

第一章 緒論

本章主要內容分為五個部分，分別就研究動機、研究目的、研究範圍與限制、研究方法與研究流程、論文內容結構予以說明。

1.1 研究動機

臺北市自改制後，於民國 90 年交通事故死亡人數降為 98 人，首度創下肇事死亡人數低於百人紀錄，並於 91 年、92 年持續維持此一成果，分別將死亡人數降為 89 人及 87 人，惟自今(93)年 5 月止，肇事死亡人數為 40 人已較去年同期(92 年 1-5 月)死亡人數增加 9 人，如何增進交通安全避免肇事發生、降低傷亡人數，將是臺北市交通相關單位，達成交通政策白皮書中所承諾的目標——每年交通事故死亡人數低於 90 人以下，所必須面對的重要課題。

提升道路交通安全所需考慮的層面相當廣，涉及工程、教育、執法等數個政府部門，不但有政府部門橫向協調的困難性，且礙於政府經費有限，如何以有限的資源完成最大的效益，將是首先面臨的問題。所以除了道路規劃、開發時必須考量各種安全的因素，另外已開發完成的道路，也要針對易肇事路段及路口進行改善，其中尤其是以易肇事路段及路口的改善最符經濟效益，以目前對於易肇事路段及路口的判定，仍是最單純的肇事率、肇事次數或肇事傷亡人數來衡量，此法最主要的優點是簡單易懂，最大的缺點則是對於新闢或改善道路只能靜待肇事發生後，才可判定易肇事地點、著手改善；另外從易肇事地點改善之方式來看，其主要手段在於進行局部道路工程及交通工程的修正，但造成交通事故的因素相當複雜，主要可歸因於「人、車、路」之三大因素，而這三項因素並非獨立性，係相互影響而產生不同的肇事型態，故現行易肇事地點的改善常必須借重專家的經驗，經由會勘的方式，就各項可能發生的肇事因素進行改善，但由於專家數量稀少，養成費時，且依據以往經驗，大部分交通事故之肇事主要原因多可歸類為「人」的因素，故對警察機關而言，執法作為或可達到治標的、速效的降低肇事，然而執法強度過高又常造成民眾抱怨，且所投入過多的警力亦有浪費之虞，如可找出會導致嚴重傷亡之違規類型，並針對該類違規加強執法，應可有效達成防制肇事、減少警力浪費及提高民眾滿意度之三項目標。

隨著人工智慧的興起，各類運用人工智慧的方法去探求交通事故件數、嚴重程度及其影響因素間相互關係之研究也因應而生，其中類神經網路被運用的次數及準確度不下於其他多變量分析的統計方法或人工智慧方法。本研究欲利用類神經網路來探討臺北市地區交通肇事嚴重地點及肇事因素的判別方式，並據以研擬可應用於執法的方案。

1.2 研究目的

本研究採用類神經網路方法，構建一個類神經網路架構用以判別肇事嚴重程度。以往對於交通事故的研究多著重於肇事次數、肇事率、受傷及死亡人數與道路幾何設計、車流特性、車輛特性等因果關係，對肇事嚴重的分析多是從總體的關點來探討，透過傳統迴歸、一般線性迴歸、卜瓦松迴歸、負二項迴歸，來探討

某一研究範圍(路段或路口)的肇事發生頻率及嚴重度，但是這樣的作法無法觀察到當事人本身屬性與肇事嚴重度的關連，並進一步以個體觀點探討肇事；且這類模式無法對危險影響因子作整體、合理的考量，模式易侷限於統計假設上，由於假定各項輸入因子具有獨立性，致無法詮釋各項因子間之非線性現象，對於上述統計模式之問題癥結，類神經網路方法恰可有效解決。

構建一個適用的類神經網路肇事嚴重程度預測模式，以預測出某路口、路段在發生肇事後，當事人之傷亡程度，及早對管理者與道路使用者提出警訊，並以適當的行政作為加以因應，此為本研究主要的目的所在。

1.3 研究範圍與限制

為滿足都市地區肇事嚴重模式構建需求，本研究將在資料可獲性及研究資源限制下，基於時間及成本考量，以臺北市 92 年 1-12 月發生的交通事故為分析對象，利用肇事記錄調查表為肇事資料來源。依肇事地點分為路口及路段，分別就其肇事資料進行研究。鑑於肇事常有數個參與的當事人，而死亡、受傷人數亦有多個，本研究僅挑選最嚴重的作為輸出項。

1.4 研究方法與研究流程

鑑於道路環境與交通工程設施對於交通肇事的影響錯綜複雜，且部分因素間相關程度頗高，欲以統計方法界定單項因素對肇事嚴重度的影響程度實有困難，再者道路幾何及交通工程因素與肇事間呈現高度非線性關係，如以傳統迴歸所能達到的解釋及預測效果不甚理想，另外，肇事的違規因素及當事人特性對於肇事嚴重程度亦有相當程度之影響；類神經網路具有處理複雜模式之能力，並可解決不易釐清之非線性關係，故本研究採用此一方法來探討該類問題。

本研究之研究流程如圖 1-1，以下分別就各流程予以說明：

一、問題界定及研究動機與目的

首先界定目前道路交通事故的問題特性，以確定研究目的與內容，釐清研究的範圍及對象。

二、文獻回顧及類神經網路回顧

蒐集國內外與交通肇事預測、分析等相關之文獻資料，包括研究報告與期刊論文，加以整理分析檢討作為後續研究之基礎，另外對於本研究所欲使用的類神經網路方法，對其理論基礎、相關文獻及應用軟體等資料進行蒐集整理。

三、資料蒐集與整理

需蒐集的肇事資料包括有道路環境、交通工程、當事人特性、肇事因素及其他相關資料等，以作為後續研究之基礎。

四、資料庫建立

依所蒐集的資料建立資料庫作為類神經網路訓練及測試之用。

五、構建類神經網路模式

類神經網路訓練、測試以評選出對交通肇事具有影響之各項因素作為輸入變數，肇事數量為輸出變數，並將資料分為兩部分，一部分作為訓練之用，一部分作為測試之用。經過上述訓練、測試過程，選擇最適隱藏層層數及處理

單元數目，建立一預測結果可接受之類神經網路模式。

六、應用研究

將模式應運於道路改善流程，並選定一肇事嚴重之路口或路段研擬改善方案。

七、結論與建議

最後依據研究成果提出結論，並建議後續研究之方向。

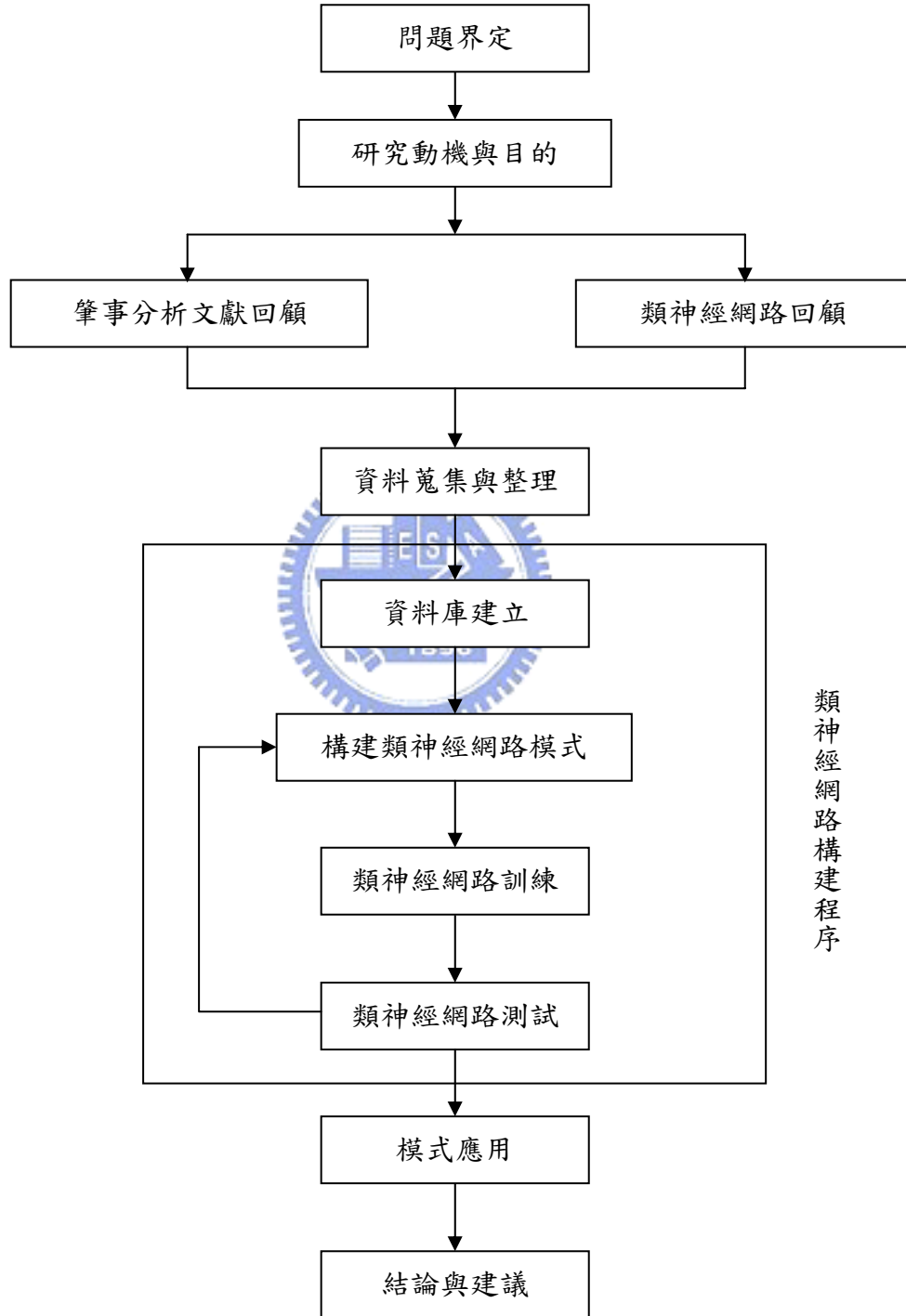


圖 1-1 研究流程圖

1.5 論文內容結構

本研究之內容架構可分為五章，第一章緒論、第二章肇事問題特性、分析方

法與相關文獻探討、第三章類神經網路概述、第四章資料整理與分析、第五章模式構建與結果分析、第六章結論與建議，以下就各章欲彙寫的內容作一簡要的說明。

- 一、緒論：主要就本研究之研究動機、研究目的、研究範圍與限制予以說明，並對所欲採用的研究方法及採用的原因予以簡述，最後則說明本研究之研究流程及論文結構。
- 二、肇事問題特性、分析方法與相關文獻探討：本章首先說明目前的肇事道路改善流程及各相關單位業務權責，其次就肇事統計資料所具有的特性加以敘述，如肇事黑數、嚴重肇事案件樣本稀少及肇事影響因素間相關性高之特性；接著簡述交通肇事常用的預測分析方法，並說明其優缺點及應用限制，最後，敘述文獻回顧心得並驗證、確認研究方向是否有偏誤及研究的必要性。
- 三、類神經網路概述：此章係簡要說明本研究所欲採用的類神經網路方法，將該方法理論基礎、運算程序、運算參數及運算過程應注意的事項予以說明，並作為後續模式構建時之根據及參考。
- 四、資料整理與分析：選取影響肇事嚴重程度之各項道路環境、交通工程設施、當事人特性及違規因素，並進行初步的統計整理，以觀察資料是否有特別或異常之處，之後再進一步將蒐集整理所得之資料建立成資料庫，以作為人工神經網路之訓練與測試之用。
- 五、模式構建與結果分析：利用整理完成之資料庫，決定輸入層、輸出層變數以及各項類神經網路訓練應設定的參數，並調整隱藏層神經元個數，進行訓練及測試，以求得最佳路口肇事嚴重程度類神路模式及路段肇事嚴重程度類神路模式，隨後將訓練所得之模式應用於肇事嚴重之道路，以尋求適當的防制方法。
- 六、結論與建議：該章主要是將研究所得以條列方式說明，並將研究過程所發現應改進及後續研究建議一併說明敘述。

第二章 肇事問題特性、分析方法與相關文獻探討

為瞭解肇事問題特性及適用的分析方法，本章就肇事問題做一整體性之文獻探討，其主要內容分為四個部分，首先就肇事嚴重道路判定及改善流程、肇事統計資料特性予以說明，其次就肇事預測分析方法及相關文獻進行探討，最後進行綜合比較分析。

2.1 肇事嚴重道路判定及改善流程現況說明

目前國內對易肇事地點之評定及改善作業係由交通部運輸研究所研訂，其改善計畫流程主要步驟如圖 2-1；臺北市各交通安全相關單位業務權責分工如表 2-1(陳高村，1997)，另外臺北市為使肇事死亡人數能有效降低，如有發生 A1 肇事案件，經初步評估可能之肇事原因後，由交通警察大隊、交通局、工務局及其他相關單位共同進行會勘，並研擬改善方案，進行肇事防制。

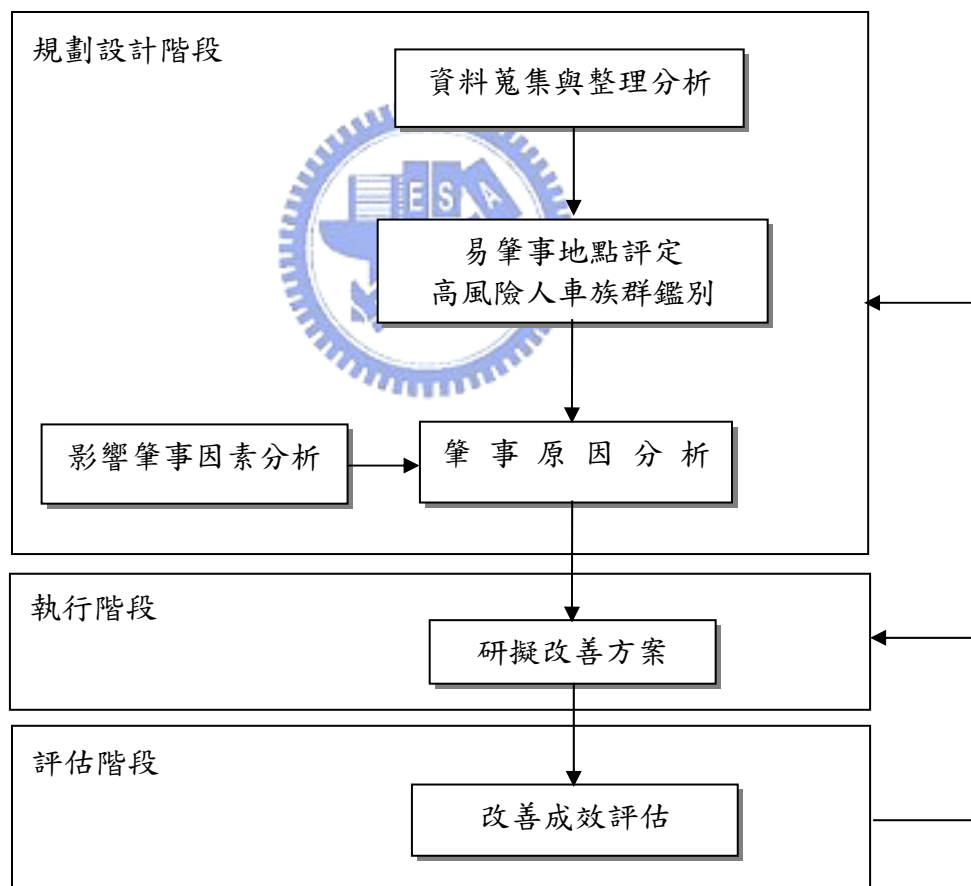


圖 2-1 交通安全改善計畫流程

表 2-1 臺北市交通安全相關單位業務權責分工表

業務項目 \ 單位名稱	交通局	警察局	道安會報	車鑑會	工務局
肇事路段改善		V		V	
肇事調查			*		
肇事鑑定				V	
道路及交通工程改善	V		V		V
車輛檢驗	V				
監理業務	V				
交通安全宣導	V	V	*		
駕駛人教育訓練	V				
運輸管理人員教育訓練	V				
交通警察專業訓練		V			
交通執法		V			
交通秩序維持		V	*		

註：*為督導單位，V為執行單位

2.2 肇事統計資料特性說明

由於肇事統計資料存在肇事黑數、樣本過度離散、樣本數不足以及各影響肇事因素之間相關性過高等特性，導致進行各項肇事預測分析時難以獲得良好的結果，以下分別就肇事統計資料特性予以說明。

2.2.1 肇事黑數

肇事案件自發生、事故調查、到最後資料庫建檔統計，這一連串的過程是否確實，有無錯漏，對於肇事資料分析應用於肇事防制均有相當大的影響，最常見問題的就是肇事黑數，目前國內事故資料登錄制度，係由警察單位負責事故現場處理，並記錄或填報事故調查報告表，再由交通主管機關進行統計分析，在實際負責處理的單位欠缺統計分析能力情況下，「肇事黑數」所造成的傷害並未及時被發現，在這樣的處理程序往往會出現實際發生事故與填報資料不符之情形，而又無從督考，陳高村(1997)在「道路交通事故處理與鑑定」乙書中曾對國外文獻進行回顧發現，世界各國對交通事故資料的統計運用也普遍存在黑數，並造成事故資料分析運用之困擾。一般而言，警察所記錄的受傷程度往往會比醫院實際的狀況輕微，如此將導致研究所得結果可能會往輕微的方向偏誤(Shankar,1996)。因此進行肇事因果預測分析時，如果肇事實際情形無法確實反映，無論使用如何先進的方法，亦無法得出正確、可用的結果。

2.2.2 肇事發生過度分散

其次導致肇事預測難以獲得正確結果的原因係肇事資料本身，一般而言，我們常將肇事案件分佈假設為 Poisson 分配，然而肇事案件發生的分佈有過度分散

情形，如此會導致肇事發生次數之變異數大於平均數，而違反 Poisson 分配平均數與變異數相等的基本假設(Miaou et al.,1992)，尤其是肇事嚴重程度高的案件數相當稀少，使得運用於統計分析時，常有樣本不足的情形 (以臺北市肇事資料為例，90 年列管交通事故共 160 件，死亡 98 人，重傷 67 人，91 年列管交通事故共 158 件，死亡 81 人，重傷 80 人，92 年列管交通事故共 142 件，死亡 87 人，重傷 56 人，3 年共計發生 460 件，分散於 423 個地點，次數最多的地點為木柵路五段，共發生 6 件，樣本資料相當離散)。故以這些肇事資料進行肇事預測分析，勢必難以獲得令人滿意的結果，陳志和 (1999)及周雍傑 (2000)在進行肇事預測研究時即得出類似的結論，也就是所構建的模式對於肇事嚴重案件的預測能力相當差，為了避免樣本過少的情形發生，通常會以長期間、大觀察區域的肇事資料進行分析，但是如此又會導致個體間之差異盡被隱藏，無法對影響行車安全因子作深入探討。

2.2.3 肇事因素相關性高

影響肇事發生的因素，約略可分成道路環境因素、交通工程因素、車輛因素、駕駛人因素等四種，各項因素間相關性非常高，尤其是交通工程因素部分，Griebe (2003)蒐集丹麥五年期間肇事資料(1990-1994 路段肇事)經過相關性分析繪製成圖 2-2，該圖顯示肇事影響因素間相關性高，因肇事資料具有此一特性，導致進行肇事預測分析時，難以構建適當模式。

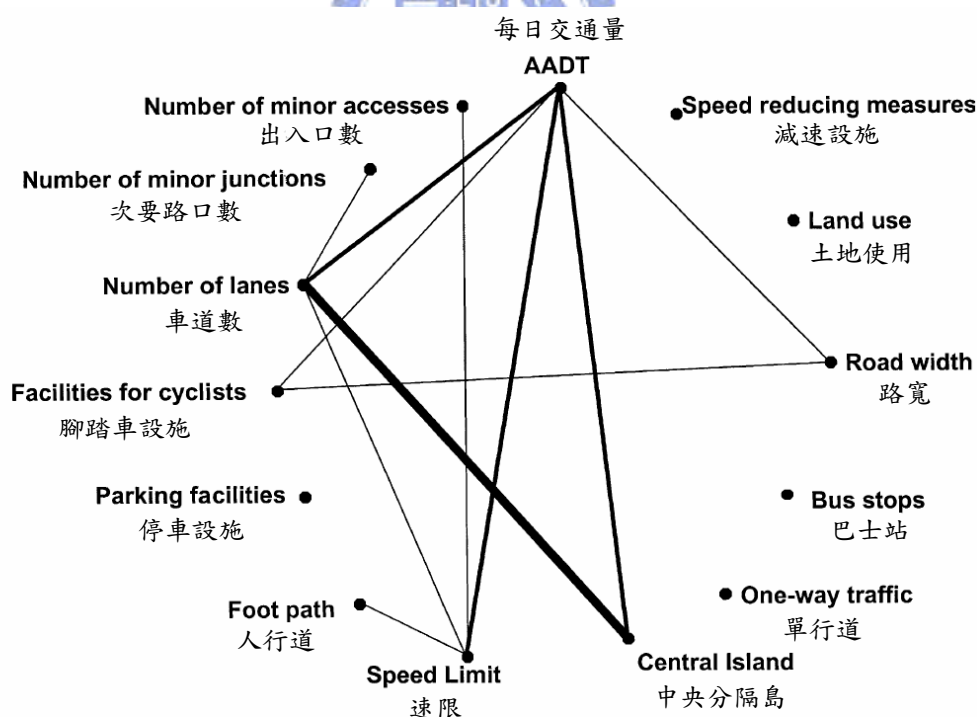


圖 2-2 路段資料之相關性表示圖

註：粗線表示相關係數 $\rho > 0.6$ ，具高度相關性，細線或無線條表示相關性較低。

2.3 肇事預測分析相關文獻探討

相關文獻中，對於各種肇事預測分析模式之構建，多數都是以交通事故的歷史資料為基礎，再利用不同的分析邏輯或數學方法進行構建，而構建的目的也或有不同，無論是易肇事路段的鑑別、危險潛因的評估或肇事率的預測，均有國內外學者進行研究，而目前較常見的肇事預測分析方法約可分類如下，茲就其模式特性歸納說明。

2.3.1 基本統計評估方法

這個方法主要是針對已發生的交通事故資料，利用基本統計分析的概念作為評估的理論基礎，因此方法簡單易懂，無需高度複雜的數學運算，故此一方法最為普及，政府部門亦用之作為決策的依據。一般常用的有肇事次數法、肇事率法、臨界肇事率法，另外亦有依上述方法，同時考量多種因素，分階段逐步進行而訂定優先篩選條件的二階段篩選法、四階段篩選法。這些方法雖然簡單明瞭，但是忽略許多與影響肇事發生及嚴重程度的重要因素，如道路幾何特性、交通量、交通管制情形。以下就各方法稍加敘述。

一、肇事次數法：係利用既有肇事資料，統計一段期間內某一路段所發生之肇事總次數，依其單位長度之事故發生數量之多寡加以排序，其公式可以 2-1 表示。

$$D_i = \frac{A_i}{L_i} \quad (2-1)$$

D_i = 路段 i 平均每單位長度肇事次數

A_i = 路段 i 肇事次數

L_i = 路段 i 長度

二、肇事率法：係以某一段期間內，將各路段肇事次數除以該路段總流量行車里程，求得每一路段之肇事率，排定危險程度大小，公式如 2-2 所示。

$$R_i = \frac{A_i}{V_i \times L_i} \times 10^6 \quad (2-2)$$

R_i = 路段 i 肇事率(百萬車公里肇事次數)

V_i = 路段 i 肇事資料取得期間車流量

三、臨界肇事率法：本法係以統計假設與檢定方法，檢定研究路段是否超過類似路段之上、下臨界肇事率，以認定是否屬危險路段，謝浩明 (1980)曾以此方法配合多變量方法進行臺北市路口、路段肇事解析及改善對策研擬，該方法公式可以 2-3、2-4 表示。其優點在於，它考慮了肇事的隨機性，並明確的訂定臨界值的設定方法，相當適用於分析交通量變化很大的公路系統，但分析過程需假設肇事之發生符合波氏分配(Poisson Distribution)。

$$R_u = \lambda + K\sqrt{\frac{\lambda}{m}} + \frac{1}{2m} \quad (2-3)$$

$$R_l = \lambda - K\sqrt{\frac{\lambda}{m}} - \frac{1}{2m} \quad (2-4)$$

- R_u = 上臨界肇事率
 R_l = 下臨界肇事率
 λ = 平均肇事率
 m = 公路行車公里數
 K = 顯著水準機率因子

四、程序性鑑別模式：程序性鑑別模式可分為二階段篩選法（程序如圖 2-3 之說明）及四階段篩選法（程序如圖 2-4 之說明）。所謂分階段，只是將肇事次數、肇事率（二階段）、事故嚴重程度、財物損失（四階段）等衡量指標依先後順序進行篩選。

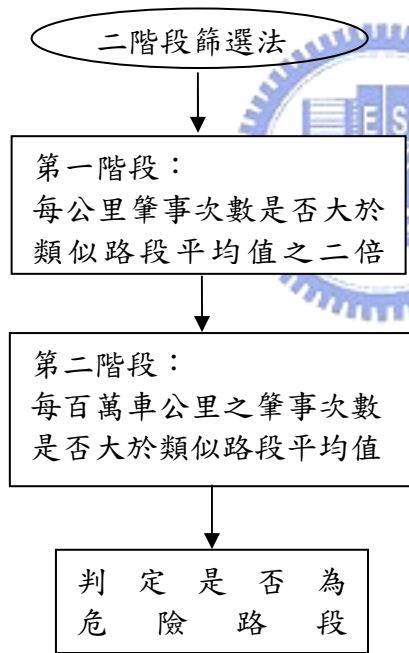


圖 2-3 二階段篩選法

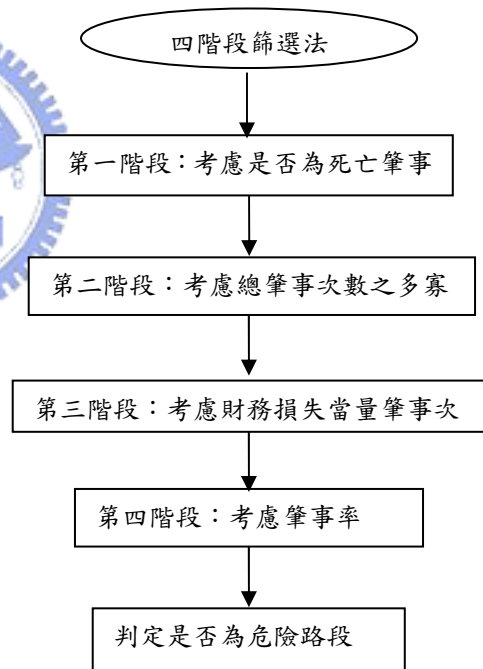


圖 2-4 四階段篩選法

2.3.2 肇事風險評估模式

肇事的發生具有隨機性，故吾人可以機率理論之概念，假設肇事次數符合 Poisson Distribution，經由統計所得之每車-公里肇事率，以機率概念建立肇事預測模式，其數學式可以公式 2-5、2-6 表示。

$$P_i = \left\{ N_i = n \mid \bar{\lambda}, V_i \right\} = \frac{(\lambda V_i)^n}{n!} e^{-\lambda V_i} \quad (2-5)$$

$$f_R(\lambda) = \frac{\beta^\alpha}{\Gamma(\alpha)} \lambda^{\alpha-1} e^{-\beta\lambda} \quad (2-6)$$

P_i = 地點 i 肇事率的機率值

$\bar{\lambda}$ = 地點 i 肇事率(隨機變數)

N_i = 地點 i 肇事次數

V_i = 通過地點 i 之車輛數

$f_R(\lambda)$ = 研究區內肇事率之機率密度函數

由於肇事的發生具有隨機性，所以每車-公里肇事率的計算概念，亦可運用於其他肇事因素(Evans, 2003)，根據所有與肇事發生有關之各項因素，如交通衝突次數、視距、坡度、駕駛操作錯誤次數、駕駛人期望值或資訊系統缺點等，藉由統計方法賦予各項因素不同的風險權數，再求其加權平均數，作為危險潛因值，其公式如 2-7 表示。

$$HI_i = \frac{\sum [W_{ij}(IV_{ij})]}{\sum W_{ij}} \quad (2-7)$$

HI_i = 地點 i 危險潛因值

IV_{ij} = 地點 i 之鑑別因子 j

W_{ij} = 地點 i 鑑別因子 j 之權重



Mountain et al. (1998) 指出依據各項因素之肇事曝光量所計算之肇事風險加權值隨著時間推進，然而國家資料卻顯示這項肇事風險評估的方法有勢微的現象。近來較常被使用的 Logit 模式，Shankar et al. (1996)及陳志和 (1999)曾運用 Logit 模式來預測肇事嚴重程度。這個方法主要是透過 Logit 轉換，將肇事發生情形(如發生或不發生，發生後是嚴重肇事或非嚴重肇事)轉換成自然對數函數式，並使用各項與肇事發生相關因素之風險權數的概念來進行肇事機率的預測。

2.3.3 統計迴歸模式

傳統線性迴歸分析可用來找出兩個或兩個以上計量變數間的關係，並進而從一群變數中預測資料趨勢，在迴歸分析中最簡單的模型是二變數的直線迴歸關係式，即所謂的簡單線性迴歸模型(Simple Linear Regression Model)，迴歸模式可以公式 2-8 表示，但是與肇事有影響之各項因素相當多，如果在模式中只引用一個自變數並不能提供正確或足夠的資訊，為充分考慮各項因素對結果的影響，大多使用多重迴歸分析(Multiple Linear Regression)進行研究，多重迴歸分析是簡單迴歸分析的延伸應用，公式如 2-9，兩者均以最小平方方法求得一個最適的線性模式，用以瞭解一組預測變項和一個目標變項的直線關係，而每個預測變項的預測能力

是研究者重要的參考指標，Miaou et al. (1992)，Miaou et al. (1993)曾使用此方法構建肇事與幾何設計間之關係模式。其他非線性迴歸模式，如多項式迴歸模式、指數迴歸模式，Greibe (2003)曾用之於肇事預測分析。

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i \quad (2-8)$$

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_{p-1} X_{ip-1} + \varepsilon_i, i = 1, 2, \dots, n \quad (2-9)$$

Y_i = 第 i 個觀察值的相依變數

X_i = 第 i 個觀察值的自變數

β_i = 參數

ε_i = 隨機誤差項

至於肇事發生，吾人可簡化地將之區分為兩類，嚴重的事故及非嚴重的事故，因為肇事可以獨立的二分法方式區分，Logistic Regression 方法則正好符合這樣的需求(Al-ghamdi,2002)，該方法主要是將肇事發生情形透過 Logit 轉換自然對數函數式，並以迴歸方法進行預測分析，目前 Logistic Regression 進行肇事預測分析的發展已相當成熟。

以統計迴歸方法進行肇事預測分析，必須假定肇事發生情形的機率分配，最常見的是將肇事分配假設為 Poisson 分配之迴歸方法，稱為 Poisson Regression，該方法是以最大概似法進行參數校估，以概似比指標及 AIC 值作模式檢定，這是該模型與傳統迴歸最大不同之處。但是由於肇事發生稀少及過度分散的特性，Miaou et al. (1992)指出，如此會導致肇事發生次數之變異數大於平均數，而違反 Poisson 分配的基本假設。近來研究指出負二項分配(Negative binomial distribution)可能更符合肇事實際的分佈狀況，因為負二項分配允許較大的變異數，故 Poch et al.(1996)即使用該方法進行交叉路口肇事頻率預測分析。

2.3.4 類神經網路方法

類神經網路自 1957 年發展以來目前已受到廣泛的應用，它主要是透過電腦運算的方式來模擬動物神經元的運作方式，具有大量平行運算、高容量記憶能力、容錯能力及學習能力等特性，因此適用於複雜的肇事因果關係推導。Mussone et al. (1996)、Sayed et al. (1998)、趙崇仁 (1996)、楊憲國 (2002)、魏開元 (1999)均曾運用類神經網路方法構建肇事預測模式以應用於易肇事路段之鑑別，並獲得令人滿意的預測績效；周雍傑 (2000)採個體觀點探究肇事之嚴重程度，以類神經網路模式對都市地區的路口與路段分別構建財物損失、受傷及死亡之預測鑑別模式，所得總正確率相當高，但是在死亡預測鑑別之正確率卻相當低。此外 Hashemi et al. (1995)、吳志文 (2002) 亦以類神經網路探討船舶事故之研究。可知交通領域中，類神經網路方法已普遍運用於肇事因果關係的預測。

2.4 文獻回顧討論

經由上述文獻回顧，我們約略可知傳統統計模式運用於肇事預測有其先天缺陷，宜改用其他方法，另外多數研究係以總體觀點進行肇事預測，以個體觀點為之的研究缺乏，這兩點是本研究採用類神經網路進行肇事嚴重程度預測之主要原

因，以下分別就文獻回顧後所得結論予以說明。

一、傳統統計模式運用於肇事預測有其先天缺陷

以傳統迴歸模式分析肇事因果之不適性，主要在於肇事黑數及肇事發生稀少等特性，導致未能有足夠數量的統計樣本，使得所構建的模式績效不佳，反之如果為使蒐集足夠樣本，則必須擴大觀察區域或加長觀察期間，如此一來，將使個體間之差異盡被隱藏，難以對影響行車安全之因子作深入之探討。雖然大觀察區域或長觀察期間會有上述缺點，但是縮小觀測區域或時間，卻會使傳統迴歸面臨「不符合連續性假設」、「無法滿足常態性分配假設」、「違反同質性假設」等問題。

以肇事率作為因變數雖可消除離散性問題，然仍無法滿足常態分配之要求，如：肇事次數為 0 次之可能甚多，使得母體分佈趨向於右偏，即違反傳統迴歸假設母體為常態分佈的不偏假設。另外，如果假設肇事符合 Poisson 分配，進行 Poisson 迴歸分析，誤差項則會存在過度離散情形，使得「平均數等於變異數」之假設無法成立；故發現確有過度離散情形時，即應以負二項迴歸方法取而代之。另一個問題是，肇事因素相當多，且各因素間多具有相關性，如果均予考慮，會使得模式過於複雜，反之，簡化模式又會損失部分有用資訊，實為兩難。

二、多數研究係以總體觀點進行肇事預測

經由文獻回顧可知，大部分的研究均是以總體的觀點去尋找肇事因素與肇事率、肇事次數或肇事風險的關係，並用之於肇事預防及易肇事地點的改善。但是以肇事因素進行肇事結果個體因素的研究相當稀少，僅陳志和 (1999)及周雍傑 (2000)分別以 Logit 模式及類神經網路模式來構建肇事後肇事者受傷嚴重程度的判別，研究結果雖顯示總正確率相當高，但是肇事死亡案件的預測正確率卻不高，尤其是以 Logit 模式(陳志和，1999)進行預測死亡案件，結果顯示效果非常差，陳君推論應係肇事嚴重案件的樣本不足所導致。

綜上所述，可知傳統統計模式運用於肇事預測分析有其不適，如再考慮肇事因素間相關性高的情形，將使得預測分析更形困難，宜採用其他更適當的方法；另外，個體肇事嚴重程度預測模式如何應用於肇事防制或易肇事地點改善的方法，尚未有實例應用提出，這可能是多數研究採用總體觀點進行肇事預測的緣故，故以個體觀點建構肇事嚴重程度預測模式仍有相當大的發展空間。因此，本研究採用類神經網路方法進行個體肇事嚴重程度預測模式構建，試圖預測、探求易肇事地點，並尋思防制策略。

第三章 類神經網路理論基礎

為明瞭本研究所使用的類神經網路的理論基礎，本章就該方法相關理論基礎及使用應注意事項予以說明，其主要內容分為三個部分，首先就類神經網路的基本概念及理論基礎予以概述，其次則就本研究構建模式所採用的倒傳遞演算法(back-propagation algorithm)詳細說明，最後再對類神經網路構建應注意事項予以解說。

3.1 類神經網路概述

建立傳統數學模式的困難在於面對複雜、非線性的問題時，必須經由一些假設、簡化環境後才能建構出方程式。因此傳統數學模式常常需要在簡化的模式與正確的模式中作一取捨，其中，正確的模式可能較難以建構取得，而簡化的模式又與現實有所差異；然類神經網路則可藉由學習與記憶的方式來處理複雜且具不確定性的問題。隨著電腦科技的進步，建立類神經網路模式所需的龐大計算過程已可輕易完成。故本研究基於此一理由，從而選用類神經網路作為研究工具，在此將針對類神經網路的基本概念與架構及所欲使用的倒傳遞網路模式，分別予以說明。

3.1.1 類神經網路之基本概念

依魏健宏(2000)與葉怡成(2000)對類神經網路的定義：「類神經網路是一種模仿生物神經網路功能的資訊處理系統，其基本原件可接受、處理並傳遞各類資訊，藉適當的連接組合，它能顯示與生物腦類似的功能。」、「類神經網路是一種模仿生物神經系統功能的資訊處理系統，是一種計算系統，包括軟體與硬體，它使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，它從外界環境或者其他人工神經元取得資訊，加以簡單的運算，並輸出其結果到外在環境或其他人工神經元。」，至於類神經網路原理介紹如下。

一、生物神經系統的原理

人類的大腦是由數百億個「神經細胞」(nerve cells)所組成的一個高度連結網狀的神經網路，而且每個神經細胞有數千根「突觸」(synapses)與其它神經細胞相連。由於神經網路具有大量的連結，若均以平行方式進行運作，則其速度將能大大地提升，且具有驚人的記憶容量。此外神經細胞藉由適當的調整使整個網路具有強大的學習能力，可處理分類、聯想等問題。再者，生物神經網路中若有部分細胞受損，對其正常的功能影響不大，此乃因網路的資訊儲存是採分散記憶方式，而該特性使得網路在正常運作下仍能對不完整或有雜訊的資訊作適當的處理。

生物的神經系統是以神經細胞為單位，神經細胞通常伸出許多細長的分支，短的叫做「樹突」(dendrite)，長的叫「軸突」(axon)，這些分支將各個神經細胞連結成一個複雜的神經網路系統。E.D.Adriant 在 20 世紀初期發現當神經細胞的輸入神經樹與外界接觸而受到刺激被激發的時候，該神經元的輸出

神經樹將放出電流脈衝(current pulse)到其他的神經細胞，神經細胞被激發時所放出的電流脈衝強度約略相同，而放出電流脈衝的頻率則與受到刺激的強度有關。根據 E.D.Adrian 對神經細胞電化學的說法：『當一個神經元受到足夠的刺激後，發出一個定值的電流脈衝。』可以用一個簡單的結構以及運算規則來模擬生物神經細胞的運作，如圖 3-1 所示。

Hebbian 在 1949 年提出神經細胞的學習規則：『當人腦在學習的時候，每個腦細胞之間的連結隨時都在改變。當一個腦神經細胞受到另一個腦神經細胞不斷的連續刺激，則這兩個腦細胞之間的連結力量就會增強。』該學習規則是否就是實際的神經細胞學習行為，目前仍有許多爭議，然此學習規則卻成為類神經網路研究的基礎。

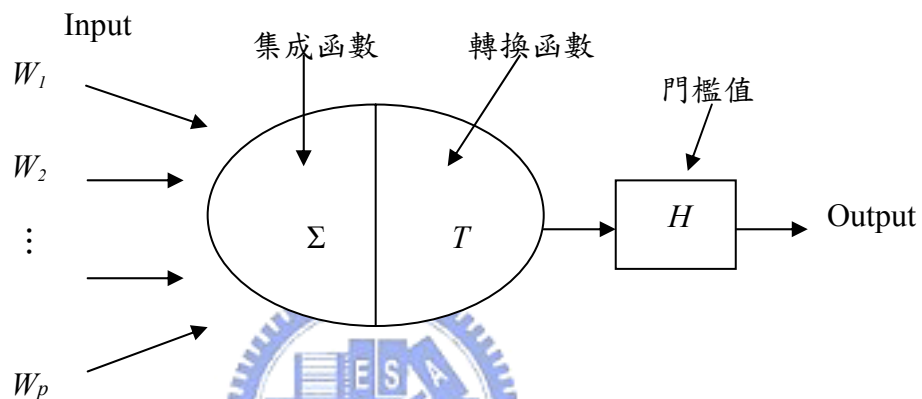


圖 3-1 處理單元的作用

二、類神經網路的特性

類神經網路(ANN)的概念是來自生物神經網路，在邏輯上的組成單元與構造皆十分類似於生物神經網路。除了其表面或外形的相似外，類神經網路也擁有許多與人類大腦相同功能的特性，其中最為重要的三項功能：(1)學習、(2)回想，與(3)歸納推演，分別概述如下：

(1)學習(Learning)

類神經網路可隨外在環境的變化來改變其行為模式。為了使類神經網路可於動態環境中發揮功效，強化或是抵銷輸入項變數的刺激，適當的學習模式是類神經網路系統的基本條件。類神經網路的學習過程可藉由下列幾個方法完成：(a)建立神經元間的連結模式、(b)修正連結神經元間的權重、(c)調整神經元活化函數中的門檻值、或是(d)結合運用上述幾種方式。目前類神經網路多樣化學習的演算法已有成熟的發展，這些演算法可分成兩類，一類是即時線上學習(on-line learning)，即依據新增或即時的資訊，不斷地修正調整網路的連結權重值；另一類是先將網路訓練好才能開始使用網路(off-line learning)，因為所有的訓練資料(training data)是一次一整批的用於修正網路的連結權重，故亦稱為批次學習(batch learning)。

(2)回想過程(Recall)

當類神經網路接受到一個輸入刺激，進而依據網路架構產生一個輸出

值時，我們便稱這一個過程為回想過程。一個類神經網路的效能及強健性與否，主要就在於其回想的過程是否快速有效。

回想過程可以在前饋式類神經網路的模型中表現出來，亦或是直接表現在輸入與輸出的過程中。前饋式類神經網路並不能記憶前一次輸入與輸出的映射過程與結果，網路的運算僅依據目前的輸入狀況來決定，反應結果與前一次的狀況可說是毫無關連。換句話說，前饋式類神經網路的表現，只是反映出目前輸入資料，直接從網路儲存的案例找到一對應的輸出，屬於瞬間表現的回想過程。

另一類回饋式類神經網路，因輸入項與輸出項之間具互相連結的作用，即本時刻(t)之輸入條件的輸出結果，可回饋至下一時刻($t+1$)，變成下一時刻的一項輸入條件，從而影響到 $t+1$ 時刻的輸出結果，故無法從訓練過程獲得單一的輸入與輸出映射關係，因輸入與輸出的映射關係不斷被修正與調整，故此一系統可視為動態系統，常用於描述非線性動態系統的行為。

(3) 歸納推演過程(Generalization)

歸納推演法是從一個系統中局部觀察描述出其整體特性的過程；不論是從特殊例子推演到整體的事件，或者是從一個或數個範例的認知去定義出物件的種類，我們稱這樣的過程為歸納推演過程。歸納推演過程是網路模式中不可或缺的部分；因為歸納推演過程能夠讓我們找出所有可以適用於全部類別物件的重要因子，而非僅將記憶中的特殊事件運用於個別物件。歸納推演過程提供了我們一個有效率的記憶與儲存模式。若是沒有歸納推演過程，我們便可能會以記憶與回想組合出無限個特殊事件、因子與關係。

類神經網路對輸入資料具有萃取其特徵的能力。經過訓練類神經網路，可以對曾見過但有些微差異的輸入訊號作出正確的輸出；重要的是這種類似內插的功能並非使用數學函數來運算的，而是由類神經網路的架構自動產生這樣的特性。

類神經網路具有上述優越的特性，並不表示類神經網路適合於各種問題上。對於一些需要準確的答案、已經存在有效率的演算法或具有確切的物理機制等問題就不適用，例如計算薪資就不適合使用類神經網路。而一些傳統電腦處理成效不好，但人類卻可以輕鬆處理好的問題，即可嘗試使用類神經網路來處理，例如辨識問題、複雜的天氣問題等。適合使用類神經網路解決的問題通常有以下的特徵：

- 問題及相關條件難以完整定義
- 需要快速得到問題解答且解答不用完全精確
- 問題非常複雜或是非線性的問題，無法由一連串已知的數學方程式來描述並求得解答者。

目前以類神經網路作為解決方案之一，且有不錯效果的問題有：語音辨識、文字辨識、天氣預測、股票指數預測、機器人控制、售貨員最佳路徑問

題(Travelling Salesman Problem, TSP)問題、雷達偵測、影像識別、汽車自動駕駛、醫學檢測、蛋白質 3D 結構檢測等。

三、類神經網路的分類

依學習法則(Learning Rule)來區分類神經網路，共可分成監督式學習網路、非監督式學習網路、聯想式學習網路、最適化應用網路等四類，說明如下：

(1)監督式學習網路(Supervised Learning Network)

從問題中取得訓練範例(有輸入和輸出變數值)，並從中學習輸入變數與輸出變數的內在對應規則，以應用新的案例(只有輸入變數值，而需推導輸出變數值)。主要的模式有感知機網路(Perceptron Network)、倒傳遞網路(Back-Propagation Network, BPN)、機率神經網路(Probabilistic Neural Network, PNN)、學習向量量化網路(Learning Vector Quantization, LVQ)、反傳遞網路(Counter-Propagation Network, CPN)。

(2)非監督式學習網路(Unsupervised Learning Network)

從問題中取得訓練範例(只有輸入變數值)，並從中學習範例的內在聚類規則，以應用於新的案例(有輸入變數值，而需推導它與哪一訓練範例屬同一聚類的應用)。主要的模式有自組織映射圖網路(Self-Organizing Map, SOM)、自適應共振網路(Adaptive Response Theory, ART)。

(3)聯想式學習網路(Associate Learning Network)

從問題領域中取得訓練範例(狀態變數值)，並從中學習範例的內在記憶規則，以應用於新的案例(只有不完整的狀態變數值，而需推論其完整的狀態變數值的應用)。包括霍普菲爾網路(Hopfield Neural Network, HNN)及雙向聯想記憶網路(Bi-directional Associative Memory, BAM)。

(4)最適化應用網路(Optimization Application Network)

對問題決定設計變數值，使其在滿足設計限制下，使設計目標達最佳狀態的應用。包括霍普菲爾-坦克網路(Hopifield-Tank Neural Network, HTN)及退火神經網路(Annealed Neural Network, ANN)。

另外依網路連結架構(Connecting Type)來區分類神經網路，可分為前向式網路與回饋式網路兩種，說明如下：

(1)前向式架構(Feed-Forward)網路：處理單元分層排列，形成輸入層、輸出層、隱藏層(若干層)。每一層只接受前一層的輸出作為輸入者，稱為前向式架構。

(2)回饋式架構(Feedback)網路：從輸出層回饋到輸入層，或者層內各處理單元間有連結者，或者處理單元不分層排列只有一層，各處理單元均可相互連結者稱為回饋式架構。

3.1.2 類神經網路的系統架構

類神經網路的架構可從網路的基本結構與構成單元分析，基本結構包括處理單元、層、網路，並透過不同的連接方式、學習規則等來構成，試摘要說明如下：

一、處理單元(Processing Element)：為類神經網路組成的基本單元或稱人工神經

元，其次由多個處理單元組成「層」，最後再由多個層以各種連結方式連合成「網路」。而處理單元的運算邏輯可用集成函數、作用函數、轉換函數等來加以說明，其運作過程如圖 3-1 所示。

(1)集成函數(Aggregation Function)：將其他處理單元輸出訊息，透過網路連結並將傳來的訊息加以綜合再輸出，一般可用公式(3-1)表示，其中以加權乘積和最為常用，如公式(3-2)。

$$I = F(W, X) \quad (3-1)$$

$$I_j = \sum_i W_{ij} X_i \quad (3-2)$$

其中 I_j = 第 j 個處理單元之加權乘積和

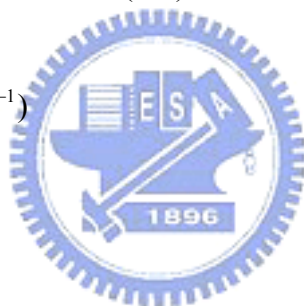
W_{ij} = 第 j 個處理單元與第 i 個處理單元間的加權值

X_i = 第 i 個處理單元傳來的訊息

(2)作用函數(Activity Function)：作用函數的目的是將集成函數值與處理單元目前的狀況加以綜合，如公式(3-3)，使用時常以直接使用集成函數輸出，如公式(3-4)。

$$net_j^n = F(I_j^n, I_j^{n-1}, net_j^{n-1}) \quad (3-3)$$

$$net_j^n = I_j^n \quad (3-4)$$



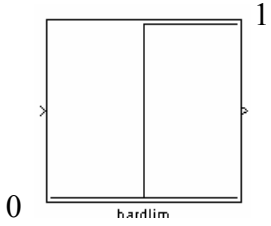
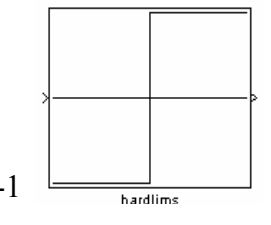
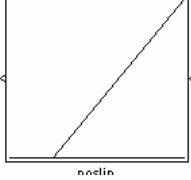
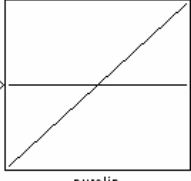
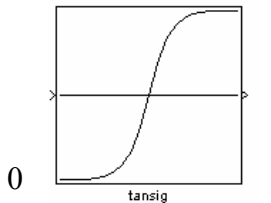
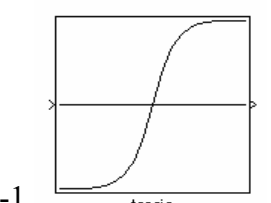
(3)轉換函數(Transfer Function)：轉換函數的作用在於適當的轉換輸入 net_j 與

輸出 Y_j ，係將作用函數的輸出值轉成處理單元的輸出，如公式(3-5)。常

用的轉換函數包括硬限函數、線性函數以及非線性函數等，見表 3-1。

$$Y_j = F(net_j) \quad (3-5)$$

表 3-1 常見的轉換函數

名稱	函數	圖形
硬 限 (Signum) 函 數	Perceptron 函數 $Y_j = \begin{cases} 1 & \text{if } net_j > 0 \\ 0 & \text{if } net_j \leq 0 \end{cases}$	Perceptron 函數 
	Hopfield-Tank 函數 $Y_j^n = \begin{cases} 1 & \text{if } net_j > 0 \\ Y_j^{n-1} & \text{if } net_j = 0 \\ 0 & \text{if } net_j < 0 \end{cases}$	
	Signum 函數 $Y_j = \begin{cases} 1 & \text{if } net_j > 0 \\ -1 & \text{if } net_j \leq 0 \end{cases}$	Signum 函數 
	Signum0 函數 $Y_j = \begin{cases} 1 & \text{if } net_j > 0 \\ 0 & \text{if } net_j = 0 \\ -1 & \text{if } net_j < 0 \end{cases}$	
	BAM 函數 $Y_j = \begin{cases} 1 & \text{if } net_j > 0 \\ Y_j^{n-1} & \text{if } net_j = 0 \\ -1 & \text{if } net_j < 0 \end{cases}$	
線性函數	門限線性函數 $Y_j = \begin{cases} net_j & \text{if } net_j > 0 \\ 0 & \text{if } net_j \leq 0 \end{cases}$	
	無門限線性函數 $Y_j = net_j$	
非 線 性 函 數	雙 彎 曲 函 數 (Sigmoid function) $Y_j = \frac{1}{1 + \exp(-net_j)}$	
	雙 曲 線 正 切 函 數 (Hyperbolic tangent function) $Y_j = \frac{e^{net_j} - e^{-net_j}}{e^{net_j} + e^{-net_j}}$	

二、層(Layer)：若干個具有相同作用的處理單元集成「層」；類神經網路主要由三層所組成，說明如下：

- (1)輸入層：用以表示網路之輸入變數，其處理單元依問題而定。
- (2)隱藏層：用以表示輸入處理單元間的交互影響，其處理單元數目並無標準方法可以決定，須以試驗方法決定其最佳數目，其所使用轉換函數常以非線性轉換函數為主，另外網路可有多層隱藏層，亦可沒有隱藏層，在進行網路構建時應從無隱藏層開始。
- (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，通常使用非線性轉換函數作為其轉換函數。

三、網路(Network)：係若干個具不同作用的層集成網路，將其處理單元分層排列，可分為輸入層(Input Layer)、輸出層(Output Layer)、隱藏層(Hidden Layer)，如圖 3-2。每一層只接受前一層的輸出作為輸入者，稱為前向式(Forward)架構。從輸出層回饋到輸入層，或者層內各處理單元間有連結者，或者處理單元不分層排列只有一層，各處理單元均可相互連結者稱為回饋式(Backward)架構。基本上，前向式架構網路相對於回饋式架構網路具有較大容量、運算較快、系統較穩定等優點；而後者可提供較佳的非線性轉換能力，兩者各有優劣點，除基本的全連接式(Fully Connected)網路外，亦可視系統特性或依採用變數之不同，研擬部份連接式(Partially Connected)網路，以提昇整體系統績效。若有關網路的「層」與「層」之間的連結型態不同，其所構成的網路模式亦不同。

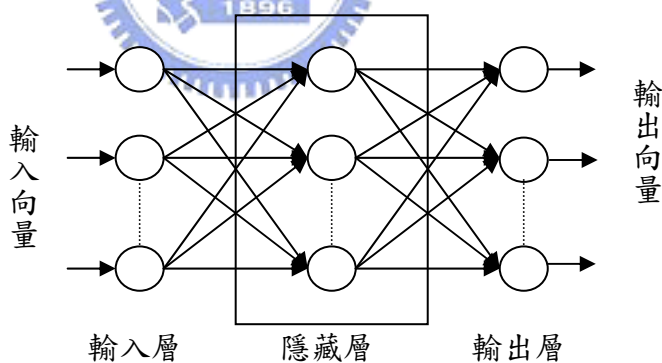


圖 3-2 類神經網路架構

3.1.3 網路學習演算法

類神經網路優於其他數學模式或方法的主要因素，就是其具有學習(Learning)的功能，模式在學習的過程中，因資訊的相互作用，逐漸修正模式，使修正後的模式反應能優於或不同於前一模式，此程序經反覆進行，直到模式可對特定資訊產生期望的特定反應為止。然而，學習過程可具體表現在處理單元連接加權值(Weights)的改變上，至於如何才能達到學習尋優的目的，則必須依賴適當的學習法則(Learning Rule)，使模式能針對外界輸入的資訊，自動調整連接加權值，達到輸出值能夠符合期望狀況。有關類神經網路之學習法則可分為監督式

(Supervised)、非監督式(Unsupervised)與聯想式(Associated)等三類。因本研究所欲採用之倒傳遞類神經網路係使用監督式學習法則，故以下僅就該演算法加以說明，其他學習演算法則不予贅述。

監督式學習目的在降低網路目標輸出值與推論輸出值間之差距。一般以能量函數(或稱誤差函數)來表示學習的品質，因此，網路的學習過程成為尋求能量函數最小化的過程，監督式學習法則說明如下：

$$E = \frac{1}{2} \sum_j (T_j - Y_j)^2 \quad (3-6)$$

其中： T_j = 目標輸出值

Y_j = 推論輸出值

欲使上述能量函數最小化，通常以最陡坡降法(The Gradient Steepest Descent Method)達成：

$$\Delta W_{ji} = -\eta \frac{\partial E}{\partial W_{ji}} \quad (3-7)$$

$$W_{ji}(n+1) = W_{ji}(n) + \Delta W_{ji}(n) \quad (3-8)$$

$W_{ji}(n)$ 與 $W_{ji}(n+1)$ ：原先及調整後的鍵結值

$\Delta W_{ji}(n)$ ：神經元受刺激後，為達學習效果，所採的改變量

η ：學習速率(Learning Rate)

回想過程則依回想演算法，輸入資料以決定網路輸出資料的過程。

3.2 倒傳遞演算法(back-propagation algorithm)

此演算法的學習過程是由正向傳遞和反向傳遞所組成；在正向傳遞過程中，輸入訊號從輸入層經隱藏層逐層處理，並傳向輸出層。每一層神經元的狀態只影響下一層神經元的狀態。如果輸出層不能得到期望的輸出，則轉入反向傳遞，將誤差沿原來的連接通路返回，透過修改各層神經元的連接線加權值，使得誤差逐漸縮小，而網路的運作過程分成二個階段：

3.2.1 學習過程(Learning)-反向傳遞

網路依學習演算法學習，以調整網路連結加權值的過程。

- 設定網路參數。
- 以均勻分布的隨機亂數設定加權值矩陣 W_{xh} 及 W_{hy} ，與偏權值向量 θ_h 及 θ_y 的初始值。
- 輸入一個訓練範例的輸入向量 X ，與目標輸出向量 T 。
- 計算推論輸出向量 Y 。
 - (1) 計算隱藏層輸出向量 H
 - (2) 計算推論輸出向量 Y

$$net_h = \sum_i W_{_xh_{ih}} \cdot X_i - \theta_{_h}$$

$$H_h = f(net_h) = \frac{1}{1 + \exp^{-net_h}}$$

$$net_j = \sum_h W_{_hy_{hj}} \cdot H_h - \theta_{_y_j}$$

$$Y_j = f(net_j) = \frac{1}{1 + \exp^{-net_j}}$$

· 計算差距量 δ

$$(3) \text{ 計算輸出層差距量 } \delta_j = Y_j(1 - Y_j)(T_j - Y_j)$$

$$\text{計算隱藏層差距量 } \delta_h = H_h(1 - H_h) \sum_j W_{_hy_{hj}} \cdot \delta_j$$

· 計算加權值矩陣修正量 ΔW ，及偏權值向量修正量 $\Delta \theta$

$$(4) \text{ 計算輸出層加權值矩陣修正量 } \Delta W_{_hy}，\text{及偏權值向量修正量 } \Delta \theta_{_y}$$

$$\Delta W_{_hy_{hj}} = \eta \cdot \delta_j \cdot H_h$$

$$\Delta \theta_{_y_i} = -\eta \cdot \delta_j$$

$$(5) \text{ 計算隱藏層加權值矩陣修正量 } \Delta W_{_xh}，\text{及偏權值向量修正量 } \Delta \theta_{_h}$$

$$\Delta W_{_xh_{ih}} = \eta \cdot \delta_h \cdot X_i$$

$$\Delta \theta_{_h_i} = -\eta \cdot \delta_h$$

· 更新加權值矩陣 W ，及偏權值向量 θ

$$(6) \text{ 計算輸出層加權值矩陣 } W_{_hy}，\text{及偏權值向量 } \theta_{_y}$$

$$W_{_hy_{hj}} = W_{_hy_{hj}} + \Delta W_{_hy_{hj}}$$

$$\theta_{_y_j} = \theta_{_y_j} + \Delta \theta_{_y_j}$$

$$(7) \text{ 計算隱藏層加權值矩陣 } W_{_xh}，\text{及偏權值向量 } \theta_{_h}$$

$$W_{_xh_{ih}} = W_{_xh_{ih}} + \Delta W_{_xh_{ih}}$$

$$\theta_{_h_i} = \theta_{_h_i} + \Delta \theta_{_h_i}$$

· 重複步驟 3 至步驟 7，直到收斂（誤差不再有明顯變化）或執行一定數目的學習循環。

3.2.2 回想過程(Recalling)-正向傳遞

網路依回想演算法以輸入資料決定網路輸出資料的過程。

- 設定網路參數。
- 讀入加權值矩陣 $W_{_xh}$ 及 $W_{_hy}$ ，與偏權值向量 $\theta_{_h}$ 及 $\theta_{_y}$ 。
- 輸入一個測試範例的輸入向量 X 。
- 計算推論輸出向量 Y 。

$$(8) \text{ 計算隱藏層輸出向量 } H$$

$$net_h = \sum_i W_{ih} \cdot X_i - \theta_h$$

$$H_h = f(net_h) = \frac{1}{1 + \exp^{-net_h}}$$

(9) 計算隱藏層輸出向量 Y

$$net_j = \sum_h W_{hj} \cdot H_h - \theta_j$$

$$Y_j = f(net_j) = \frac{1}{1 + \exp^{-net_j}}$$

上述之網路學習過程通常以一次一個訓練範例的方式進行，直到所有的學習範例均學習完畢，此過程稱為一個學習循環(Learning Cycle)，一個網路可以將訓練範例反覆學習直到網路收斂為止。為了檢驗網路之學習成果，通常將所收集之範例隨機分成二部份，一部份當作訓練範例，另一部份則留作測試範例，在網路學習幾個循環後，可將測試範例載入網路中，檢驗網路的誤差程度是否收斂。

3.3 類神經網路構建應注意事項

3.3.1 參數

倒傳遞網路中有幾個重要參數，包括隱藏層處理單元數目、隱藏層層數與學習速率，進一步說明如下：

一、隱藏層處理單元數目

通常隱藏層處理單元之數目越多收斂越慢，但可達到更小的誤差值，特別是「訓練範例」誤差。但超過一定數目後，再增加則對降低「測試範例」誤差幾乎沒幫助，徒然增加執行時間。這可解釋成隱藏層處理單元之數目太少，不足以反映輸入變數間的交互作用，因而有較大的誤差，而數目越多，雖然可達到更小的誤差值，但因網路較複雜，因而收斂較慢。為平衡品質與成本，以取適當的數目為宜。一般而言，隱藏層處理單元數目的選取原則如下：(葉怡成，1997，1998)

- 公式 1：隱藏層單元數目=(輸入層單元數+輸出層單元數)/2
- 公式 2：隱藏層單元數目=(輸入層單元數 · 輸出層單元數)^{1/2}
- 問題雜訊高，隱藏層單元數目宜少。
- 問題複雜性高，隱藏層單元數目宜多。
- 測試範例誤差遠高於訓練範例誤差，隱藏層單元數目宜減少；反之，宜增加。

另外確定隱藏層神經元個數亦可使用網路修剪法(pruning algorithm)或網路增長法(constructive algorithm)，網路修剪法是由 1998 年 Abrahart 等人提出，這個方法是在一開始即設定大數量的隱藏層神經元個數進行訓練，隨後逐一減少，直到找到適當的神經元數為止，其缺點是浪費大量計算時間，較不經濟；而網路增長法則是由 1997 年 Kwok 和 Yeung 提出，該法與網路修剪法相反，一開始設定較少數量的神經元個數，逐一增加至適當的神經元數 (張斐章等，2003)。

二、隱藏層層數

通常隱藏層之數目為一層到二層時有最好的收斂性質，太多層或者太少層其

收斂結果均較差。通常有一或二層隱藏層已足以反應輸入單元間的交互作用，太多的隱藏層反而使網路過度複雜，造成更多局部最小值，使得在修正網路加權值時易掉入誤差函數的局部最小值，而無法收斂。至於設一層隱藏層好還是二層，在不同研究或問題中有不同結論(張斐章等，2003)，依據經驗，一般使用一層隱藏層的網路架構，即可描述大部分的問題。通常，在用倒傳遞網路求解問題時，一定要先用無隱藏層架構作作看，如果其精確性比有隱藏層者為佳，則此問題必不適合用倒傳遞網路解，其理由為無隱藏層架構的倒傳遞網路其效果接近統計學上一些有「線性」假設的方法，例如線性迴歸、區別分析，因此無理由用一個複雜的網路解答去取代一個有明確公式的統計學解答。

三、學習速率

學習速率的大小，會影響收斂過程的快慢，通常學習速率太大或太小均不利於網路的收斂。因較大的學習速率，有較大的網路加權值修正量，可較快逼近函數最小值，但也可能導致網路加權值修正過量，造成數值振盪，難以收斂，而較小的學習速率，則收斂過程速率緩慢，因此學習速率的大小對學習有很大的影響。學習速率在相當大的範圍內均有良好的收斂性，一般常將學習速率的值設定在 0.01 到 1.0 間，對學習的結果並不會過於敏感，但仍有些問題的適當學習速率可能低到 0.01 以下或高到 10 以上。

上述這些參數都有一些經驗值，然而，假使參數的設定一旦變動，學習的效果與精度可能會有很大差異，因此，本研究將先以經驗值進行實證，然後調整各參數值以比較結果之差異，再決定適合本研究之參數值。

3.3.2 範例

一、網路輸入變數與輸出變數的選用原則：

- 獨立性(Independence)：每一個輸入變數的值應是互斥的，即一個案例中，一個輸入變數只有一個值。
- 必需性(Necessity)：每一個輸入變數應是對決策有影響的，提供太多沒必要的輸入變數，反而會降低網路學習的品質。
- 充份性(Sufficiency)：如果訓練範例中，有任兩個範例具有相同的輸入變數值，卻有不同的輸出變數值，則此訓練範例所考慮到的輸入變數可能不充份，應考慮更多的輸入變數來解決此問題。
- 可用性(Availability)：每一個輸入變數均應可很清楚地觀察到，而不需要高深的經驗、冗長時間、或昂貴成本。

當變數是連續值時，須注意：

- 正規化：當連續變數的值域很大時，應加以正規化，以利網路學習。
- 比率化：比率化的連續變數比非比率化的變數更有意義。

二、範例取樣方法

範例取樣是指將原始資料檔中的範例，分成訓練範例與測試範例，方法可分為二種：

- 隨機取樣法：以隨機方式決定一範例是作訓練範例或測試範例。

- 循序取樣法：以循序方式決定一範例是作訓練範例或測試範例，即前面的範例作訓練範例，後面的則作測試範例，適用於有時間先後關連性的問題。

3.3.3 測試及驗證

網路訓練方式通常將範例隨機分成兩部份，一部份作為訓練範例，另一部份作為測試範例，在網路學習階段，可每學習幾個學習循環，即將測試範例載入網路。所得到的輸出向量，與實際結果比對評估其收斂效果以判斷網路成效，收斂結果可分為下述四類。

第一類：訓練範例的誤差收斂良好，測試範例誤差收斂良好。

意義：表示訓練範例具代表性，範例數量充分。

第二類：訓練範例的誤差收斂良好，測試範例誤差收斂不好。

意義：表示可能是訓練範例過於稀少，不足以學得分類規則，或者訓練範例的訓練規則根本就與測試範例不同。

第三類：訓練範例的誤差收斂不好，測試範例誤差收斂良好。

意義：表示訓練範例只能學部份分類的分類規則，因而誤差大。而在測試範例中，其分類大多為訓練範例能學得分類原則，因而誤差小。

第四類：訓練範例的誤差收斂不好，測試範例誤差收斂不好。

意義：表示可能訓練範例本身就存在矛盾範例，無法學得分類或預測原則，或網路的參數設定不當。

