

林金賢、周世玉、楊良基、謝宇宣(2021),「就要醬玩—運用資料探勘技術規劃您所喜愛的旅遊行程」,《資訊管理學報》,第二十八卷,第四期,頁 395-418。

就要醬玩—運用資料探勘技術規劃您所喜愛的 旅遊行程

林金賢

國立中興大學企業管理學系

周世玉*

國立臺灣師範大學企業管理學系

楊良基

國立中興大學企業管理學系

謝宇宣

國立中興大學企業管理學系

摘要

近年來我國旅遊市場不斷成長，而自助旅行人次也屢創新高，加上資通訊科技的進步對旅遊市場造成的衝擊，如何針對自由行旅客提出同時符合旅遊動機以及消費者個人特性的旅遊行程，已成為一個重要的議題。本研究嘗試結合決策樹、羅吉斯、及類神經網路等資料探勘技術提出令旅客滿意的行程推薦。實證結果顯示旅遊動機的確可以提高對旅遊行程滿意度的預測能力，而本研究所提出之方法可以排除個人的主觀意識，從過去的旅遊行為中找出特定的旅遊行程，建構旅客旅遊動機、個人特性、與旅遊行程間的函數關係，進一步精準的預測自由行旅客對不同行程的滿意程度。此方法可以讓目的地行銷公司更有效地提供令旅客滿意的行程推薦。

關鍵字：旅遊動機、旅遊行程、決策樹、羅吉斯葉、類神經葉

* 本文通訊作者。電子郵件信箱：sychou@ntnu.edu.tw
2021/02/08 投稿；2021/05/12 修訂；2021/06/13 接受

Lin, C.S., Chou, S.Y., Yang, L.C. & Hsieh, Y.H. (2021), 'That's How We Play - Using Data Mining Techniques to Design Your Favorite Tour', *Journal of Information Management*, 28(4), 395-418.

That's How We Play-Using Data Mining Techniques to Design Your Favorite Tour

Chin-Shien Lin

National Chung Hsing University

Shih-yu Chou*

National Taiwan Normal University

Liang-Chi Yang

National Chung Hsing University

Yu-Hsuan Hsieh

National Chung Hsing University

Abstract

With the continuing growth of Taiwan's tourism market in recent years, especially the number of self-guided travelers has repeatedly hit new heights, and the huge impact caused by the advancement of Information and Communication Technology, how to propose a travel itinerary which can both consider the traveler's motives and demographic characteristics is an essential issue. This study is trying to construct a model by combining Decision Trees, Logit, and Neural Network to provide a trip recommendation that could satisfy the self-guided travelers. The empirical results show that travel motives can indeed help increase the prediction accuracy of travel satisfaction. The method proposed in this study can get rid of the individual subjectivity, finding out specific tourism types from travelers' past tourism behaviors. It constructs a relationship among the travelers' motives, personal characteristics, and the tourist attractions to precisely predict the satisfaction of different tourism itineraries for each self-guided traveler. This proposed method can be applied by the destination marketing organization to effectively provide the customers satisfying travel itinerary.

Keywords: Travel motive, travel itinerary, decision trees, logit leaf, neural network leaf

* Corresponding author. Email: sychou@ntnu.edu.tw
2021/02/08 received; 2021/05/12 revised; 2021/06/13 accepted.

壹、緒論

資通訊技術的進步使得消費者對於旅遊行程的規劃與決策起了很大的變化。在過去資訊不對稱的情況下，大多數消費者都仰賴旅行社專業導遊為其規劃旅遊，例如行程推薦、住宿旅館預訂、交通安排、機位預訂、費率確認等。然而自從網路與手機使用普及之後，消費者對於旅遊目的地的訊息可以非常容易取得，對行程的規劃有很大的自主性，而導遊的角色也被減縮，使整個旅遊產業的商業模式產生很大的改變。

為了因應旅遊產業競爭環境的改變，觀光業者也各使出不同的方法來吸引消費者。有些提供網路評論機制，讓消費者可以知道其他消費者對各景點旅遊完後的評價；有些積極推出新景點以吸引新的客群；有些跟其他業者間作策略聯盟，提出絕對優惠的價格來吸引消費者；更有些則推動忠誠計畫，讓消費者入會後可以有折扣，甚至於可享退費等。在現有目的地行銷業者的激烈競爭下，要想能夠勝出必須要能夠了解消費者的旅遊動機與滿意度間的關係(Devesa et al. 2010; Yoon & Uysal, 2005; Zabkar et al. 2010)。目的地行銷者必須能夠依據消費者的認知來作市場定位與差異化(Chattalas et al. 2008)。如果可以讓消費者在適當的時機獲取相關的旅遊相關資訊，就可以讓他們更清楚的知道他們想要的行程並增加其滿意度(Wheeler 1995)。

依據相關的研究結果，旅遊動機是旅遊的最主要決定因子(March & Woodside 2005)，而旅遊的動機不外乎建立關係，聲譽，以及求知 (Sung et al. 2016)，參訪獨特位置的景點 (Sirisack et al. 2014)，以及海邊的休閒活動(Abodeeb et al. 2015)。動機會因為人口統計變項的不同而不同(Jensen 2011; Suttikun et al. 2018; Alshammari et al. 2019)，使得人口統計變項在旅遊的研究中不可或缺(Kara & Mkwizu 2020; Allaberganov & Preko 2021)。另外一方面，動機也是影響旅遊滿意度的一個重要因素(Gidey & Sharma 2017)。而遊客滿意度對於一個觀光景點的永續發展更是扮演一個非常重要的角色(Prayag 2009)。Chi & Qu (2008)也指出滿意度是影響遊客做旅遊規劃，購買產品或服務，以及是否重遊決策時的重要因素。

另外一方面，資料探勘技術主要目的在從大量的數據資料中萃取出有用的訊息以利相關的決策制定。雖然說已經有很多的資料探勘以及人工智慧工具被發展出來，然而由於每一項工具都有其優缺點，如何選擇一個好的工具來解決所對應的問題並不是一件容易的事 (You et al. 2015)。例如，類神經網路(artificial neural network, ANN)在模型建構的過程有很多的參數要決定(Zhang et al. 2005)，包括學習速率，隱藏層的層數，每一層的神經元個數等。使得在解釋變數間的關係時會有黑箱的質疑，以及訓練時間跟過度配適的問題。然而其優點是它可以偵測出自變數與應變數間的複雜函數關係。而決策樹工具的優點是，簡單易懂也容易操作，相較於其他工具在使用上比較友善。然而，其缺點則包括不穩定性，不夠準確。整體來說，沒有一個工具適用於所有的問題(Akin 2015)。因此有很多的學者企圖結合不同工具的優點來解決不同的問題 (Guo et al. 2013; Yan & Ma 2015)。

成功的目的地旅遊除了可令遊客感到滿意之外，也可以使其未來仍然願意重遊，並推薦此行程給親朋好友 (Assaker et al. 2011; Chi & Qu 2008)。雖然已有很多的研究對不同的旅遊階段面對的不同問題提出解決方法，然而對於同時考量遊客旅遊動機、景點選擇、人口統計變項資料以及滿意度提出旅遊行程建議或推薦之相關研究則比較缺乏。Watfa & Sobh (2017) 指出雖然先進的資通訊技術已經應用在觀光產業，並也已提出符合某些準則的旅遊行程推薦系統(例如某年齡層女性比較喜歡的旅遊行程)，但大多數遊客對現行的推薦系統仍感到不滿意，尤其

是在個別旅客的特性與偏好並沒有被充分考慮到。換言之，旅遊行程推薦文獻上存在一個明顯的缺口，亦即如何將個人旅遊動機、人口統計變項資料以及個人旅遊滿意度同時納入客製化行程推薦系統。

本研究嘗試結合決策樹、羅吉斯迴歸以及類神經網路模型，建構一精準的客製化旅遊行程推薦系統，針對個別遊客的旅遊動機與人口統計變項資料，找出令遊客滿意的旅遊行程(景點組合)。本論文結構如下：第二節進行相關文獻探討；第三節說明資料來源與取得方式並介紹分析方法；第四節為資料分析與實證結果；第五節則提出研究結論、說明研究限制以及建議未來研究方向。

貳、文獻探討

一、行程推薦

旅遊行程推薦的相關研究文獻可以概分旅行前與旅行中兩類，在旅遊前的行程推薦方面，除了旅行社現成的套裝行程可以選擇之外，相關研究包括單一景點推薦 (Lin & Huang 2009)、多景點推薦 (Shapoval et al. 2018)、給定行程資源限制下極大化其效用的行程推薦 (Chiang & Huang, 2015) 以及考量滿意度的行程推薦(彭定國 2017)等。對自由行的遊客而言，最受歡迎的旅遊行程是哪些景點 (Lin & Huang 2009)？何種景點組合才是遊客最想要的 (Shapoval et al. 2018)？以及如何做到因遊客特性不同而有不同的行程推薦等 (Lin & Huang 2009)都受到研究者的重視。另外一方面旅行中的景點推薦通常會考量旅客所在的位置、距離、成本及偏好，強調在地化以及即時性的資訊 (Fan et al. 2018)。Chen et al. (2013)使用旅客興趣點(points of interest)的方式為背包客自動設計行程，而 Cai et al. (2018)以地理標記(Geo-tagged)方式為旅客建構行程推薦系統或在已知出發點與終點的前提下，想知道這個行程中可能比較偏好的旅館或遊樂景點等 (Bin et al. 2019; Chang et al. 2016)。這些研究強調行程推薦必須考量所在地的即時情境資訊包括氣候、交通等以提供旅遊規劃參考，而非只是網路上最受歡迎、點閱率最高、或推銷頻率最高的景點。鑒於每個景點即時訊息的普及性在現階段仍有其侷限性，本研究將聚焦於行前的旅遊規劃。

二、旅遊動機

動機是解釋個人行為的最終驅使力 (Mayo & Jarvis, 1981)，可能是來自於個人內在需要不滿足所導致心理或生理上不安的緊張，而旅遊便是用來舒緩此不安緊張的方法之一 (Fodness 1994)。Dann (1977)調查旅遊者的態度找出兩種旅遊動機：逃離與自我強化。Dann (1981)進一步加入外在的影響因素，認為旅遊動機應包括拉力(目的地吸引力)與推力(心理需要)。不同的消費者會因為不同的推拉動機來選擇不同的目的地(Baloglu & Uysal 1996; Kozak 2002; Yoon & Uysal 2005)，推拉理論認為消費者會先受到內部的慾望或著情緒因子的推力而想旅遊，然後會受到外部的有形因子(目的地的屬性)的拉力所吸引而決定去那裏旅遊。而這兩種力量可能是獨立也可能是相互影響。行銷人員必須知悉某些旅遊行為背後的動機，包括為何選擇某一景點，以及該景點的屬性(Yoon & Uysal, 2005)。

Fodness(1994)則提出四種功能性面向來解釋旅遊動機，而此四動機面向也被使用在酒莊遊客(Brown & Getz, 2005)、腳踏車遊客(Ritchie et al. 2010)、及高爾夫遊客(Kim & Ritchie 2012)等的旅遊動機研究。另外一方面，Pearce (1988)以馬斯洛需求層級理論為根基，發展出旅行生涯階梯(Travel Career Ladder, TCL)模型，將旅遊動機分為五種不同的需求，呈現出層級或階梯形式，強調較高層級的動機會包含較低層級的動機；大部分遊客都會系統性地經過各個需求階段，因而可以預測他們的旅遊動機模式。Pearce (2005), Pearce & Lee (2005), Pearce (2011)則修

正 TCL 模型中的層級看法，認為動機也有可能穩定在特定的需要層級上，且取決於遊客健康、財務等因素的穩定性或侷限性，強調在旅遊的型態，可以更有彈性的解釋旅遊動機而形成旅遊生涯模式(Travel Career Pattern, TCP)的內容。

上述在旅遊領域中對於動機的闡述反應了旅客可以有不同的旅遊動機來滿足內在需要或舒緩內心緊張。可見動機對於旅遊而言，是一個極重要的決定因素。Cohen (1972)是最早利用動機來對遊客做分群。Fodness (1994)利用旅遊動機的差異來對旅客做分類，對不同旅遊行程的遊客做分析來強化旅遊行銷。Chi (2006), Kim & Ritchie (2012)以及 Yuan et al. (2005)亦同樣利用動機面向將酒莊遊客、釣客以及高爾夫球遊客，各分成不同的類別。由以上文獻可以得知，好的旅遊推薦必須要考慮到旅客的動機，才能滿足其內心的需求。

三、滿意度與旅遊動機

滿意度在行銷研究中扮演非常重要的角色。要想成功的行銷某一目的地，遊客滿意度非常重要(Devesa et al. 2010; Yoon & Uysal 2005; Zabkar et al. 2010)，直接跟景點的選擇，景點的屬性，以及是否再重遊(Kozak & Rimmington 2000)有很大的關係。Fang et al. (2008)認為消費者的旅遊滿意度代表目的地屬性滿足消費者需求的程度，該訊息可以幫助目的地行銷人員改善產品或服務品質。

遊客的滿意度代表在消費完旅遊產品之後所反應出來的正面感受或喜悅(Beard 1980)，包括衡量旅遊的體驗或者旅遊產品跟遊客的旅遊動機(Buckley 2009)是否符合預期。一般比較常用的理論就是 expectancey-disconfirmation 理論(Oliver 1980)，消費者的滿意度決定於實際的體驗與預期間的差距。Meng & Tepanon (2008)把滿意的面向定義成友善的服務，戶外活動，寄宿品質，以及自然的景色。Tsiotsou & Vasioti (2006)的旅遊滿意度則包括個人滿意度，食物滿意度，行程滿意度，社交滿意度，景色滿意度。Lee (2015)的旅遊滿意度則包括資訊，娛樂，安全以及環境永續等面向。Adam et al. (2019)的滿意度則包括教育，社交，衛生，以及放鬆等面向。旅遊滿意度的衡量面向包括社交，服務，設施等不同的面向，因為不同的旅遊內容而有不同的衡量。

另外，也有不同的研究指出不同的旅遊動機會有不同的預期。動機會影響遊客在評估預期與實際體驗時的比較，而產生不同的滿意程度 Huang et al. (2014)。換句話說旅遊動機會影響滿意度(Lee 2009; Meng & Tepanon 2008; Lee 2007)。綜合以上的文獻，旅遊行程可以視為是一種產品，每一個行程或景點有其不同的特色，旅客會因為不同的動機而購買不同的行程。如果行銷人員可以知悉遊客的旅遊動機以及會選擇某一景點是因為該景點的何種屬性(Uysal et al. 2008)，就可以設計不同的行程來滿足不同客戶的需求。

四、人口統計變項，旅遊動機，與旅遊滿意度

社會人口統計變項，例如年齡，教育程度，收入等是旅遊動機的預測變數，而這些變數與旅遊動機間的關聯性也在過去的研究中被討論過。例如，比較年長的遊客比較喜歡新奇(Jönsson & Devonish 2008; Jang 2007)，但是也有相反的研究結果(Luo & Deng 2008; Kim et al. 2008)，亦即年紀越輕越喜歡新奇。對年輕人來說，他們比較傾向於「娛樂」，「享受」，「尊榮與印象」，以及「尋找社交」(Kim et al. 2008; Jönsson & Devonish 2008; Chen et al. 2013; Jensen 2011)。針對教育程度來看，教育程度越高的遊客會比較喜歡「放鬆與逃避」，「知識尋找」，以及「社交」；而教育程度較低者會比較偏向「尊榮與印象」，以及「新奇」(Kim et al. 2008; Jensen 2011)。就收入來看，高收入者傾向於「崇尚大自然」，「放鬆與逃避」，而低收入者會比較偏向「增廣見聞」(Kim et al. 2008; Jensen 2011)。Ozdemir et al. (2012) 指

出年紀越大，教育程度越高，且收入越低的女遊客有較高的滿意度。Tsiotsou & Vasioti (2006)也指出年長且教育程度較高的遊客的旅遊滿意度比較高。從上述的相關文獻，可以知悉旅遊動機會影響滿意度，人口統計變項會影響旅遊動機與滿意度，因此在討論影響旅遊滿意度的重要因素時，人口統計變項與旅遊動機的影響性不容忽略。

五、資料採礦

科技的進步已使得各行業的經營環境隨著快速改變。硬體技術的進步使資料的蒐集變得更加快速，因而對於如何從不斷累積的龐大資料庫中擷取有用的資訊的技術也產生了更大需求。資料探勘技術的應用目的便是從大量資料中擷取重要的訊息，並轉換成有用的知識來供決策參考 (Fayyad et al. 1996)。現今已經有很多產業利用資料採礦方法來解決不同的問題，例如消費者行為分析 (Hsieh & Chu 2009)、顧客流失預防 (Tsai & Lu 2010)、客戶保留 (Ranjan & Bhatnagar 2010)、及產品銷售的市場菜籃分析 (Shahrabi & Neyestani 2009)等。事實上，旅遊業的經營情境更是需要做大量資料處理 (Buhalis 1998; Sheldon 1994)。遊客的資料以極快的速度累積，要想在激烈競爭環境中存活，目的地管理組織不得不使用資料採礦技術來因應 (Pyo et al. 2002)。

近年來，有很多資料探勘技術被廣泛的討論與應用，包括決策樹，羅吉斯回歸，文字探勘，關聯分析法，以及類神經網路等。而資料探勘的步驟中，資料準備以及簡化是兩個基本的步驟，資料準備指的是處理資料的正確性，就如同“garbage in, garbage out”所描述的現象，資料品質會影響預測的準確度，而且資料必須縮減成可以分析的資料格式。一旦資料被整理乾淨(正確)並簡化成可以分析的資料格式後，就可以選擇適當的工具來分析，例如決策樹，羅吉斯回歸，或著類神經網路等。

Shapoval et al. (2018) 利用決策樹技術分析日本的外國觀光客的旅遊滿意度與重遊意願，其用意是提供給目的地行銷決策的參考依據。彭定國(2017)利用決策樹技術從台中市觀光局七七亮點中篩選出三個主要景點，據以產生伴隨國家歌劇院的旅遊行程。陳建民等人(2016) 利用多項式羅吉斯模型(multinomial logit model) 以及混合式羅吉斯模式(mixed logit model) 探討金門酒廠觀光酒廠遊程服務設計，其中混合式羅吉斯模式考量到遊客可能存在的異質行為，從模型的結果得知哪些參訪行程屬性的內容必須考慮遊客的異質性而有不同的特別設計。Juwattanasamran et al. (2013)利用關聯法則分析找出遊客的旅遊行為與其個人資料間的關係，使旅遊業者可以更有效的回應旅客需求。Wong et al. (2006)利用RFM分析(Recency, Frequency and Monetary Value)找出有價值的遊客，並使用決策樹技術找出這遊客的特性，例如人口統計變數資料、購物行為、造訪旅遊景點等；最後再使用關聯法則分析找出下一次可能的旅遊目的地。Liao et al. (2010)則利用關聯法則分析與集群分析做旅遊新產品開發以及顧客關係管理的應用。

上述資料採礦技術應用的文獻，分別解決了某些特定的問題，對理論與實務都有其重要貢獻，然而對本研究的研究問題「如何針對消費者的不同旅遊動機以及人口統計變項，提出滿意度高的行程規劃」卻沒有具體的回答。例如，Shapoval et al. (2018)的決策樹分析結果能提供遊客願意重遊的旅遊元素組合，而彭定國(2017)的決策樹技術分析也能提供最受歡迎的特定旅遊行程，但對於特定景點的旅遊行程適合具有哪些特性的消費者，並無法精準地回答。

六、新演算法的邏輯

從以上的相關文獻可以得知人口統計變項，旅遊動機對旅遊滿意度有所影響。提供不同的旅客滿意的行程規劃，基本上是人口統計變項，旅遊動機間的配適問題。搭配得好滿意度就高，搭配得不好，滿意度就低。因此如何建構人口統計變項與旅遊動機間的搭配是本研究的中心議題。過去的做法，大都對顧客先依據人口統計變項或著旅遊動機分群(Wong et al. 2006; Chi 2006; Kim & Ritchie 2012; Yuan et al. 2005)，再進一步了解各群組的特性後，針對群組特性來提供行程規劃。或著把人口統計變項與旅遊的行程與經驗一同分類，來解釋旅遊特性與人口統計變項的關聯性(Shapoval et al., 2018)。本研究異於過去的思維，先針對行程規畫做分類，依照旅客的滿意度與否來分類行程，再進一步針對每一個行程了解何種人口統計特性是適合該行程。在分類的過程，必須滿足決策樹的分類準則才會產生子葉，因此可以把非常龐大的行程組合先依據滿意與不滿意做分類，簡化成適度的行程規劃(子葉)。進一步利用羅吉斯回歸或著類神經網路了解甚麼樣的人口統計變項可以對旅遊滿意度增加解釋能力。本研究的特點在於對行程推薦問題的處理是從行程滿意與否的分類開始，採用決策樹的分類作法，對行程類別的簡化有很大的幫助。而進一步針對決策樹子葉(各組行程規劃)做羅吉斯或類神經網路的預測模型建構，除了可以改善決策樹在預測上比較簡化的缺點之外，可以發揮羅吉斯與類神經的模型優勢。

參、研究方法

一、研究流程

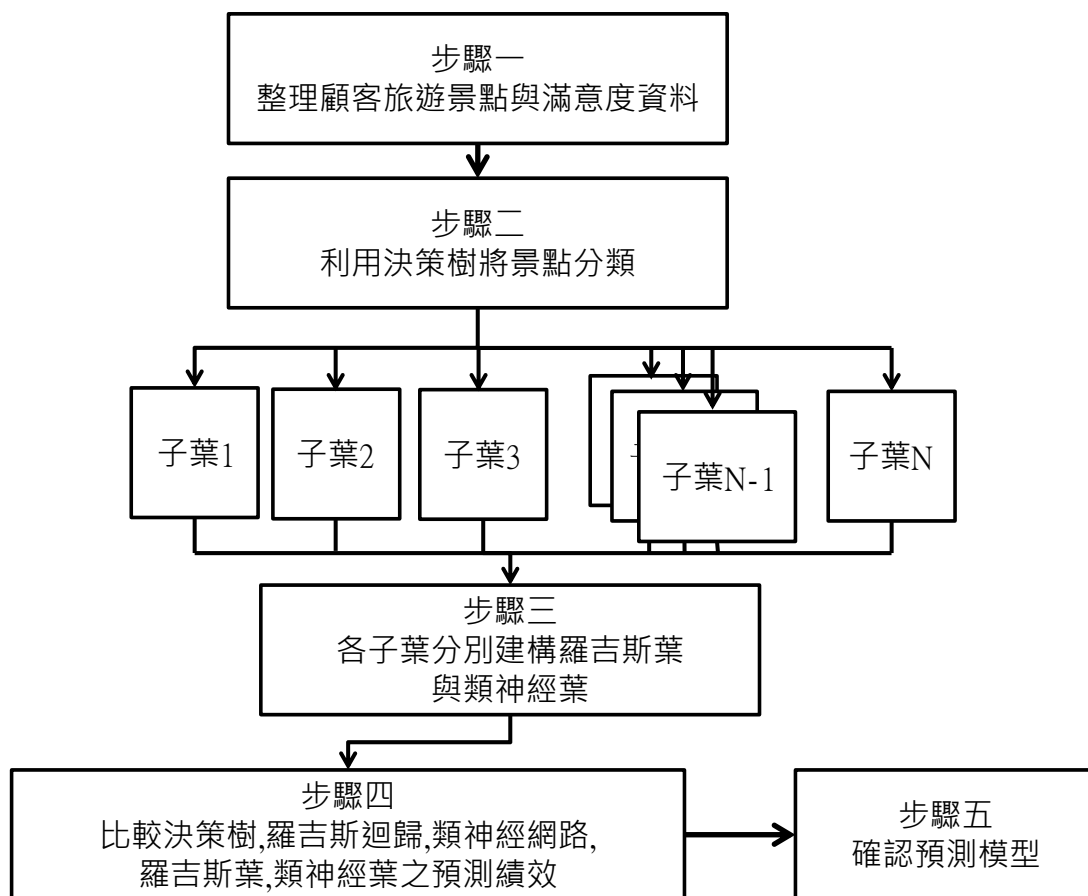


圖 1：預測模型建構

本研究的研究流程可以分為兩個階段。第一階段是建模階段包含五個步驟如圖 1，第二階段則是行程推薦包含四個步驟如圖 2。第一階段的五個步驟如下。步驟一:整理顧客的旅遊景點與滿意度資料。步驟二:以滿意度與否為應變數，景點為解釋變數，利用決策樹將景點組合分類成好幾個子葉，每個子葉代表一個可能的行程規劃。步驟三:針對每個子葉，利用人口統計變項與個人的旅遊動機為解釋變數來配置羅吉斯回歸與類神經網路來預測滿意度，亦即建構羅吉斯葉與類神經葉。步驟四:比較決策樹，羅吉斯回歸，類神經網路，羅吉斯葉以及類神經葉等五種方法的預測績效。步驟五:決定使用預測績效最好的預測模型來預測滿意度。

第二階段的四個步驟如下。步驟一，針對每一位自由行遊客，蒐集其人口統計變項資料及其旅遊動機。步驟二，針對每一子葉所配出的最佳預測模型，代入每一位遊客的人口統計變項資料及旅遊動機資料計算出該遊客對每一子葉所對應旅遊行程的滿意度機率。步驟三，再依滿意度機率排序挑出前五名。步驟四，前五名機率所對應的行程就是對該遊客的最佳旅遊行程規劃推薦。

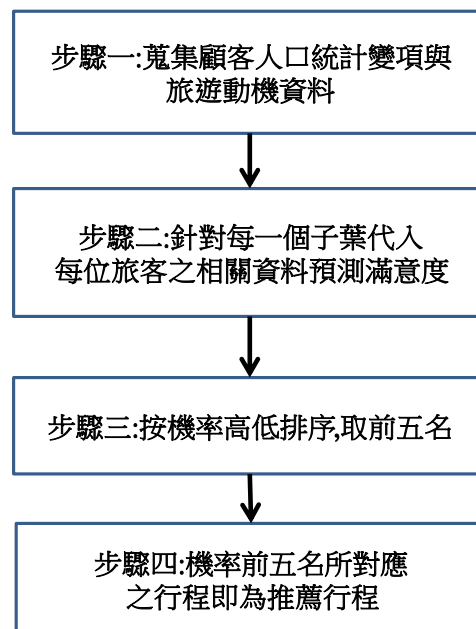


圖 2：行程推薦

本研究選擇用決策樹來對景點分類而不採用常用的集群分析將景點分群，主要原因如下：第一，決策樹的分類層級可以調整，不一定要使用全部的變數，因此分類結果會有不同景點數的子葉(行程)產生，例如有些子葉(行程)包含 4 個景點有些子葉則包含 6 個景點，而且分析人員也可以自行決定是否包含某特定景點。第二，決策樹的分類是以增加整體預測準確度的前提下來進行，除了分類之後的新子葉同質性提高外，整體歸類預測準確度也會增加。然而，一般的集群分析技術是以降低組內變異以及提高組間變異為其目標函數來做分群，因此所得到的結果大大不同。另外，羅吉斯葉與類神經葉的模型建構，主要目的在於改善決策樹的分類結果。就決策樹所得到的子葉而言，每一個子葉成員的歸類預測值都相同，即每一成員都將被歸類到多數成員所屬的類別，即使每一個成員的人口統計變項以及個人旅遊動機可能差異很大。因此，如何建構一個有效的對應函數來改善預測績效是一個非常必要的工作。由於本研究的應變數為滿意與否，我們使用羅吉斯迴歸及類神經網路模式來捕捉人口統計變項及個人旅遊動機等自變數

與滿意度應變數間可能的函數關係。

二、樣本來源

Tsaur et al. (2010)提及自助旅行屬於大類別的旅遊型態，包括所有非購買套裝行程的遊客。在此大類別下所衍生出來的旅遊型態包含半自助旅行、背包客、單車旅行及徒步旅行等。由於學者間對自助旅行還沒有一致的定義，本研究採 Tsaur et al. (2010)的定義，只要不是購買旅行社所推出的套裝行程或對旅行行程擁有自主性的旅遊方式皆屬自助旅行，其中包含背包客與半自助等。為取得本研究運用資料採礦技術所需的遊客資料，本研究針對過去三年曾經到過台中旅遊且其行程屬於非正式化(非套裝)的遊客，於網路上發放問卷。

基本上這是發放問卷蒐集資料的研究，希望可以在特定的時間內搜集到越多的樣本越好，因此考量到樣本大小的問題把資料的蒐集時間拉到多年而非一年內，另外又擔心時間太長會有回溯性誤差的問題，斟酌下將樣本的時間設定在三年內。而目的地的選擇可以是任何的景點，本研究以台中景點為主的原因是因為台中為目前的旅遊熱點之一，在資料蒐集上會比較容易。而實際上，本研究所提之方法適用於各個觀光景點的資料，以到台中地區遊玩的旅客為樣本的研究並不會失去本研究對工具訴求的通用性。另外，本研究強調提出旅途規劃的方法，觀光局的調查資料每年也都有可能不一樣，本研究是針對該結果提出解決旅途規劃的方法，不同的調查結果只要把選項內容作調整即可。

三、變數衡量

本研究的問卷內容主要包括人口統計變項、旅遊動機量表、到訪目的地以及滿意度等。人口統計變項包括性別、年齡、教育程度、職業、婚姻狀況、平均月收入、目前居住地等。旅遊動機量表以 Pearce(1988)提出的「逃離/放鬆」、「新奇/刺激」、「自我發展」、「人際關係」及「自我實現」等五個動機構面為基礎，參考 Pearce et al.(2005)的內容發展出 22 個題項，以 Likert 五點量表衡量。

到訪目的地則以 2017 年觀光局調查出遊客到訪台中最常遊玩的 19 個景點，包括台中國家歌劇院，宮原眼科，台中文學館，分子藥局，勤美誠品綠園道，鳶嘴山，高美濕地，大坑風景區，台中文化創意產業園區，秋紅谷廣場，國立自然科學博物館，彩虹眷村，東海大學(含路思義教堂)，國立台灣美術館，新社，麗寶樂園，梨山(含福壽山農場)，大雪山國家森林遊樂區，武陵國家森林遊樂區。

Bigne et al. (2001) 指出個別屬性滿意度並不是總體滿意度的唯一前提，衡量遊客感受和遊後滿意度都屬於購買後的判斷，因此較適合使用整體衡評估量方式，而不是分析各個屬性的滿意度。Day (1977)認為消費者對購買前、購買後以及使用過程的各方面做出評價，並產生滿意、不滿意或不關心等的感受，這些感受因人而異；以往的不同經歷，及使用環境的改變都可能影響消費者的評價和購買後的行為。許多屬性構面都可以作為衡量消費者對消費體驗的解釋變數，但要找出一組適當的變數來充分衡量消費者的感受並不容易；因此，以涵蓋範圍廣泛的總體滿意度可以概念化消費者的完整評估過程，並降低誤解研究結果的風險。鑒於以上的理由，本研究使用整體滿意度單一變項來衡量滿意度。

四、分析工具

(一)羅吉斯葉模式

羅吉斯葉模式(Logit Leaf Model, LLM)是一種結合決策樹模式及羅吉斯迴歸模式的兩步驟演算法。De Caigny, Coussement, and De Bock (2018)指出，使用 LLM 模式能夠減少決策樹和羅吉斯迴歸的缺點，而同時保留有其原有優勢。在 LLM 模式的第一演算步驟是使用決策樹進行資料分類，亦即將相同性質的資料歸於同

類；在第二演算步驟時，則針對決策樹分類後的各分類資料進行羅吉斯迴歸模式參數估計。LLM 的價值在於它能提高決策樹模式和羅吉斯迴歸模式的預測準確度；其主要原因在於決策樹模式所提供的分類資料之後，使我們易於觀察各分類資料的主要影響變數以及進行後續資料操作。

決策樹模式是一種常用在各種領域的分類預測模型，其優點除具備不錯準確度之外，分析結果也容易解讀，而且其演算過程簡易。決策樹模式主要是使用樹狀分枝的概念來作資料分類，決策樹從頂端的節點(稱為樹根)開始往下分支，每次分支是以最能提升分類後資料集純度的某一變數來將資料集分群，即把該節點分成兩個或多個子節點。如此，重複此分支的過程直到某一特定的停止準則被滿足。最後無法再分支的節點稱為子葉，每個子葉都代表符合某特定法則的所有觀察值所形成的集合，而所有子葉的聯集即是整個原始的資料集。令 S 代表原始的整體資料集， T 為所有的子葉形成的集合，整體資料集 S 和每一個子葉用 S_t 的關係表達如下：

$$S = \cup_{t \in T} S_t; \forall t \neq t'; S_t \cap S_{t'} = \emptyset \quad (1)$$

當子葉越分越多時，代表資料集越分越細，決策樹模式的內容也越來越複雜，且有越來越多的規則來分割資料集。如此，非常容易伴隨著模式過度配適的風險，即模式效度僅能及於樣本內的資料。對於決策樹模式過度配適的問題，可以使用一些方法來避免，例如樹的修剪(pruning)，即設定子葉最大個數或最大分支層級數，抑或每個子葉最少應該有的觀察值個數等。這些參數的設定也是一般決策樹模式用來停止分支的準則。關於每一個子葉中的觀察值，一般是用比例高的類型來作為觀察值的預測群組歸屬。例如某個子葉中有 80% 遊客有重遊意願而其餘則沒有，在歸類上，運用此子葉的法則我們將每位遊客歸類為有重遊意願。

(二)類神經葉模式(Neural Network Leaf Model, NNLM)

類神經網路是常用來建構非線性預測模型的資料採礦工具之一(Bellazzi and Zupan, 2008)。一般而言，類神經網路可以有三種層次的類神經運算元以形成一個模式，即輸入層、中間層以及輸出層，其中間層可以不只一層。類神經網路的每一層都有若干類神經元並與其他層的類神經元分別連結。輸入層的節點代表自變數，而輸出層的節點則代表應變數。在類神經網路模式中，最常被使用的類型就屬倒傳遞類神經網路(Backward Propagation Neural Network, BPNN)。BPNN 可以藉由一種學習機制來記憶及調整輸入與輸出變數間的非線性關係，它已被廣泛的應用在不同領域。原則上，輸入與輸出間的關係並不容易解釋，因此被稱為黑箱。

類神經網路的中間層與輸出層的神經元節點都有活化函數(Activation Function)，而中間層或輸出層的每一個神經元節點將所有與它連節的前一層的神經元節點的輸出值與其連結權重相乘後加總，經過該節點的活化函數運算再輸出成為下一層神經元的輸入值。每一層的神經元個數要依所要求解問題的複雜程度來決定。以只有一個中間層的類神經網路為例，假設有 n 筆觀察值， M 表示中間層的神經元數、 p 表示輸入層的神經元數、 w_j 與 w_{js} 為連結權重，則第 i 筆觀察值的網路輸出值為：

$$\hat{y}_i = \phi_2 \left(b_0 + \sum_{j=1}^M w_j \phi_1 \left(b_{j0} + \sum_{s=1}^p x_{is} w_{js} \right) \right) \quad i = 1, 2, \dots, n \dots (2)$$

其中 ϕ_1 與 ϕ_2 為活化函數， b_{j0} 與 b_0 為每個節點的門檻值。中間層的活化函

數通常都是非線性，而輸出層的活化函數可以是線性或非線性。類神經網路的訓練過程就是要計算與調整節點與節點間的連結權重，而倒傳遞類神經網路的訓練過程通常包含兩個階段。第一階段是由前往後，即從輸入層開始計算所有輸入層變數與其對應的連結權重的乘積和，再進行中間層節點的非線性轉換，其結果則成為中間層的輸出，再經同樣的演算步驟計算中間層的輸出與其連結權重的乘積和做為輸出層節點的輸入，經再一次轉換即可得到對應的輸出。這個輸出值與實際值之間的誤差，將做為第二階段調整連結權重的依據。換言之，第二階段是由後往前運算，亦即在算出實際值與輸出值的誤差後，由後往前檢視，看看中間層與輸出層間那一條連結的權重調整會使誤差項獲得最有效的修正，然後再往前一層探究，輸入層與中間層的哪一條連結的權重調整可以使誤差項的修正最有效，然後調整這些參數，包括活化函數中的截距(門檻值)與斜率(連結權重)。調整過程的步伐大小與學習慣性的選擇都是學習過程中的重要參數，整個學習過程的目標函數就是誤差項的平方和 (Adavi et al. 2016; Sedehi et al. 2011)。

肆、實證結果

一、資料集

本研究使用網路問卷方式調查近三年內曾到訪台灣中部地區旅遊的台灣民眾。總計回收 440 份問卷，扣除填答不完整、重覆填答、相同網路 IP 之無效問卷外，有效問卷總共有 411 份，有效問卷比例為 93.41%。每一份問卷的資料，除了需要回答問卷填答者的人口統計變項如性別、年齡、職業別之外，還包括填答者最近這次的台中旅遊到過哪些景點、旅遊動機及整體滿意度等。問卷發放期間從 2019 年 1 月 16 日至 2 月 16 日；網路問卷邀請張貼於 Facebook 旅遊社團、PTT 台中版等自助旅行討論區。

二、基本人口統計變項

基本人口統計變項的分析結果如表 1 所示。填答者的男女比例分別為 45.2% 與 54.8%，女性人數比男性稍為偏多。在年齡分佈上，最多為 20-29 歲佔 59.1%，其次為 30-39 歲佔 27.3%，兩者合計共 86.4%，顯示自由行遊客仍以年青族群為主。在教育程度方面，大學生佔 64.1%，研究所以上佔 30.9%，高中職學歷佔則只佔 4.3%。在職業分佈方面，學生佔 31.8%、服務業佔 20.9%、製造業佔 18.4%、軍公教佔 11.4%、自由業佔 7%、家管及已退休者 2.6%、農林漁牧業 0.4%、其他則佔了 7.5%。在婚姻部分，未婚族群佔 86.4%、已婚且小孩未獨立佔 7%、已婚無子女佔 5.5%、已婚且小孩已獨立佔 1.1%。在月收入的方面，20000 元以下佔 34.5%、20001-40000 元佔 38.4%、40001-60000 元佔 19.3%。在居住地部分，中部地區佔 42%、北部地區佔 43%、南部地區佔 15%。

旅遊滿意度的平均為 3.99，標準差 0.54，顯示受訪者對於最近一次到訪台中旅遊的整體滿意度還算不錯。五個旅遊動機面向的 KMO 值、Bartlett 球形檢定卡方值及單一因素解釋變異量彙整如表 2 所示。由表 2 可知，每個構面的 KMO 值都大於 0.5，卡方檢定也都顯著，代表每個構面的變數適合作資料簡化。每一個構面因素分析後的第一個因素的解釋變異量都超過 50%，顯示因素分析後的五個構面因素可以相當程度的代表整個動機資料集的變異程度。我們依序為五個構面命名為 M01~M05，以利進行後續的羅吉斯迴歸以及類神經網路模型運算。

表 1：人口變數統計量

變數	項目	次數	百分比	累計百分比
性別	男	199	45.2%	45.2%
	女	241	54.8%	100.0%
年齡	19 歲以下	27	6.1%	6.1%
	20-29 歲	260	59.1%	65.2%
	30-39 歲	120	27.3%	92.5%
	40-49 歲	29	6.6%	99.1%
	50 歲以上	4	0.9%	100.0%
教育程度	國中及以下	3	0.7%	0.7%
	高中	12	2.7%	3.4%
	專科	7	1.6%	5%
	大學	282	64.1%	69.1%
	研究所及以上	136	30.9%	100.0%
職業	學生	140	31.8%	31.8%
	軍、公、教	50	11.4%	43.2%
	農林漁牧業	2	0.4%	43.6%
	製造業	81	18.4%	62.0%
	家管、已退休	11	2.6%	64.6%
	服務業	92	20.9%	85.5%
	其他	33	7.5%	100.0%
婚姻狀況	未婚	380	86.4%	86.4%
	已婚(無子女)	24	5.5%	91.9%
	已婚(小孩未獨立)	31	7.0%	98.9%
	已婚(小孩已獨立))	5	1.1%	100.0%
平均月收入	20000 元以下	152	34.5%	34.5%
	20001-40000 元	169	38.5%	73.0%
	40001-60000 元	85	19.3%	92.3%
	60001-80000 元	26	5.9%	98.2%
	80001 元以上	8	1.8%	100.0%
居住地	北部地區	185	42.0%	42.0%
	中部地區	189	43.0%	85.0%
	南部地區	66	15.0%	100.0%

表 2：旅遊動機構面 Bartlett 及 KMO 值分析

構面	KMO值	卡方值(顯著性)	解釋變異量
逃離/放鬆	0.749	510.942***	61.662%
新奇/刺激	0.683	620.264***	61.489%
人際關係	0.781	789.339***	66.584%
自我發展	0.860	1594.198***	67.576%
自我實現	0.820	1152.270***	78.201%

註：*p <.1；**p <.05；***p <.01

三、決策樹模式分析結果

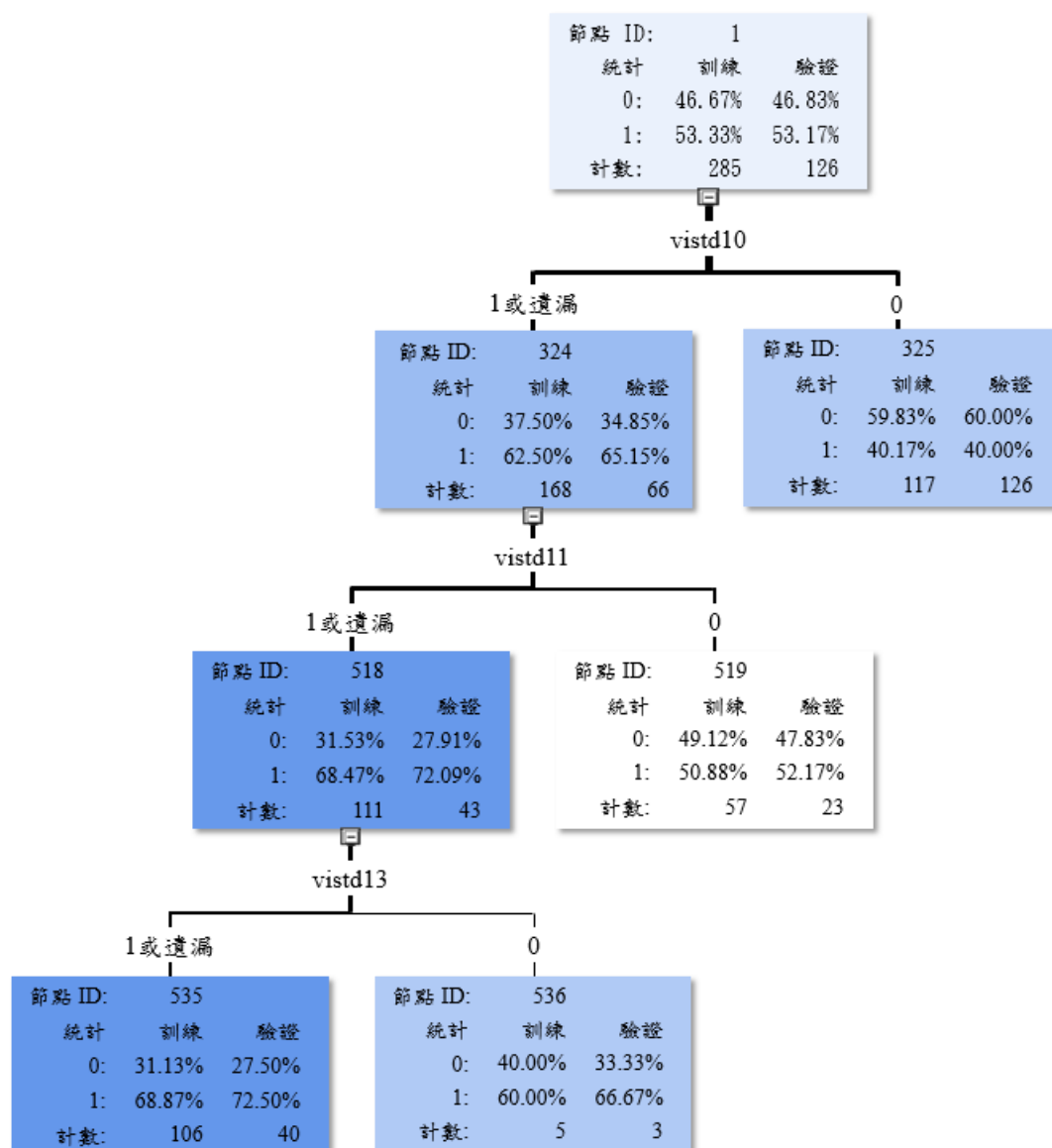


圖 3：決策樹樹狀圖

本研究先利用決策樹模式針對景點來對整個資料做分類。每筆資料代表一位旅客，應變數 y 代表滿意程度，令 $y=1$ 代表滿意(滿意度大於等於 3.5 者)， $y=0$ 代表不滿意(滿意度小於 3.5 者)。自變數 $D1\sim D19$ 為代表台中熱門景點的 19 個虛擬變數，例如 $D1=1$ ，代表此次旅遊該遊客有到景點 1，若 $D1=0$ 則代表此次旅遊沒有到過景點 1。

我們將資料集輸入到決策樹分析軟體之後，得到的樹狀圖如圖 3 所示。我們將整個資料集以 70% 與 30% 的比例隨機分成訓練資料集與測試資料集。第一層的根節點 ID1 顯示對所此次旅遊行程滿意的百分比為訓練集 53.33%，驗證集 53.17%，這個數字代表還未作任何子節點分割時的滿意預測準確度。

在資料蒐集過程，19 個景點有到過的打勾，在編碼過程都轉換成 1，因此沒有到過的景點一率都填 0，不會有遺漏值產生。由於 sas 軟體在決策樹的分枝表達上，0 的對應會寫成「1 或遺漏」，實際上所代表的就是 1。因此在解釋上可以把遺漏的情況忽略。

節點 ID324 表示到過景點 10 的遊客中，在訓練集有 62.5% 的遊客滿意，驗證集則為 65.15%。針對節點 ID324 進一步分割，得到下一層子節點之一的 ID518 表示到過景點 10 與景點 11 的遊客中，在訓練集有 68.47% 的遊客滿意，驗證集則有 72.09%。再從節點 ID518 往下一層做子節點分割，所得到子節點之一 ID535 表示到過景點 10、11 與 13 的遊客中，在訓練集有 68.87% 的遊客滿意，驗證集則有 72.50%。在完整的決策樹中，整個資料集可以用四個葉節點來表示，即 ID535、536、519 及 325。除了節點 ID519 之外，其餘三個葉節點的滿意預測準確度都比根節點 ID1 的準確度 53.3% 還要高，而所提高的預測準確度正是決策樹演算帶來的效用。葉節點 ID535 的規則為 $IF(\text{景點 } 10=1)AND(\text{景點 } 11=1)AND(\text{景點 } 13=1)$ 則(滿意=是)，歸類預測準確度為 68.87%。

四、羅吉斯葉分析結果

為進一步改善決策樹分析所得子葉的歸類預測準確度，我們分別對每個子葉中的遊客，以人口統計變項、旅遊動機及旅遊行為作為解釋變數，滿意與否作為應變數，建構一羅吉斯迴歸模式。以葉節點 ID535 為例，其羅吉斯葉模式估計如表 3 所示。

表 3 的羅吉斯迴歸結果顯示，參加 ID535 旅遊行程的旅客，若其旅遊動機越是偏向尋找新奇/刺激或著追求自我實現者，滿意度將會越高。因此行銷機構可以據此設計對應的行銷方案。

五、類神經葉分析結果

同樣為了改善第一階段決策樹模式的準確度，我們也針對決策樹子葉節點配置類神經網路預測模型。在此階段的類神經模式建置仍然以遊客的人口統計變項以及旅遊動機作為模式的解釋變數，也以滿意與否作為模式的類別應變數，來進行類神經模式的訓練。以決策樹模式的葉節點 ID535 為例，其第二階段類神經網路模式(類神經葉)的訓練集與驗證集歸類預測正確率分別為 85.85% 及 85.00%。如果我們不做第一階段的決策樹模式建置而直接進入第二階段，亦即直接用整個資料集的所有變數(包括 19 個旅遊景點虛擬變數、人口統計變項及旅遊動機構面)，及滿意與否作為類別應變數，則所配置的類神經網路模式的歸類預測準確率分別為訓練集 72.63% 與驗證集 73.81%，明顯低於類神經葉模式的預測績效。

表 3：羅吉斯迴歸之結果

變數	類別	係數	標準誤	P 值	勝算比
截距		2.1977	23.5847	0.9258	
逃離/放鬆		0.1658	0.1801	0.3572	1.180
新奇/刺激		0.7208	0.2410	0.0028***	2.056
自我發展		-0.3260	0.2636	0.2161	0.722
自我實現		0.7756	0.2387	0.0012***	2.172
婚姻狀況 1 ^a	1 vs 4	-1.4742	23.5976	0.9502	0.002
婚姻狀況 2 ^b	2 vs 4	-2.5505	23.5967	0.9139	<.001
婚姻狀況 3 ^c	3 vs 4	-0.8169	23.5942	0.9724	0.003
年齡 1 ^d	1 vs 5	-0.2083	0.9102	0.8189	0.134
年齡 2 ^e	2 vs 5	-0.9982	0.7500	0.1832	0.061
年齡 3 ^f	3 vs 5	-0.5918	0.7386	0.4230	0.092
性別 1 ^g	1 vs 2	0.2085	0.1467	0.1552	1.517
旅遊次數		0.0461	0.0482	0.3398	1.047

註：* $p < .1$ ；** $p < .05$ ；*** $p < .01$

^a 未婚；^b 已婚(無子女)；^c 已婚(小孩未獨立)；^d 19 歲以下；^e 20-29 歲；^f 30-39 歲；^g 男

六、模式預測準確度比較

為了清楚說明本研究所提的方法的確有比較好的預測準確度，我們將比較不同方法的預測準確度。首先，我們直接將所有 411 筆資料分別進行決策樹模式、羅吉斯迴歸模式與類神經網路模式的預測模型建構訓練。三個模式所使用的輸入變數(自變數)皆相同，包含 19 個旅遊景點虛擬變數、人口統計變項及旅遊動機構面，而應變數為滿意與否類別變數。三個模式的歸類預測結果如表 4 單一模式列所示，訓練集與驗證集的預測準確度約略相同，訓練集大致上為 73% 左右，而驗證集大約為 74% 左右。

一般來說訓練集的預測準確度應該會比驗證集的預測準確度還要高。因為模型是用訓練集的資料學習而建構出來，用原來訓練集的資料代進所訓練出來的模型去作預測，通常會有不錯的預測結果，因此若用訓練集的準確度來評比，會高估該模型的預測準確度。然而，驗證集的準確度比訓練集的準確度高的情況也還是會發生。如果驗證集的準確度遠大於訓練集的準確度，有可能兩個樣本的屬性並不相同。如果兩個資料集的準確度沒有差太多，就代表該模型可以描述該資料集的特性。如果訓練集的準確度遠高於驗證集的準確度就代表該模型有過度配適的疑慮。本研究的結果，雖然測試集的答案比訓練集好，但相差不大，代表該模型可以描述該資料集。

接著則比較羅吉斯葉模式與類神經葉模式此二個結合模式(決策樹結合羅吉斯、決策樹結合類神經)的歸類預測績效表現。二步驟結合模式的第一步驟皆為建置決策樹模式，所使用的分枝變數為 19 個旅遊景點虛擬變數，而應變數為滿意與否類別變數。如前所述，此階段共得到 4 個葉節點(代表 4 種旅遊行程)，亦即將原始資料集分為 4 個子集。接著在第二步驟則分別針對每個子葉節點分別配置羅吉斯迴歸與類神經網路模式，此步驟使用的自變數皆為人口統計變項及旅遊動機構面，而應變數仍為滿意與否類別變數。表 4 的結合模式列顯示，羅吉斯葉

模式的歸類預測準確率訓練集 78.30%、驗證集 77.50%，而類神經葉模式的歸類預測準確率訓練集 85.85%、驗證集 85.00%。

表 4：模式歸類預測準確度比較

類別	模型	訓練集正確率	驗證集正確率
單一模式	決策樹模式	72.98%	74.60%
	羅吉斯迴歸模式	73.33%	75.40%
	類神經網路	72.63%	73.81%
結合模式	羅吉斯葉模式	78.30%	77.50%
	類神經葉模式	85.85%	85.00%

為了讓方法間的差異可以被清楚的辨識出來，本研究分別針對訓練集與測試集作方法間預測準確度的兩兩比較作檢定，檢定結果如表 5 與表 6 所示。表 5 中的第一列第一欄的 -0.0035 是決策樹的預測準確率減羅吉斯預測準確率所得到的值，檢定結果不顯著，代表兩種方法的預測績效沒有顯著差異。另外，第一欄第 4 列的 -0.1287 代表決策樹與類神經葉的預測準確度差，有顯著差異，代表類神經葉的預測績效顯著的優於決策樹，其餘的比較結果如表所列。整體來看，類神經葉顯著的優於羅吉斯葉以及其他單一的方法。羅吉斯葉雖然準確率都比單一方法要好，但統計上並不顯著。可以確認的是，決策樹加上類神經網路(類神經葉)的績效顯著優於決策樹與類神經方法的獨自表現。換句話說，在行程推薦的這個議題上，類神經葉的作法可以較過去的單一作法(亦即純粹決策樹或純粹類神經網路)有顯著的績效提升。

表 5：訓練集資料各方法的預測準確度的 t 檢定結果

	決策樹	羅吉斯迴歸	類神經網路	羅吉斯葉
回歸	-0.0035			
類神經網路	0.0035	0.007		
羅吉斯葉	-0.0532	-0.0497	-0.0567	
類神經葉	-0.1287***	-0.1252***	-0.1322***	-0.0755 **

註：* $p < .1$ ；** $p < .05$ ；*** $p < .01$

表 6：測試集資料各方法的預測準確度的 t 檢定結果

	決策樹	羅吉斯回歸	類神經網路	羅吉斯葉
羅吉斯回歸	-0.008			
類神經網路	0.0079	0.0159		
羅吉斯葉	-0.029	-0.021	-0.0369	
類神經葉	-0.104***	-0.096***	-0.1119***	-0.075**

註：* $p < .1$ ；** $p < .05$ ；*** $p < .01$

七、討論

本研究欲解決的問題是要能夠針對遊客個人的人口統計變項與旅遊動機，精準地推薦出喜愛的旅遊行程景點組合，以使遊客有高滿意度而願意再度重遊。我們所建議的方法，基本精神是在於建構自變數與應變數間的對應關係，而自變數包含遊客的人口統計變項資料、旅遊動機構面與到訪旅遊景點，應變數則為滿意與否類別變數。一般來說簡單的作法是將所有變數資料投入，建構單一預測模式，表 4 的單一模式列即呈現此種做法的結果，亦即決策樹模式、羅吉斯模式及類神經模式等三種模式訓練集準確度分別為 72.98%、73.33% 及 72.63%，而驗證集準確度則分別為 74.60%、75.4% 及 73.81%。

而本研究所主張的二階段旅遊行程推薦法的基本想法在於不同類群的遊客應該會有不同類型的旅遊需要，而影響旅遊滿意度的原因也不會一樣，因而各類群遊客的人口統計變項、旅遊動機影響旅遊是否滿意的函數關係也該不同。因此，在建構旅遊滿意與否的預測模式之前，若能對旅客作一個有效的分群或著對景點可以有一個預測前的分類，對於旅遊滿意度的預測準確度將會有很大的助益。

現存有許多分群方法，例如多變量集群分析技術又可再分為階層式與非階層式兩類。其中，階層式集群分析演算法中又有融合式或裂解式的不同做法，甚至也會因為要融合或裂解的兩群間距離如何認定而使最終分群結果產生差異；非階層式集群分析方法也會因為起始群的中心點或演算終止準則等參數設定的不同而使分群結果有所差異。因此，如何決定一個合理的分群做法，對於問題解決成效會是一個重要的決定性步驟。異於過去對旅客分群的邏輯，本研究使用決策樹對旅遊景點先做第一階段的分類。一般而言，決策樹模式的自變數選擇非常具有彈性，選擇不同自變數的分析結果對實務有不同的意涵，進而影響決策制定或實務操作。本研究以是否到訪某旅遊景點做為決策樹模式的自變數(二元虛擬變數)，決策樹模式分析結果所得到的子葉將代表一個旅遊行程景點組合，對於旅遊實務的管理意涵有較大貢獻。

接著，我們針對決策樹所得到的子葉分別建置羅吉斯迴歸與類神經網路模式。換言之，我們將所有自變數分成兩部分，第一部分的旅遊景點造訪與否先投入決策樹模式分析，而第二部分則包含人口統計變項及旅遊動機構面投入羅吉斯迴歸或類神經網路模式的模型建構。從表 4 的結合模式列可以清楚看出，不論是羅吉

斯葉或類神經葉模式，在訓練集資料與驗證集資料都顯著比第一階段決策樹的歸類預測準確度要高，代表結合模式的羅吉斯迴歸或類神經模式對歸類預測準確度可以有進一步的提升。

在第二階段，我們針對每一個想要做行程規劃的旅客，將其人口統計變項資料及旅遊動機構面資料輸入各個子葉(不同的景點組合行程)的羅吉斯葉或類神經網路葉模式，計算該名旅客在各個子葉(旅遊行程)的滿意機率，再將各行程滿意度的機率依序排列，推薦前五名的行程清單就是對該遊客的最佳旅遊行程的精準推薦。

伍、結論與建議

一、結論

近年來雖然資通訊科技硬體技術成熟進步及旅遊網站與旅遊行程系統蓬勃發展，能針對遊客的個人旅遊動機及人口統計變項資料，精準推薦與提供令遊客真正滿意的軟體應用仍嫌不足。本研究嘗試組合運用資料探勘技術，針對精準行程推薦的問題本質提出一個二階段方法來解決此問題。第一階段為使用特定旅遊景點到訪與否的二元虛擬變數做為自變數，旅遊行程滿意與否做為應變數，利用決策樹模式的技術，試圖找出不同的旅遊行程分類。決策樹模式分析所得到的子葉節點即代表一個特定的旅遊行程景點組合。接著用各別遊客的人口統計變項與旅遊動機資料作為自變數，建構與旅遊行程滿意與否間的預測模型，並比較不同預測工具的預測績效。第二階段則是使用個別遊客的人口統計變項與旅遊動機資料，帶入各個子葉(一個旅遊行程)計算滿意度的機率並加以排序，最後將前五名機率所對應的旅遊行程推薦給該遊客。實證分析結果顯示本研究所建議的方法確實可以提供因人而異的旅遊行程推薦建議，而達到精準行銷的目的。

二、研究發現與管理意涵

近年來已有愈來愈多的研究從不同角度來討論遊客旅遊行程設計與推薦，例如從效用極大化的觀點來考慮旅遊行程設計，在給定時間、距離及成本等限制下，使用適當演算法找出效用最大的旅遊行程路線 (Chiang & Huang 2015)。本研究則以遊客的對旅遊景點偏好為主，試圖找出遊客會滿意的旅遊行程。由於旅遊動機為影響遊客旅遊行程選擇與旅遊行為的驅動力，本研究亦將旅遊動機納入旅遊行程推薦的重要考量因素。實證結果顯示，旅遊動機的確有助於解釋旅遊行程滿意的歸類預測正確率的顯著改善，此結果也驗證了研究文獻所指出的旅遊動機會顯著影響遊客對於旅遊行程的選擇。

雖然資通訊科技的進步成熟，智慧旅遊與精準行銷的觀念也甚囂塵上，但網路上常見的旅遊行程推薦系統仍然偏向於只是現成套裝行程的選擇，而這種套裝旅遊行程大部分都是由旅行業者主動決定並沒有真正參考旅客個人的差異或旅遊需要，這也是系統使用者不滿意的主要原因之一。本研究所提出的二階段旅遊行程推薦方法，以遊客個人的人口統計變項資料與其旅遊動機資料為基礎，經由所建置的資料探勘技術組合模式的數據運算，而向遊客推薦令其滿意的客製化旅遊行程，這種精準與客製化的方法相較於實務上常用的大眾觀光之行銷區隔作法而言，可以更精準地找到令顧客滿意的行程。精準旅遊行程推薦除了可以贏得遊客喜愛及降低行銷成本之外，也可針對每一旅遊行程的預測模式所指出的滿意與不滿意的遊客進一步做差異分析，以制定更有效的行銷策略來贏得顧客忠誠。除此，決策樹模式所找出的旅遊行程，除可以用來作包括景點間往返的交通、旅館住宿及餐飲等異業結合的依據，也可以有助於旅遊景點間的聯合行銷。

三、研究限制與建議

本研究基於時間及成本的限制，於 2019 年 1 月初於網路進行調查問卷之發放，而選定的網路旅遊社團及網路討論版或許會有樣本代表性的疑慮。但本研究主要的目的乃在於提供一個關於資料探勘工具組合運用的客製化精準旅遊行程推薦對策，以利於解決精準行銷的問題，而不在於提供明確的解答。若未來能建置一個包含遊客到訪景點、人口統計變項及旅遊動機等資料的大型資料庫，並且運用本研究所建議的方法，則將可以產生更多、更深入的管理意涵。在決策樹模式的分析結果將產生更多不同景點組合的葉節點，例如包含 7 個景點或 10 個景點組合的旅遊行程推薦等，對遊客提供更多的選擇。另外，影響旅遊滿意度的因素非常的多，本研究聚焦在旅遊行程的規劃對滿意度的影響，遺漏的重要變數對現存變數影響滿意度可能造成的偏誤只能盡量地以抽樣的隨機性來降低，未來研究除了可以再增加其他的重要變數之外，也可以藉由更大量的資料與抽樣的隨機性來增加其準確性。另外，在過度配適的辨識過程，本研究採取訓練集與測試集的 cross validation 作法而沒有採用 k-fold 的作法，主要原因在於 19 個旅遊景點的自由組合下，將造成各個子葉下樣本數太少的問題，未來如果有大量的旅客資料，在過度配適的辨識過程可以考慮使用 k-fold 的作法。

本研究的問題聚焦於出發前的行程規劃，而每一位遊客的行程規劃會因其旅遊目的與個人喜好及人口統計變項的差異而不同。然而在旅遊過程中，當遊客到達某一景點時，可能對附近的景點也需要有類似的推薦。因此，如何調整既有的資料探勘模式且兼顧即時性與在地性，進一步提出針對遊客個人旅遊動機、人口統計變項的差異來推薦附近景點，可作為一個未來研究的有趣議題。此外，雖然本研究將遊客旅遊動機資料納入資料探勘模式，建議後續研究者可以再增加其他可能的重要變數，例如遊客個人人格特質、過去旅遊行為等。

參考文獻

- 陳建民、蔡宗憲、李珮榕(2016),「遊客需求導向之觀光酒廠遊程服務設計」, *觀光休閒學報*, 第二十二卷, 第二期, 頁 183-212。
- 彭定國(2017),「利用資料探勘技術規劃文化設施為主題之旅遊行程-以台中國家歌劇院為例, *觀光休閒學報*」, 第二十三卷, 第一期, 頁 33-65。
- Abodeeb, J., Wilson, E., & Moyle, B. (2015). Shaping destination image and identity: insights for Arab tourism at the Gold Coast, Australia. *International Journal of Culture, Tourism and Hospitality Research*, 9(1), 6-21.
- Adavi, M., Salehi, M., & Roudbari, M. (2016). Artificial neural networks versus bivariate logistic regression in prediction diagnosis of patients with hypertension and diabetes. *Medical journal of the Islamic Republic of Iran*, 30(1), 1-5.
- Adam, I., Adongo, C. A., & Amuquandoh, F. E. (2019). A structural decompositional analysis of eco-visitors' motivations, satisfaction and post-purchase behaviour. *Journal of Ecotourism*, 18(1), 60-81.
- Akin, M. (2015). A novel approach to model selection in tourism demand modeling. *Tourism Management*, 48, 64-72.
- Allaberganov, A., & Preko, A. (2021). Inbound international tourists' demographics and travel motives: views from Uzbekistan. *Journal of Hospitality and Tourism Insights*.
- Alshammari, F., Whaley, J., Hur, S., & Kim, Y. K. (2019). Gender differences in motivations to attend festivals in Saudi Arabia. *International Hospitality Review*. 33(2), 126-141.

- Assaker, G., Vinzi, V. E., & O'Connor, P. (2011). Examining the effect of novelty seeking, satisfaction, and destination image on tourists' return pattern: A two factor, non-linear latent growth model. *Tourism management*, 32(4), 890-901.
- Baloglu, S., & Uysal, M. (1996). Market segments of push and pull motivations: A canonical correlation approach. *International journal of contemporary Hospitality Management*, 8(3), 32-38
- Beard, J. G., & Ragheb, M. G. (1980). Measuring leisure satisfaction. *Journal of leisure Research*, 12(1), 20-33.
- Bellazzi, R., & Zupan, B. (2008). Predictive data mining in clinical medicine: current issues and guidelines. *International journal of medical informatics*, 77(2), 81-97.
- Bigne, J. E., Sanchez, M. I., & Sanchez, J. (2001). Tourism image, evaluation variables and after purchase behaviour: inter-relationship. *Tourism management*, 22(6), 607-616.
- Bin, C., Gu, T., Sun, Y., Chang, L., & Sun, L. (2019). A travel route recommendation system based on smart phones and IoT environment. *Wireless Communications and Mobile Computing*, 2019.
- Brown, G., & Getz, D. (2005). Linking wine preferences to the choice of wine tourism destinations. *Journal of travel research*, 43(3), 266-276.
- Buckley, R. (2009). *Ecotourism: Principles and practices*. CABI.
- Buhalis, D. (1998). Strategic use of information technologies in the tourism industry. *Tourism management*, 19(5), 409-421.
- Cai, G., Lee, K., & Lee, I. (2018). Itinerary recommender system with semantic trajectory pattern mining from geo-tagged photos. *Expert Systems with Applications*, 94, 32-40.
- Chattalas, M., Kramer, T., & Takada, H. (2008). The impact of national stereotypes on the country of origin effect: A conceptual framework. *International Marketing Review*, 25(1), 54-74.
- Chang, H. T., Chang, Y. M., & Tsai, M. T. (2016). ATIPS: automatic travel itinerary planning system for domestic areas. *Computational intelligence and neuroscience*, 2016, 1-13.
- Chen, G., Wu, S., Zhou, J., & Tung, A. K. (2013). Automatic itinerary planning for traveling services. *IEEE transactions on knowledge and data engineering*, 26(3), 514-527.
- Chi, C. G. Q., & Qu, H. (2008). Examining the structural relationships of destination image, tourist satisfaction and destination loyalty: An integrated approach. *Tourism management*, 29(4), 624-636.
- Chi, Y. N. (2006). Segmenting fishing markets using motivations. *E-Review of Tourism Research*, 4(3), 64-73.
- Chiang, H. S., & Huang, T. C. (2015). User-adapted travel planning system for personalized schedule recommendation. *Information Fusion*, 21, 3-17.
- Cohen, E. (1972). Toward a sociology of international tourism. *Social research*, 39(1), 164-182.
- Dann, G. M. (1977). Anomie, ego-enhancement and tourism. *Annals of tourism research*, 4(4), 184-194.
- Dann, G. M. (1981). Tourist motivation an appraisal. *Annals of tourism research*, 8(2), 187-219.
- Day, R. L. (1977). Extending the concept of consumer satisfaction. *ACR North American Advances*, 4(1), 149-154.
- De Caigny, A., Coussement, K., & De Bock, K. W. (2018). A new hybrid classification algorithm for customer churn prediction based on logistic

- regression and decision trees. *European Journal of Operational Research*, 269(2), 760-772.
- Devesa, M., Laguna, M., & Palacios, A. (2010). The role of motivation in visitor satisfaction: Empirical evidence in rural tourism. *Tourism management*, 31(4), 547-552.
- Fan, L., Bonomi, L., Shahabi, C., & Xiong, L. (2018). Optimal group route query: Finding itinerary for group of users in spatial databases. *GeoInformatica*, 22(4), 845-867.
- Fang, M., Yodmanee, T., & Muzaffer, U. (2008). Measuring tourist satisfaction by attribute and motivation: The case of a nature-based resort. *Journal of Vacation Marketing*, 14(1), 41.
- Fayyad, U., Piatetsky-Shapiro, G., & Smyth, P. (1996). From data mining to knowledge discovery in databases. *AI magazine*, 17(3), 37-37.
- Fodness, D. (1994). Measuring tourist motivation. *Annals of tourism research*, 21(3), 555-581.
- Gidey, Y., & Sharma, K. (2017). Tourists satisfaction in tourist destination (A study of Tigray-Ethiopia). *International Journal of Research in Finance and Marketing*, 7(4), 138-151.
- Guo, Z. X., Wong, W. K., & Li, M. (2013). A multivariate intelligent decision-making model for retail sales forecasting. *Decision Support Systems*, 55(1), 247-255.
- Hsieh, N. C., & Chu, K. C. (2009). Enhancing consumer behavior analysis by data mining techniques. *International journal of information and management sciences*, 20(1), 39-53.
- Huang, Y., Luo, S., Ding, P., & Scott, N. (2014). Impressions of Liusanjie: a study of motivation, theatrical performance evaluation, and satisfaction. *Current Issues in Tourism*, 17(3), 280-296.
- Jang, S. S., & Feng, R. (2007). Temporal destination revisit intention: The effects of novelty seeking and satisfaction. *Tourism management*, 28(2), 580-590.
- Jensen, J. M. (2011, November). The relationships between socio-demographic variables, travel motivations and subsequent choice of vacation. In *2nd International Conference on Economics, Business and Management*, 22, 37-44.
- Jönsson, C., & Devonish, D. (2008). Does nationality, gender, and age affect travel motivation? A case of visitors to the Caribbean island of Barbados. *Journal of Travel & Tourism Marketing*, 25(3-4), 398-408.
- Juwattanasamran, P., Supatranuwong, S., & Sinthupinyo, S. (2013). Applying data mining to analyze travel pattern in searching travel destination choices. *International Journal of Engineering Science*, 2(4), 38-44.
- Kara, N. S., & Mkwizu, K. H. (2020). Demographic factors and travel motivation among leisure tourists in Tanzania. *International Hospitality Review*. 34(1), 81-103.
- Kim, J. H., & Ritchie, B. W. (2012). Motivation-based typology: An empirical study of golf tourists. *Journal of Hospitality & Tourism Research*, 36(2), 251-280.
- Kim, S. S., Kim, M., Park, J., & Guo, Y. (2008). Cave tourism: Tourists' characteristics, motivations to visit, and the segmentation of their behavior. *Asia Pacific Journal of Tourism Research*, 13(3), 299-318.
- Kozak, M., & Rimmington, M. (2000). Tourist satisfaction with Mallorca, Spain, as an off-season holiday destination. *Journal of travel research*, 38(3), 260-269.
- Kozak, M. (2002). Destination benchmarking. *Annals of tourism research*, 29(2), 497-519.
- Lee, T. H. (2007). An ecotourism behavioural model of national forest recreation areas

- in Taiwan. *International Forestry Review*, 9(3), 771-785.
- Lee, T. H. (2009). A structural model to examine how destination image, attitude, and motivation affect the future behavior of tourists. *Leisure sciences*, 31(3), 215-236.
- Lee, C. F. (2015). Tourist satisfaction with forest recreation experience: a segment-based approach. *Anatolia*, 26(4), 535-548.
- Liao, S. H., Chen, Y. J., & Deng, M. Y. (2010). Mining customer knowledge for tourism new product development and customer relationship management. *Expert Systems with Applications*, 37(6), 4212-4223.
- Lin, C. T., & Huang, Y. L. (2009). Mining tourist imagery to construct destination image position model. *Expert Systems with Applications*, 36(2), 2513-2524.
- Luo, Y., & Deng, J. (2008). The New Environmental Paradigm and nature-based tourism motivation. *Journal of Travel research*, 46(4), 392-402.
- March, R. S., & Woodside, A. G. (2005). *Tourism behaviour: travellers' decisions and actions*. Cabi Publishing.
- Mayo, E. J., & Jarvis, L. P. (1981). *The psychology of leisure travel. Effective marketing and selling of travel services*. CBI Publishing Company, Inc.
- Meng, F., Tepanon, Y., & Uysal, M. (2008). Measuring tourist satisfaction by attribute and motivation: The case of a nature-based resort. *Journal of vacation marketing*, 14(1), 41-56.
- Oliver, R. L. (1980). A cognitive model of the antecedents and consequences of satisfaction decisions. *Journal of marketing research*, 17(4), 460-469.
- Ozdemir, B., Aksu, A., Ehtiyar, R., Çizel, B., Çizel, R. B., & İçigen, E. T. (2012). Relationships among tourist profile, satisfaction and destination loyalty: Examining empirical evidences in Antalya region of Turkey. *Journal of Hospitality Marketing & Management*, 21(5), 506-540.
- Pearce, P. L. (1988). *The ulysses factor: Evaluating visitors in tourist settings*, New York: Springer-Verlag.
- Pearce, P. L. (2005). *Tourist behaviour: Themes and conceptual schemes*. Channel View Publications.
- Pearce, P. L. (2011). Travel motivation, benefits and constraints to destinations. In Y. Wang and A. Pizam (Eds.), *Destination marketing and management: Theories and applications*, CABI, 39-52.
- Pearce, P. L., & Lee, U. I. (2005). Developing the travel career approach to tourist motivation. *Journal of travel research*, 43(3), 226-237.
- Prayag, G. (2009). Tourists' evaluations of destination image, satisfaction, and future behavioral intentions-the case of Mauritius. *Journal of Travel & Tourism Marketing*, 26(8), 836-853.
- Pyo, S., Uysal, M., & Chang, H. (2002). Knowledge discovery in database for tourist destinations. *Journal of Travel Research*, 40(4), 374-384.
- Ranjan, J., & Bhatnagar, V. (2010). Application of data mining techniques in the financial sector for profitable customer relationship management. *International Journal of Information and Communication Technology*, 2(4), 342-354.
- Ritchie, B. W., Tkaczynski, A., & Faulks, P. (2010). Understanding the motivation and travel behavior of cycle tourists using involvement profiles. *Journal of Travel & Tourism Marketing*, 27(4), 409-425.
- Sedehi, M., Mehrabi, Y., & Khodabakhshi, A. (2011). Using principal component analysis to increase accuracy of prediction of metabolic syndrome in artificial neural network and logistic regression models. *Journal of Shahrekord University of Medical Sciences*, 13(4), 18-27.

- Shahrabi, J., & Neyestani, R. S. (2009). Discovering iranians' shopping culture by considering virtual items using data mining techniques. *Journal of Applied Sciences*, 9(13), 2351-2361.
- Shapoval, V., Wang, M. C., Hara, T., & Shioya, H. (2018). Data mining in tourism data analysis: inbound visitors to Japan. *Journal of Travel Research*, 57(3), 310-323.
- Sheldon, P. (1994). Information technology and computer systems, In S. Witt and L. Moutinho (Eds.), *Tourism marketing and management handbook*, 126-130. London: Prentice Hall.
- Sirisack, D., Xayavong, S., & Vongsanga, S. P. N. (2014). The characteristics and motivations of foreign tourists who visit Luang Prabang Province, Lao PDR. *International Journal of Business and Social Science*, 5(9), 262-275.
- Sung, Y. K., Chang, K. C., & Sung, Y. F. (2016). Market segmentation of international tourists based on motivation to travel: A case study of Taiwan. *Asia Pacific Journal of Tourism Research*, 21(8), 862-882.
- Suttikun, C., Chang, H. J., Acho, C. S., Ubi, M., Bicksler, H., Komolsevin, R., & Chongsithiphol, S. (2018). Sociodemographic and travel characteristics affecting the purpose of selecting Bangkok as a tourist destination. *Tourism and Hospitality Research*, 18(2), 152-162.
- Tsai, C. F., & Lu, Y. H. (2010). Data mining techniques in customer churn prediction. *Recent Patents on Computer Science*, 3(1), 28-32.
- Tsaur, S. H., Yen, C. H., & Chen, C. L. (2010). Independent tourist knowledge and skills. *Annals of Tourism Research*, 37(4), 1035-1054.
- Tsiotsou, R., & Vasioti, E. (2006). Using demographics and leisure activities to predict satisfaction with tourism services in Greece. *Journal of Hospitality & Leisure Marketing*, 14(2), 69-82.
- Uysal, M., Li, X., & Sirakaya-Turk, E. (2008). Push-pull dynamics in travel decisions. In H. Oh & A. Pizam (Eds.), *Handbook of Hospitality Marketing Management*, Elsevier, 412-439.
- Watfa, M. K., & Sobh, D. (2017). Generative smart tourism, the road for big data. *Journal of Advanced Management Science*, 5(6), 424-439.
- Wheeler, M. (1995). Tourism marketing ethics: an introduction. *International Marketing Review*, 12(4), 38-49.
- Wong, J. Y., Chen, H. J., Chung, P. H., & Kao, N. C. (2006). Identifying valuable travelers and their next foreign destination by the application of data mining techniques. *Asia Pacific Journal of Tourism Research*, 11(4), 355-373.
- Yan, H. B., & Ma, T. (2015). A group decision-making approach to uncertain quality function deployment based on fuzzy preference relation and fuzzy majority. *European Journal of Operational Research*, 241(3), 815-829.
- Yoon, Y., & Uysal, M. (2005). An examination of the effects of motivation and satisfaction on destination loyalty: a structural model. *Tourism management*, 26(1), 45-56.
- You, Z., Si, Y. W., Zhang, D., Zeng, X., Leung, S. C., & Li, T. (2015). A decision-making framework for precision marketing. *Expert Systems with Applications*, 42(7), 3357-3367.
- Yuan, J. J., Cai, L. A., Morrison, A. M., & Linton, S. (2005). An analysis of wine festival attendees' motivations: A synergy of wine, travel and special events?. *Journal of Vacation Marketing*, 11(1), 41-58.
- Zabkar, V., Brencic, M. M., & Dmitrovic, T. (2010). Modelling perceived quality, visitor satisfaction and behavioural intentions at the destination level. *Tourism*

management, 31(4), 537-546.

Zhang, D., Jiang, Q., & Li, X. (2005, May). A hybrid mining model based on neural network and kernel smoothing technique. *In International Conference on Computational Science* (801-805). Springer, Berlin, Heidelberg.