

利用類神經網路於酒後駕車交通事故之研究

林俊伸¹

郭廣洋²

吳信宏³

摘要

依據行政院衛生署民國 98 年死因統計資料，交通事故為國內十大死因第六位。影響交通事故的因素很多，不同的事故因素會造成不同的傷亡程度，其中因酒醉駕駛或酒駕肇事過失致死的悲劇更是時有所聞，所以透過資料的顯現讓駕駛人有所警惕，減少造成事故發生的不良因素，有其必要性。因此本研究主要在探討因為酒醉肇事而產生事故受傷程度受到哪些因素影響，利用資料探勘技術挖掘重要資訊。過去在交通事故領域進行資料探勘，常因受限於資料取得不易，而容易造成預測失準的情況。本研究為求謹慎與準確，採用行政院警政署民國 94 年至 96 年全國交通事故資料進行分析，目標變數為受傷程度，投入文獻收集而來的 13 個因素，利用類神經網路建立預測模型。本研究再投入自行整理的 25 項因素，利用類神經網路歸納出的 21 個重要因素進行維度縮減動作，最後以 9 項因素為類神經網路維度縮減的預測模型。經由因素降維的動作，研究結果也證實具有一定的效益。透過文獻回顧整理出的 13 個變數當成投入變數時，類神經網路的預測正確率為 66.24%；投入自行選取的 25 項因素時，預測正確率為 77.55%；透過因素降維的動作之預測正確率為 75.34%，此結果顯示利用較少的維度可以有效的代表原本的因素，可使資料探勘過程更具效率，並且找到一些原本未發現的重要影響因素。

關鍵詞：資料探勘、類神經網路、維度縮減、酒後駕車。

一、前言

依據行政院交通部民國 99 年 10 月統計月報統計顯示，台閩地區機動車輛登計數已超過 2166 萬輛，在道路空間極為有限的情況下，隨之而來的是交通事故日益嚴重（中華民國交通部統計處，2010）。根據行政院衛生署 98 年死因統計資料，民國 98 年十大死因中，意外事故位居第六大死亡原因，其中因交通運輸事故而死亡的人數多達 7,358 人，顯見交通事故已成為危害我們日常生活安全一大隱憂（行政院衛生署，2011）。因此如何有效運用現有資源，以達到降低車禍傷亡，促進交通安全之目的，成為當前重大課題。

¹ 國立彰化師範大學行銷與流通管理研究所畢業生。

² 中央警察大學交通學系副教授。

³ 國立彰化師範大學企業管理學系教授（聯絡地址：彰化市師大路 2 號，電話：04-7232105 轉 7412，E-mail: hhwu@cc.ncue.edu.tw）。

事故發生肇因的不同往往會造成不同程度的傷亡，其中因酒後駕車而釀成的悲劇更是屢見不鮮，酒後駕車不僅會造成自己傷亡的風險，更會影響他人的生命財產安全，研究更指出飲酒者一旦發生交通事故，其損害為未飲酒者之 3.6 倍(吳秀英，1988)。雖然相關單位近年來積極宣導酒後駕車的危險性，但酒後駕車肇事案件仍不斷發生，根據行政院內政部警政署 100 年第 12 號警政統計通報(行政院內政部警政署，2011)，民國 99 年交通事故的主要肇因仍以酒後駕車 299 件(占 20.22%)最多，因此透過現有資料來探討酒後駕車所造成的事故傷害有其必要性，可作為警示駕駛人避免不當的駕駛行為，以達到減少事故發生機會與事故傷害。

近年來隨著資料探勘(Data Mining)技術的日益精進，資料探勘技術已經被廣為利用於各領域，包括將資料探勘技術應用在車禍事故的領域上。例如李燕鳴、白璐(2000)以花蓮地區機動車事故急診傷者為研究對象探討飲酒相關交通事故之損失，調查飲酒相關事故之醫療費用及傷者的死亡率，該研究發現血液中酒精濃度(Blood Alcohol Concentration)呈陽性的傷者之死亡率會比全體傷者高出約 1.3 倍。吳冠宏等人(2006)則透過資料探勘技術中不同的分群方法，將大量車禍資料進行分群工作，發現在快車道數目較多段所發生的車禍事故較為嚴重，以供工務局及駕駛人作為決策參考及警惕之用。

本研究主要探討影響酒後駕車交通事故傷亡嚴重程度之因素，由於資料探勘技術運用在交通事故上具有相當不錯的效果，最常使用的資料探勘技術有類神經網路、決策樹、關聯法則等技術，此類的分析工具主要是能處理大量的資料，而且無須太專業的統計背景即可分析的優點(王秀雯，2004)。相較於使用傳統的統計工具，類神經網路有較強大的能力在於預測的功能上(Su, 2010)，本研究將藉由交通事故當下環境因素以及肇事者本身情況，透過類神經網路來分析大量的事故資料，探討酒後肇事產生的受傷程度。從交通事故資料中找出一些特徵，提供相關單位或決策者以減少車禍發生。

在資料探勘過程中，當資料變數的數目過多，其中若是彼此間存在高度相關性，則會使形成的模式對於應變數的估算中，反映的訊息有所重疊，亦會增加計算上的複雜及難度(劉雪芬等，2010；粘雅琪等，2010；Wei et al., 2011)。本研究將透過維度縮減(Dimension Reduction)觀念，辨識出資料集中之相關屬性並剔除無相關或是重複的屬性，以提升類神經網路預測績效。

二、文獻回顧與探討

2.1 資料探勘的定義

資料探勘是從一個龐大的資料庫中將以前未發現卻非常重要的資訊加以抽離出來，並且利用這些資訊做出重要決策的過程(陳文華，2000；粘雅琪等，2010；李明蕊，2011)。Han and Kamber (2006)提出資料探勘即是在資料庫中，利用各種分析方法與技術，將過去所累積的大量歷史資料，進行分析、歸納、預測與整合，尋找出有價值的隱藏事件，加以分析，以萃取出有用的資訊或找出使用者有興趣的樣式與知識，提供企業管理階層，作為訂定決策的依據。

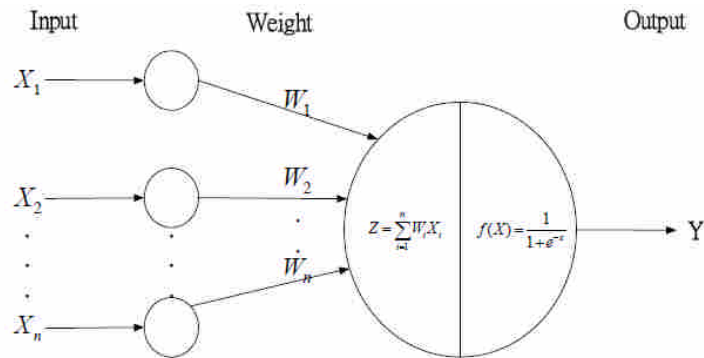
Berson et al. (2000)則認為資料探勘是從儲存著大量資料的倉儲中進行挖掘，以發現資料間有意義的新關係、型樣和趨勢的過程。利用自動或半自動學習等方式進行分析，以尋得具有意義之關係或法則(Berry and Linoff, 1997)，是一種新的且不斷循環的決策過程，它能夠從組合在一起的資料中，發現出隱藏價值的知識(Kleissner, 1998; 李明蕊，2011)。

Fayyad et al. (1996)則認為資料探勘是資料庫知識發現(Knowledge Discovery in Database)的一部分，如圖 1 所示。而知識發現的運作流程包含資料選取、資料前處理、資料轉換、資料探勘、解釋評估等階段。資料選取是從資料庫中選取分析的資料，再整合為目標資料；資料前處理是由目標資料中，清除不需要和不一致性的資料，對多餘的、遺失的、錯誤的、無關係之資料做刪除或修正處理；資料轉換是經轉換或合併成適合探勘的格式；資料探勘是運用演算法來挖掘並取得資料樣式與解釋評估是經評估或辯認資料樣式是否令人感到興趣，將經過評估而有意義之資料樣式，依視覺化或其他技術將知識呈現(劉宜妝，2002)。

2.2 類神經網路

類神經網路是由生物學所發展出來，主要是模仿人腦會藉由外界的刺激調整本身的行為這項特點，希望人工神經網路的組成單元也能夠像生物的神經元一樣，可從過去的經驗中學習而得到一個新的結果(顏瑋廷，2001)。類神經網路利用大量簡單的相連人工神經元來模仿生物神經網路的能力，人工神經元是生物神經元的簡單模擬，從外界或從其他神經元取得資訊並加以非常簡單的運算，而後以扇狀輸出結果給其他人工神經元，以此機制來處理資訊，並對外部的輸入以網路動態來回應(嚴仕偉，2005)。這種模式是由許多

的處理單元所組成，處理單元又被稱之為神經細胞，每個神經細胞間有不同強度的權重相互連結(吳振晃，2003)，如圖 1 所示。圖 1 為模擬生物神經細胞的神經元構造， $X_1 \dots X_n$ 代表輸入值，依輸入值的重要性給予不同的權重 (Weight)，即圖中 W_n 所表示的連接鍵權數，並加總後再經由人工神經元中的數學函數轉換， Y 代表神經元的最後輸出。



資料來源：吳振晃(2003)

圖 1 神經元構造

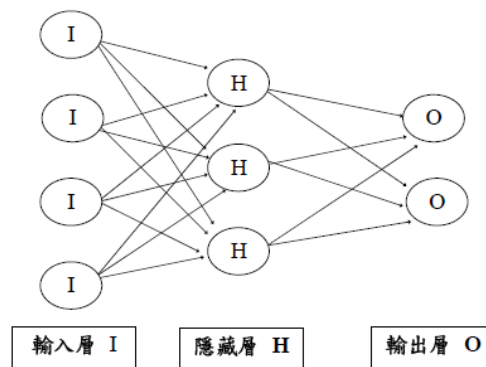
類神經網路可藉助本身的學習力、訓練力，由輸入的資料中歸納出隱含關係，一旦學習完成之後，以各神經元間連結強弱程度來表示，並且將此關係儲存於各人工神經元中。如果再碰到相似的輸入，就會自動依照所學習成的關係，來產生最近似的輸出，就算是資料不完整或是混亂時，類神經網路也會作出最佳的預測(甘堯昌，2006)。

楊金聲(2005)指出傳統人工智慧思考是根源於人們所建立的資料庫與規則，而這種法則基礎是無法適應善變的環境，更無法處理輸入不明確的訊號。就因人腦有容錯、學習及處理模糊性問題等優點，模仿生物神經網路之類神經網路繼承了這些優點，類神經網路之所以吸引各領域的矚目，主要是在於下列特性：1.學習能力：能修正本身行為以反映出環境的變化，更能依輸入自動調整，以產生正確的輸出。類神經網路是藉由輸入的圖樣，依其網路架構，自動彙整出規則來，而非經由人類輸入的規則。2.儲存能力：傳統人工智慧的儲存方式是將資料庫及規則全部儲存於電腦中，而類神經網路是將資料庫及規則分散於連結各神經元的權重上，大量減少資料儲存空間並降低資料存取時間。3.容錯性：傳統人工智慧若接受部分錯誤的輸入或雜訊干擾，將導致輸出錯誤。而類神經網路在訓練後，就算遭遇上述問題仍可得到合理輸出。4.歸納能力：能依本身的網路架構，將以前不曾見過或不完整的輸入加以歸納分類，並不一定需要明確的輸出作參考。5.平行處理：類神經網路本質上即是屬於平行處理的架構，處理速度更快速。

依網路學習策略而言，可以分為以下四類(顏能通，2005)：1.監督式學習網路：所取得的訓練範例當中，包含輸入值以及輸出值，在學習的過程當

中，輸入值經由網路的運算，會計算出一組推論值，將此推論值與範例的輸出值比較，來調整網路內的權重值，如此不斷地反覆計算，以達到推論值與輸出值之差距達到最小，倒傳遞類神經網路即屬於此類網路。2.非監督式學習網路：所取得的訓練範例當中，只有輸入值，從學習的範例去尋找範例內在聚類規則，自行找出輸入資料的規則或趨勢。3.聯想式學習網路：所取得的訓練範例當中，學習範例的內在記憶規則，以應用於新的輸入樣本中。4.最適化應用網路：對一模擬問題所設計的變數值，不僅使其滿足設計限制條件，並且使得我們所定義的目標達到最佳狀態的設計目標。

倒傳遞類神經網路(Back-Propagation Neural Network)在 1985 年由 Rumelhart 及 McClelland 所提出的，這種網路具有學習與平行處理訊息的能力(朱記民，1999)，屬於監督式學習網路，是目前應用最為廣泛且最成功的一種類神經元網路(胡明強，2001；Han and Kamber, 2006)。一般常用的倒傳遞類神經網路架構如圖 2，倒傳遞網路具有一層輸入層，一層輸出層，以及 1 到無限多的隱藏層神經元上，且每一層皆由一些神經元建構而成。在圖中同一層中的神經元彼此並不相連，不同層間的神經元則彼此相連，且信號的流向是由輸入層向輸出層單向傳播(陳艷星，2007)。



資料來源：甘堯昌(2006)。

圖 2 倒傳遞類神經網路架構

倒傳遞類神經架構共分成三個主要部分(簡辰丞，2001)，分別為：1.輸入層：用以表示網路的輸入變數，其處理單元(Processing Element)數目依問題而定，使用線性轉換函數。2.隱藏層：用以表現輸入處理單元間交互影響，其處理單元數目並無標準方法可以決定，經常需以試驗方式決定其最佳數目。使用非線性轉換函數。理論上隱藏層不無限制，通常為一或二層，有些學者則認為要解決複雜之問題則應要三至四層才行。3.輸出層：用以表現網路的輸出變數，其處理單元數目依問題而定。

倒傳遞演算法可歸納為三個重要的步驟(Han and Kamber, 2006)：1.資料輸入網路後，網路使用隨機產生的初始權重去計算輸出值；2.網路接著計算

輸出值與實際值的誤差與 3. 誤差由輸出層反向傳回網路中，藉以調整權重，以減少誤差。

2.3 酒後駕車

酒後駕車乃指喝酒後駕駛動力交通工具，不論體內酒精含量多寡，在體內酒精尚未完全消除的狀態下駕駛者，皆屬之(陳明志，2008)。然而，血液中酒精濃度更與交通事故之間有絕對的關係(汪明生等，2005)。喝酒後駕駛績效降低，駕駛者維持車道之能力降低導致車道偏移量增大、方向盤偏角度增大，而且反應時間也會變長(Lenne et al., 1999)。因為酒精會影響中樞神經系統，而中樞神經系統又與人的集中、協調、思考、記憶與決策判斷有關，隨著血液中酒精濃度的增加，會使人的身體功能減損，語言變得模糊(林大煜等，1992)。血液酒精濃度對駕駛能力之影響如表 1 所示。研究亦發現喝酒後駕駛者的視覺能力會變差，正常情況下的清醒的駕駛者週邊視界可以達到 180 度，喝完酒後的視覺角度將會減少，酒類飲料飲用量愈多對周邊的景物愈看不清楚，對光的適應亦會變差，使駕駛者無法正確操控車輛，駕駛者的反射神經也會變得遲鈍導致動作變慢(蔡中志，2000)。

表 1 血液酒精濃度對駕駛能力之影響

血液中酒精濃度值(%)	對駕駛績效的影響
< 0.03	無明顯影響
0.03 ~ 0.05	視覺與反應靈敏度減弱，對速度及距離判斷力較差
0.05 ~ 0.08	反應遲鈍、駕駛能力受損，對事情之判斷開始猶豫不決
0.08 ~ 0.15	判斷力嚴重受損、體能與精神協調力受損
0.15 ~ 0.50	精神恍惚，判斷及理解與事實不符，駕駛不穩定
> 0.50	無法開車

資料來源：謝益銘(1998)。

國內外相關文獻多以法令政策制定，探討酒後駕車行為直接管制效果或間接的影響，張新立、葉純志(2003)即以台北市之酒後駕車防治措施進行成效的監控與評估，檢討我國改善酒後駕車肇事措施介入時效之適宜性。Kenkel (1993)則以探討有效率抑制民眾酒醉駕車政策為題，發現提高法定飲酒年齡政策、採用課稅或提高酒品價格政策，皆能藉由酒類消費量的減少，達到抑制酒醉駕車的效果。

Schechtman et al. (1999)提出飲酒習慣與酒醉駕車繫安全帶、以及遵守速限規定等駕駛行為間的相關性。最後發現隨著飲酒頻率增加，酒後駕車的機率也會增加；每次飲酒數量越高，酒後駕車的機率也越高，但使用安全帶的機率卻越低。因此研究認為防治政策雖然可以遏止酒後駕車的行為，但是對於飲酒頻率較高以及每次飲酒數量較多的駕駛者而言卻是無效的。

三、研究方法

本研究資料來自於警政署 A1 類事故資料(A1 類係指造成人員當場或二十四小時內死亡之交通事故)，此資料為車禍發生時由警察機關人員依據車禍當下情況記錄，具有民國 94 年至 96 年度 A1 類的車禍現場與肇事人員資料，資料筆數為 20,824 筆，從中抽取出經警察機關酒測有酒測值反應之肇事記錄，資料筆數總共為 3,095 筆，每一筆肇事記錄皆包含了環境因素、人為因素及車輛因素三大類。樣本資料選取過程如表 2 所示，X 代表該筆資料未選取，O 則表示選取該筆資料，其中資料類別號碼直接對應至表 2 欄位項，資料並未經過轉換過程。

表 2 選取事故樣本資料

性別	年齡	受傷程度	保護裝備	職業	飲酒情形	選取
1	54	1	1	22	未飲酒	X
1	26	1	1	22	檢測超過0.25%	○
1	36	1	1	22	檢測超過0.41%	○
2	29	1	2	22	檢測超過0.55%	○

本研究係以類神經網路，將民國 94 年至 96 年度之車禍肇事記錄，分為訓練與測試資料。在資料探勘過程中，訓練及測試樣本的比例大小可以從 50% 訓練資料、50% 測試資料，到 90% 訓練資料、10% 測試資料(Larose, 2006)。由於本研究事故肇事筆數僅 3,095 筆，因此將訓練樣本設置較多的筆數，設定 80% 為訓練資料、20% 為測試資料，建立訓練資料與測試資料，以進行預測模型之建立。由於本研究主要是分析影響事故受傷程度，因此就受傷程度之類別項目進行分析，共分成死亡、受傷及未受傷，共三個程度指標。

本研究使用 SPSS Clementine 12.0 當中的類神經網路模型建立，在類神經網路模型節點參數的設定中，採用徹底修剪方法(Exhaustive Prune)，以一個大型類神經網路開始，隨著訓練的進行剔除隱藏層和輸入層的最差單元，由於是徹底修剪，類神經網路訓練參數的選擇要確保對可能模型空間進行徹底搜索以選出最好的模型，訓練時間長，但能生成最好的結果。其餘設定亂數種子(Set Random Speed)設為 7，終止準則(Stop On)設為 Default，在此設置下，網路將在其看起來達到最佳訓練狀態時停止訓練。

首先，本研究透過文獻所整理影響車禍肇事之影響因素，包含性別、駕駛資格、天候、事故位置等 13 項因素，建立受傷程度之預測模型。其次，除了文獻中所提及的因素外，其他的因素亦有可能影響酒駕肇事受傷程度，因此本研究為求準確，自行挑選包含天候、光線、道路類別、速限、道路型態、事故位置、路面鋪裝、路面狀態、路面缺陷、障礙物、視距、號誌種類、號誌動作、車道劃分設施、事故類型及型態、性別、年齡、保護裝備、行動電話、當事者行動狀態、駕駛資格情形、最初撞擊部位、其他撞擊部位、職業與旅次目的，共計 25 項因素，建立受傷程度預測模型。最後，本研究針對前述 25 項因素採用維度縮減觀念，再進行酒駕肇事受傷程度之預測模型。所有因素欄位項及說明整理於表 3。

表 3 輸入因素各欄位項說明

因素	說明
性別	(1)男、(2)女
年齡	肇事者年齡
天候	(1)暴風、(2)強風、(3)風沙、(4)霧或煙、(5)雪、(6)雨、(7)陰、(8)晴
光線	(1)日間自然光線、(2)晨或暮光、(3)夜間(或隧道、地下道、涵洞)有照明、(4)夜間(或隧道、地下道、涵洞)無照明
道路類別	(1)國道、(2)省道、(3)縣道、(4)鄉道、(5)市區道路、(6)村里道路、(7)專用道路、(8)其他
速限	發生車禍肇事地點之速限
道路型態	(1)平交道有遙控器、(2)平交道無遙控器、(3)三岔路、(4)四岔路、(5)多岔路、(6)隧道、(7)地下道、(8)橋樑、(9)涵洞、(10)高架道路、(11)彎曲路及附近、(12)坡路、(13)巷弄、(14)直路、(15)其他、(16)圓環、(17)廣場
事故位置	(1)交岔路口內、(2)交岔口附近、(3)機車待轉區、(4)機車停等區、(5)交通島(含槽化線)、(6)迴轉道、(7)快車道、(8)慢車道、(9)一般車道(未劃入快慢車道)、(10)公車專用道、(11)機車專用道、(12)機車優先道、(13)路肩和路緣、(14)加速車道、(15)減速車道、(16)匝道、(17)行人穿越道、(18)穿越道附近、(19)人行道、(20)收費站附近、(21)其他
路面鋪裝	(1)柏油、(2)水泥、(3)碎石、(4)其他鋪裝、(5)無鋪裝
路面狀態	(1)冰雪、(2)油滑、(3)泥濘、(4)濕潤、(5)乾燥
路面缺陷	(1)路面鬆軟、(2)突出(高低)不平、(3)有坑洞、(4)無缺陷
障礙物	(1)道路工事(程)中、(2)有堆積物、(3)路上有停車、(4)其他障礙物、(5)無障礙物
視距	(1)彎道、(2)坡道、(3)建築物、(4)樹木、農作物、(5)路上停放車輛、(6)其他、(7)良好
號誌種類	(1)行車管制號誌、(2)行管制號誌(附設行人專用號誌)、(3)閃光號誌、(4)無號誌
號誌動作	(1)正常、(2)不正常、(3)無動作、(4)無號誌

表 3(續) 輸入因素各欄位項說明

車道劃分設施	(一)中央分向島：(1)寬式(50公分以上)、(2)窄式附柵欄、(3)窄式無柵欄 (二)雙向禁止超車線：(4)附標記、(5)無標記 (三)單向禁止超車線：(6)附標記、(7)無標記 (四)行車分線線：(8)附標記、(9)無標記 (五)無：(10)無分向設施
事故類型及型態	(一)人與汽(機)車：(1)對向通行中、(2)同向通行中、(3)穿越道路中、(4)在路上嬉戲、(5)在路上作業中、(6)衝進路中、(7)從停車後(或中)穿出、(8)佇立路邊、(9)其他 (二)車與車：(10)對撞、(11)對向擦撞、(12)同向擦撞、(13)追撞、(14)倒車撞、(15)路口交岔撞、(16)側撞、(17)其他 (三)汽(機)車本身：(18)路上翻車摔倒、(19)衝出路外、(20)撞護欄、(21)撞號誌桿、(22)撞收費亭、(23)撞交通島、(24)撞非固定設施、(25)撞橋樑建築物、(26)撞路樹電桿、(27)撞動物、(28)撞工程施工、(29)其他 (四)平交道事故：(30)衝過遮斷器、(31)正越過平交道中、(32)暫停位置不當、(33)在平交道內無法行動、(34)其他
保護裝備	(1)戴安全帽或繫安全帶、(2)未戴安全帽或未繫安全帶、(3)不明、(4)其他
行動電話	(1)未使用、(2)使用手持、(3)使用免持、(4)不明、(5)非汽(機)車駕駛人
駕駛資格情形	(1)有適當之駕照、(2)無照(未達考照年齡)、(3)無照(已達考照年齡)、(4)越級駕駛、(5)駕照被吊扣、(6)駕照被吊銷、(7)不明、(8)非汽機車駕駛人
當事者行動狀態	(一)車的狀態：(1)起步、(2)倒車、(3)停車操作中、(4)超車(含超越)、(5)左轉彎、(6)右轉彎、(7)向左變換車道、(8)向右變換車道(9)向前直行中、(10)插入行列、(11)迴轉或橫越道路中、(12)靜止(引擎熄火)、(13)停等(引擎未熄火)、(15)其他 (二)人的狀態：(16)步行、(17)靜立(止)、(18)奔跑、(19)上、下車、(20)其他 (三)不明：(21)不明
車輛撞擊部位	(一)汽車：(1)前車頭、(2)右側車身、(3)後車尾、(4)左側車身、(5)右前車頭(身)、(6)右後車尾(身)、(7)左後車尾(身)、(8)左前車頭(身)、(9)車頂、(10)車底 (二)機車：(11)前車頭、(12)右側車身、(13)後車尾、(14)左側車身 (三)其他：(15)不明、(16)非汽(機)車
職業	(1)民意代表、行政主管、企業主管及經理人員、(2)專業人員、(3)技術員及助理人員、(4)事務工作者、(5)服務工作者、(6)售貨員、(7)農林漁牧工作者、(8)保安工作者(不含警察人員)、(9)技術工、(10)汽車、火車駕駛員及船員、(11)機械設備操作工及組裝工、(12)非技術工及體力工、(13)未就學兒童、(14)小學生、(15)國中生、(16)高中生、(17)專科生、(18)大學(研究生)、(19)家庭主婦(夫)、(20)無業者、(21)其他、(22)不明、(23)警察人員
旅次目的	(1)上、下班、(2)上、下學、(3)業務聯繫、(4)運輸、(5)社交活動、(6)觀光旅遊、(7)購物、(8)其他、(9)不明

表 3(續) 輸入因素各欄位項說明

車種	(一)大客車：(1)公營公車、(2)民營公車、(3)公營客運、(4)民營客運、(5)遊覽車、(6)自用大客車 (二)小客車：(7)營業用小客車、(8)自用小客車 (三)大貨車：(9)營業用大貨車、(10)自用大貨車 (四)全聯結車：(11)營業用全聯結車、(12)自用全聯結車 (五)半聯結車：(13)營業用半聯結車、(14)自用半聯結車 (六)曳引車：(15)營業用曳引車、(16)自用曳引車 (七)小貨車：(17)營業用小貨車、(18)自用小貨車 (八)特種車：(19)救護車、(20)消防車、(21)警備車、(22)工程車、(23)其他特種車 (九)軍車：(24)大客車、(25)載重車、(26)小型車 (十)機車：(27)大型重型、(28)普通重型、(29)輕型 (十一)慢車：(30)腳踏車、(31)人力車、(32)獸力車、(33)兒童專用車、(34)其他慢車 (十二)其他車：(35)併裝車、(36)農耕用車、(37)動力機械、(38)拖車(架)、(39)火車、(40)其他車 (十三)行人：(41)行人、(43)其他人
----	--

四、研究結果

4.1 透過文獻整理投入因素的類神經網路之結果

本研究針對受傷程度透過類神經網路建構模式，投入參考文獻之 13 種因素，將資料分成訓練及測試資料進行交叉驗證，類神經網路預測模式建構結果如表 4 所示，訓練及測試預測準確率為 66.24% 及 65.75%。表 5 為針對受傷程度預測模型所得出的 4 個重要因素，以駕駛資格(0.312)為最重要因素，其次為年齡(0.279)。

表 4 類神經網路 13 因素受傷程度模型準確度

	訓練樣本		測試樣本	
	正確	1,617	66.24%	476
錯誤	824	33.76%	178	34.25%
總計	2,441	100.00%	654	100.00%

表 5 類神經網路 13 因素受傷程度重要因素

	因素名稱	重要性
1	駕駛資格	0.312
2	年齡	0.279
3	道路型態	0.223
4	保護裝備	0.183

4.2 本研究自行整理投入因素的類神經網路之結果

利用自行選取的 25 項因素，透過類神經網路所建立的預測模式結果如表 6 所示，訓練及測試預測準確率為 77.55% 及 77.06%。表 7 為類神經網路針對受傷程度預測模型由 25 項因素中所得出的 22 個重要因素，以最初撞擊部位 (0.199) 為最重要因素，其次為行動狀態 (0.103)。

表 6 類神經網路 25 因素受傷程度模型準確度

	訓練樣本		測試樣本	
	正確	1,893	77.55%	504
錯誤	548	22.45%	150	22.94%
總計	2,441	100.00%	654	100.00%

表 7 類神經網路 25 因素受傷程度模型重要因素

	因素名稱	重要性
1	最初撞擊部位	0.199
2	行動狀態	0.103
3	保護裝備	0.077
4	旅次目的	0.076
5	光線	0.052
6	年齡	0.050
7	道路型態	0.046
8	分向設施	0.045
9	道路類別	0.042
10	其他撞擊部位	0.035
11	駕駛資格	0.029
12	天候	0.027
13	視距	0.025
14	障礙物	0.021
15	號誌種類	0.020
16	性別	0.020
17	路面狀態	0.020
18	路面鋪裝	0.014
19	號誌動作	0.012
20	路面缺陷	0.011
21	速限	0.003

4.3 應用維度縮減觀念於類神經網路之結果

本研究透過前述 25 項因素的預測模型，所萃取出的 21 項重要因素中，先投入前 8 項的因素後，逐步增加因素數量，分別建立受傷程度的類神經網路模型，直至預測率無明顯增加為止。如表 8 所示，模型在投入第 9 項因素可得最佳的預測率，而後的模型預測率則逐步下降，因此本研究將以 9 項因素為維度縮減之模型。投入的因素為最初撞擊部位、行動狀態、保護裝備、旅次目的、光線、年齡、道路型態、分向設施及道路類別，共計 9 項因素。針對受傷程度投入 25 項因素模型中前 9 項重要因素，透過類神經網路預測模式建構結果如表 9 所示，訓練及測試預測準確率為 75.34% 及 74.62%。表 10 為類神經網路針對受傷程度預測模型由 9 項因素中所得出的 7 個重要因素，以最初撞擊部位(0.303)為最重要因素，其次為行動狀態(0.163)。

表 8 類神經網路維度縮減測試

	訓練樣本預測率	測試樣本預測率
8 項因素	74.48%	75.23%
9 項因素	75.34%	74.62%
10 項因素	74.72%	75.23%
11 項因素	74.11%	74.77%

表 9 類神經網路 9 因素受傷程度模型準確度

	訓練樣本		測試樣本	
	正確	1,853	75.91%	506
錯誤	588	24.09%	148	22.63%
總計	2,441	100.00%	654	100.00%

表 10 類神經網路 9 因素受傷程度模型重要因素

	因素名稱	重要性
1	最初撞擊部位	0.303
2	行動狀態	0.163
3	保護裝備	0.144
4	旅次目的	0.133
5	年齡	0.088
6	道路型態	0.087
7	道路類別	0.081

4.4 研究結果比較

透過文獻找尋的 13 項因素發現，只有駕駛資格、年齡、道路型態與保護裝備為重要因素，預測正確率為 66.24%。透過自行找尋的 25 項因素時，共有 21 項重要因素包括最初撞擊部位、行動狀態、保護裝備、旅次目的、光線、年齡、道路型態、分向設施、道路類別、其他撞擊部位、駕駛資格、天候、視距、障礙物、號誌種類、性別、路面狀態、路面鋪裝、號誌動作、路面缺陷與速限，預測正確率為 77.55%。因此，投入 13 項因素所找出的重要因素皆包括在使用自行找尋的 25 項因素時所找出的 21 項重要因素當中。最後透過維度縮減觀念後，找出最初撞擊部位、行動狀態、保護裝備、旅次目的、光線、年齡、道路型態、分向設施及道路類別為重要因素，雖然預測正確率略為下降至 75.34%；但是考量的因素從原先的 21 項大幅減少至 9 項。此外，也比透過文獻找尋的 13 項所得到的預測正確率來得更高，因此透過維度縮減觀念於本研究確實帶來好處，包括找出較為重要的因素於酒駕車禍與找出文獻回顧未討論到的其它重要因素。

五、結論

藉由類神經網路模式探討酒後駕車肇事事務資料，預測正確率則介於 66.24% 至 77.55%。本研究發現在進行資料探勘技術過程中，除了透過文獻來搜尋相關影響因素之外，以本身的主觀意志來挑選可能的因素，亦可以達到不錯的效果，也可以避免文獻中漏掉關鍵的因素。不過當因素過多時，所花的時間亦會較長，因此經由維度縮減來達到因素降維的動作，研究結果也證實具有一定的效益。此結果顯示利用較少的維度可以有效的代表原本的因素，可使資料探勘過程更具效率。本研究在類神經網路的模型中，原本需花費 4.5 小時的建構時間，在維度縮減後，僅約 2 個小時就可完成。然而本研究中只使用三千多筆資料就有如此差異，當資料量越大時，所減少的時間成本將更明顯。

參考文獻

- 中華民國交通部統計處 (2010)，中華民國交通部統計月報，交通部統計處。
- 王秀雯 (2004)，應用資料挖掘技術於交通事故傷亡嚴重程度之研究，國立嘉義大學運輸與物流工程研究所碩士論文。
- 甘堯昌 (2006)，以資料採礦技術探討影響個股績效之關鍵因素-以台灣上市個股為例，大同大學資訊經營研究所碩士論文。
- 朱記民 (1999)，類神經網路技術應用在中央空調可變風量控制系統之研究，國立彰化師範大學工業教育系碩士論文。
- 行政院內政部警政署 (2011)，內政部警政署 100 年第 12 號統計通報，擷取

- 日期：2011 年 7 月 5 日，網站：
<http://www.npa.gov.tw/NPA/Gip/wSite/ct?xItem=55978&ctNode=12594&mp=1>。
- 行政院衛生署 (2011)，98 年度死因統計完整統計表，擷取日期：2011 年 7 月 5 日，網站：
http://www.doh.gov.tw/CHT2006/DM/DM2_2_p02.aspx?class_no=440&now_fod_list_no=11397&level_no=4&doc_no=76512。
- 吳秀英 (1988)，交通傷害事故危險因子的探討，國立陽明大學公共衛生研究所碩士論文。
- 吳冠宏、吳信宏、郭廣洋 (2006)，「應用分群技術於交通事故資料分析」，品質學報，第十三卷第三期，頁 305-312。
- 吳振晃 (2003)，資料採礦技術於銀行授信之應用以消費者貸款為例，中國文化大學資訊管理研究所碩士論文。
- 李明菴 (2011)，應用資料探勘技術於顧客關係管理之個案研究，國立彰化師範大學行銷與流通管理研究所碩士論文。
- 李燕鳴、白璐 (2000)，「飲酒相關交通事故之醫療花費及死亡率情況分析」，台灣家庭醫學雜誌，第十卷三期，頁 148-161。
- 汪明生、黃國良、郭文俊 (2005)，「酒後駕車風險知覺之實驗研究-資訊整合理論之應用」，管理學報，第二十二卷第四期，頁 429-447。
- 林大煜、鄭俊明、曹瑞和、袁正平 (1992)，駕駛人行為反應之研究-違規駕駛人性向測驗分析與矯正模式建立之研究，交通部運輸研究所研究報告。
- 胡明強 (2001)，應用倒傳遞類神經網路於象棋棋子之辨識，國立彰化師範大學工業教育學系碩士論文。
- 張新立、葉純志 (2003)，「酒後駕車防治措施成效之監控與評估-以台北市為例」，運輸計劃季刊，第三十二卷第一期，頁 131-150。
- 陳文華 (2000)，「運用資料倉儲技術於顧客關係管理」，能力雜誌，第五百二十七期，頁 132-138。
- 陳明志 (2008)，酒後駕車者之問題行為症候群研究-以臺北市為例，國立台北大學犯罪學研究所碩士論文。
- 陳艷星 (2007)，倒傳遞類神經網路法於烏山頭水庫壩體滲流量觀測值分析之研究，國立成功大學水利及海洋工程研究所碩士論文。
- 粘雅琪、郭廣洋、吳信宏 (2010)，「應用資料採礦技術於輕型機車交通事故之研究」，九十九年道路交通安全與執法研討會論文集，頁 159-177。
- 楊金聲 (2005)，利用類神經網路與線性迴歸進行成本預測之研究-以印刷電路板產業為例，中原大學資訊管理學系碩士論文。

- 劉宜妝 (2002), 資料採礦之應用研究-台灣地區漁市場行情資料庫之關聯法則分析, 國立中興大學行銷系碩士論文。
- 劉雪芬、郭廣洋、吳信宏 (2010), 「應用分類迴歸樹於尖峰時段交通事故之研究」, 中華民國品質學會第四十六屆年會暨 2010 國際品質管理研討會, E2-3。
- 蔡中志 (2000), 血液唾液連液吐氣酒精濃度檢測方法實用性評估, 交通部道路安全委員會報告。
- 謝益銘 (1998), 提昇酒後駕車執法品質之研究, 中央警察大學警政研究所碩士論文。
- 簡辰丞 (2001), 結合 MACD 與類神經模糊技術之股票預測模型-以金融股為例, 靜宜大學企業管理研究所碩士論文。
- 顏能通 (2005), 應用類神經網路於模鑄型變壓器部份放電音頻信號之辨識, 國立成功大學電機工程研究所碩士論文。
- 顏瑋廷 (2001), 應用類神經網路法於表皮及肩胛骨之運動分析與模式建立, 國立成功大學工業設計學系碩士論文。
- 嚴仕偉 (2005), 整合類神經網路與資料採礦於金融最佳決策代理系統之規劃與設計, 國立交通大學資訊管理研究所碩士論文。
- Berry, J. A. and Linoff, G. (1997), *Data Mining Techniques for Marketing, Sales, and Customer Support*, New York: John Wiley & Sons Inc.
- Berson, A., Smith, S., and Thearling, K. (2000), *Building Data Mining Application for CRM*, New York: McGraw-Hill Inc.
- Fayyad, U., Piatetsky, G., and Smith, P. (1996), "From Data Mining to Knowledge Discovery in Databases," *AI Magazine*, Vol. 17, No. 3, pp. 37-54.
- Han J. and Kamber, M. (2006), "Data Mining: Concepts and Techniques", San Francisco: Morgan Kaufmann.
- Kenkel, D. S. (1993), "Drinking, Driving, and Deterrence: the Effectiveness and Social Costs of Alternative Policies," *Journal of Law and Economics*, Vol. 36, No. 2, pp. 877-913.
- Kleissner, C. (1998), "Data Mining for the Enterprise," *Proceedings of the Hawaii International Conference on System Sciences*, Vol. 7, pp. 295-304.
- Larose, D. T. (2006), *Data Mining Methods and Models*, Canada: Wiley.
- Lenne', M. G., Triggs, T. J., and Redman, J. R. (1999), "Alcohol, Time of Day, and Driving Experience: Effects on Simulated Driving Performance and Subjective Mood," *Transportation Human Factors*, Vol. 1, No. 4, pp.

331-346.

- Schechtman, E., Shinar, D., and Compton, R. C. (1999), "The Relationship between Drinking Habits and Safe Driving Behaviors," *Transportation Research Part F: Traffic Psychology and Behaviour*, Vol. 2, No. 1, pp. 15-26.
- Su, C. C. (2010), "Media Consumption in Taiwan: Mining the Shin Hsin Communication Database," *Asia Pacific Management Review*, Vol. 15, No. 1, pp. 113-131.
- Wei, J.-T., Kou, K.-Y., and Wu, H.-H. (2011), "Using Feature Selection to Reduce the Complexity in Analyzing the Injury Severity of Traffic Accidents," *2011 International Joint Conference on Service Sciences*, pp. 329-333.