所有變項
教育程度 (對照組：國中以下)					1.071*** (.015)	
專科程度		15.498*** (.477)	10.734*** (.479)	9.811*** (.482)		9.430*** (.484)
大學以上		16.812*** (.481)	12.646*** (.479)	10.862*** (.486)		10.293*** (.488)
Constant	.057***	.010***	.008***	.010***	.362**	.009***
N	1559	1557	1628	1557	1557	1557
-2Log Likelihood	1296.342	1284.224	1294.336	1262.454	1268.166	1250.492
Ominibus test Chi-Square	201.530***	212.825***	255.462***	234.595***	228.883***	246.556***
Cox & Snell R <sup>2</sup>	.121	.128	.145	.140	.137	.146
Nagelkerke R <sup>2</sup>	.196	.207	.237	.226	.221	.237

註：括弧內的數字為標準差；\* $P < 0.05$ ；\*\* $P < 0.01$ ；\*\*\* $P < 0.001$ 。

緻的關聯 (association)。尤其是當不同自變數之間的類別不是 complete (or quasi-complete) separation，類別之間在計算 logistic regression 係數時會互相干擾。而連續變數和類別變數也是會互相干擾。本文所關注的網路算命這樣的新興議題，一些人口變項之間就是高度相關。以 logistic regression 來處理的另一缺失是計算出來的數值，很難向讀者解釋說明清楚。「勝算=事件會發生的機率／事件不會發生的機率」，不是那麼容易理解的方式。勝算值是要比機率值大，但一般人還是把勝算想成機率。此外，勝算比值是經過自然對數轉換而得，當比值距離 1.00 越遠，會有越強的膨脹效果 (inflate effect sizes) (Osborne 2015: 34-35)。因如此，表 1 中的模型(1)的年齡類別，會有 12.744，相當高的勝算比值。

本文使用同樣的變遷調查，但以貝氏網路的資料探勘方式折減出重要的變項，並以決策樹分析方法選出使用網路算命的規則。以 2004 年為例，本文使用資料探勘的貝氏網路篩選出年齡、教育程度、婚姻狀況為重要變項。其次，本文的決策樹也選出兩條使用網路算命的規則，第一條規則為年齡為 25 歲以下、婚姻狀況為未婚、性別為女、教育程度為大學以上、月收入為 3 萬元以下，規則準確率為 57.8%，涵蓋率為 27.1%。第二條規則為年齡為 25 歲以下、婚姻狀況為未婚、性別為女、月收入為 3 萬元以下、教育程度為專科，規則準確率為 52%，涵蓋率為 15.2%。本文採用的方法，可以避開迴歸模型的共線性問題，也很適用在許多類別變項的分析。且視覺化的決策樹圖，一目了然。

## 參考文獻

- 1111 人力銀行，2007，企業求才命理大調查 (<https://winner.1111.com.tw/zone/pr/headline.asp?autono=1577>，取用日期：2015 年 10 月 15 日)。
- Yes123 求職網，2015，半年職場命理調查 ([https://www.yes123.com.tw/admin/white\\_paper/index\\_list.asp](https://www.yes123.com.tw/admin/white_paper/index_list.asp)，取用日期：2015 年 10 月 15 日)。
- 台灣網路資訊中心，2015，2014 年臺灣寬頻網路使用狀況調查 (<https://www.twNIC.net.tw/download/200307/20140820e.pdf>，取用日期：2015 年 10 月 20 日)。
- 江羿臻、林正昌，2014，〈應用決策樹探討中學生學習成就的相關因素〉。《教育心理學報》45(3): 303-327。
- 吳怡瑾、李睿傑、陳子立，2014，〈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〉。《管理與系統》85: 41-182。
- 林彥錡，2009，《網路命理儀式參與者之意願成因探討》。嘉義：南華大學傳播學研究所碩士論文。
- 帥嘉珍、高振源、蘇宜芬，2009，〈從使用者觀點討論科技產業應用企業系統之綜合效益——以 ERP 對 SCM 之影響為例〉。《電子商務研究》7(3): 269-290。
- 張國偉，2003，《算命網站顧客信任感影響因素之研究》。高雄：國立中山大學資訊管理學系研究所碩士論文。
- 章英華、傅仰止，2004，《台灣地區社會變遷基本調查計畫：第四期第五次調查計畫執行報告》。行政院國家科學委員會專題研究計畫報告，NSC 93-2420-H-001-002-B1。台北：中央研究院社會學研究所。
- 郭貞，2004，《從多重學理取徑探討網路之算命行為》。行政院國家科學委員會專題研究計畫報告，NSC 91-2412-H-004-024-SSS。台北：國立政治大學廣告學系。
- ，2005，〈影響網路算命行為因素與動機之探討：找尋網路算命行為之動力心理模式〉。《新聞學研究》85: 41-182。
- 陳建榮，2004，〈網路社會中的巫術文化——台灣網路算命之初探〉。《崇右學報》10: 263-281。
- 傅仰止、杜素豪，2009，《台灣地區社會變遷基本調查計畫：第五期第五次調查計畫執行報告》。行政院國家科學委員會專題研究計畫報告。台北：中央研究院社會學研究所。
- 傅仰止、章英華、杜素豪、廖培珊，2015，《台灣社會變遷基本調查計畫：第六期第五次調查計畫執行報告》。行政院國家科學委員會專題研究計畫報告。台北：中央研究院社會學研究所。

- 鄒小蘭、王琬蓁，2015，〈國中一般智能資賦優異學生鑑定資料分析與預測力研究〉。《特殊教育學報》42: 87-110。
- 劉素君，2012，〈西洋占星命理網站使用者行為之研究〉。台北：實踐大學企業管理學系碩士論文。
- 薄懷武，2009，〈網路算命下的生活策略〉。台北：東吳大學社會學研究所碩士論文。
- 瞿海源，1999，〈術數流行與社會變遷〉。《臺灣社會學刊》22: 1-45。
- 顏任儀，2012，〈玩命之徒的玩命之途：女性算命行為的性別面向分析〉。《性別教育平等季刊》61: 108-114。
- 魏劭楷，2010，〈網路算命使用者的動機、滿足、信任與評估的初探性研究——以批踢踢實業坊命理研究院為例〉。嘉義：南華大學生死學研究所碩士論文。
- Friedman, Nir, Dan Geiger, and Moises Goldszmidt, 1997, "Bayesian Network Classifiers." *International Machine Learning* 29: 131-163.
- Gürbüz, Feyza, Lale Özbakira, and Hüseyin Yapicib, 2009, "Classification Rule Discovery for the Aviation Incidents Resulted in Fatality." *Knowledge-Based Systems* 22(8): 622-632.
- Kuo, Cheng, 2009, "A Study of the Consumption of Chinese Online Fortune Telling Services." *Chinese Journal of Communication* 2(3): 288-306.
- Kuzey, Cemil, Ali Uyar, and Dursun Delen, 2014, "The Impact of Multinationality on Firm Value: A Comparative Analysis of Machine Learning Techniques." *Decision Support Systems* 59: 127-142.
- Lövheim, Mia, 2008, "Rethinking Cyberreligion? Teens, Religion and the Internet in Sweden." *Nordicom Review* 29(2): 205-217.
- Maciocha, Agnieszka, and Jerzy Kisielnicki, 2009, "Intangible Assets in a Polish Telecommunication Sector—Rough Sets Approach." Pp. 169-196 in *Transactions on Rough Sets X*, edited by James F. Peters, Andrzej Skowron, Marcin Wolski, Mihir K. Chakraborty and Wei-Zhi Wu. Heidelberg, Berlin: Springer-Verlag.
- Osborne, Jason W., 2015, *Best Practices in Logistic Regression*. New York: Sage Publications.
- Pawlak, Zdzistaw, 1982, "Rough Sets." *International Journal of Computer and Information Science* 11(5): 341-356.
- Pawlak, Zdzistaw, 2002, "Rough Sets and Intelligent Data Analysis." *Information Sciences* 147(1): 1-12.
- Predki, Bartlomiej, and Szymon Wilk, 1999, "Rough Set Based Data Exploration Using ROSE System." Pp. 172-180 in *Foundations of Intelligent Systems: 11th International Symposium, ISMIS'99 Warsaw, Poland, June 8-11, 1999 Proceedings*. (Lecture Notes in Artificial Intelligence 1609), edited by Zbigniew W. Raz, and Andrezej Skowron. Berlin/Hei-

- delberg: Springer.
- Predki, Bartłomiej, Roman S. Iowinski, Jerzy Stefanowski, Robert Susmaga, and Szymon Wilk, 1998, "ROSE—Software Implementation of the Rough Set Theory." Pp. 605–608 in *Rough Sets and Current Trends in Computing: First International Conference, RSCTC'98 Warsaw, Poland, June 22–26, 1998 Proceedings*. (Lecture Notes in Artificial Intelligence 1424), edited by Lech Polowski, and Abdrezej Skowron. Berlin/ Heidelberg: Springer.
- Quinlan, J. Rose, 1993, *C4.5: Programs for Machine Learning*, San Francisco, CA: Morgan Kaufmann.
- Shuai, Jia-Jane and Han-Lin Li, 2005, "Using Rough Set and Worst Practice DEA in Business Failure Prediction." Pp. 503–510 in *Rough Sets, Fuzzy Sets, Data Mining, and Granular Computing: 10th International Conference, RSFDGrC 2005, Regina, Canada, August 31–September 3, 2005, Proceedings, Part II*. (Lecture Notes in Computing Science 3642), edited by Dominik Slezak, Jing Tao Yao, James F. Peters, Wojciech Ziarko, and Xiaohua Hu. Berlin/Heidelberg: Springer.
- Walczak, Beata, and D.L. Massart, 1999, "Tutorial Rough Sets Theory." *Chemometrics and Intelligent Laboratory Systems* 47(1): 1–16.
- WEKA: Data Mining Software in Java. <http://www.cs.waikato.ac.nz/ml/weka/> (Date visited: July 17, 2015)
- Wu, Wei-Wen, Yu-Ting Lee, Ming-Lang Tseng, and Yi-Hui Chiang , 2010, "Data Mining for Exploring Hidden Patterns between KM and Its Performance." *Knowledge-Based Systems* 23(5): 397–401.