

南華大學  
資訊管理學系碩士班  
碩士論文

以資料萃取技術探索  
天氣、污染、氣喘病發作的關連性  
—以類神經網路 BPN 模型為例

指導教授：鍾志明 博士

韓欽銓 博士

研究生：李修宇

中華民國九十年六月

## 摘要

近百年來科技發達，創造了許許多多的產品改善人們的物質生活，但也同時製造了大量的污染物質，讓我們賴以維生的地球遭到破壞，而有所謂之文明病的產生，人們因此嚐到自己所種下的苦果。本研究欲利用資料萃取的技術找出天氣、空氣污染、健保記錄間的關聯，提供學者做進一步的相關研究。

本研究中將資料萃取技術應用在三個異質資料庫：天氣氣象（中央氣象局）、空氣污染（環保署）、健保記錄（健保局），這些資料彼此之間存在著微妙的關係，利用類神經網路中的 BPN 模型，建立起模擬預測的模型，藉以從中發掘知識。其方法是利用天氣氣象、空氣污染等因子以建立預測模型，藉著模型的建立來預測氣喘病的發作，並考量區域性的因子以加強預測的正確性。

最後，我們建立出全台灣 16 個地區的氣喘病預測模型，並將所有研究成果列於結論部份，咸信此項成果可以讓專家學者避免我們曾經走過的歧路，進而找出未來研究的方向，對於氣喘病的預防及治療有所裨益。

關鍵字：天氣氣象、空氣污染、健保、資料萃取、氣喘病、類神經網路、BPN(Back-propagation Network)

## 作者簡歷

姓名：李修宇 (Shiou-Yu Lee)

生日：1973.11.02

學歷：新埔技術學院機械工程科一般組 (1994 年畢)

大華技術學院資訊管理學系 (1999 年畢)

南華大學資訊管理系碩士班

住址：(242)台北縣新莊市思源路 419 巷 20 號 5 樓

E-mail Address：[jason008@kimo.com.tw](mailto:jason008@kimo.com.tw)

著作及論文發表：

- 戴崑崧，李修宇，葉彥伶，董國龍，“金鑰密碼系統 - 秘密金鑰 DES 系統”。(大華技術學院畢業專題)1999。
- 謝建成，史孟祥，李修宇，謝馥安，“非關聯性資料庫之資料倉儲建立 - 以天氣氣象與空氣污染為例”，第六屆資訊管理研究暨實務研討會，2000。
- 韓欽銓，李修宇，“以資料萃取技術探索天氣、污染、氣喘病發作間的關連性 - 以類神經網路中 BPN 模型為例(樣本空間：大台北地區)”，2001 年資訊管理學術暨實務研討會，2001。

研究領域：Data Mining, Security, Neural Network

## 誌謝

以前的我劃地自限，以為我僅能念到專科畢業，在因緣際會的巧合下，我在退伍後得已由二技一路念上碩士學位，這是當初始料未及的。憑著一股不服輸的衝勁，親人、朋友的鼓勵，以及韓欽銓恩師、鍾志明老師時時予以關鍵性的協助，使得學生得已順利完成碩士論文。在學期間，因為有謝建成老師、蔡鴻旭老師及邱宏彬老師的授業，豐富了我專業方面的學養，由於這些老師的啟發及教導，使學生得已在論文撰寫上更得心應手；並且蔡老師、邱老師在擔任口試委員期間，提供了許多寶貴的建議，使得論文更趨完備。

碩士學業的求學過程，一路走來歷經風風雨雨，心中有著許許多多的感觸，幸好身旁有著關心我的朋友，妳（你）們不斷地噓寒問暖，滋潤、修補我受創的心，讓我擁有繼續前進的動力。而在論文撰寫上，我要特別感謝富宏學長、苑荏學姐，因為有你們提供的經驗及專業知識，並不斷地予以指正，才有如今論文的完稿。

感謝這兩年間陪我走過的同學及學弟（妹）- 秀女、玉美 ... .. 等等，沒有你們的關懷及聆聽，我的人生不會這麼地順遂；感謝我的父母，沒有您們的苦心栽培及細心呵護，就沒有今日的我。謝謝您們陪我走過這段歲月，也感激您們所做的一切，不但豐實了我的生活，更留下許多珍貴美好的回憶。

最後謹以本文獻給所有愛我、關心我的親友，我誠摯期盼將這份喜悅及小小的成就，與您們一同來分享。

## 目錄

<b>第一章 緒論</b> .....	<b>1</b>
1.1 研究動機 .....	1
1.2 研究目的 .....	2
1.3 研究範圍與限制.....	2
1.4 論文結構 .....	3
<b>第二章 相關研究</b> .....	<b>6</b>
2.1 氣象知識 .....	6
2.2 空氣污染的產生及危害 .....	9
2.3 健保及氣喘病 .....	13
2.4 類神經網路 .....	17
2.5 資料萃取技術 .....	22
<b>第三章 研究方法</b> .....	<b>32</b>
3.1 BPN 模型說明 .....	32
3.2 BPN 演算法.....	34
3.3 研究設計 .....	39
3.4 研究建置步驟 .....	40
3.5 實例說明 .....	48
<b>第四章 實驗結果</b> .....	<b>52</b>
<b>第五章 結論與未來研究方向</b> .....	<b>60</b>
5.1 結論.....	60
5.2 未來研究方向 .....	61
<b>參考書目</b> .....	<b>63</b>

附錄 .....69

## 圖目錄

圖 1.1 : 研究流程	4
圖 2.1 : 氣象地面觀測站分布圖	9
圖 2.2 : 空氣污染與氣喘的關連性	13
圖 2.3 : 氣象與氣喘的關連性	14
圖 2.4 : 氣象、空污、氣喘病間關聯性	14
圖 2.2 : 迴歸式網路	18
圖 2.3 : 前授式網路	18
圖 2.4 : 動態式網路	19
圖 3.1 : BPN 網路架構	32
圖 3.2 : 人工神經元圖	34
圖 3.3 : 研究設計圖	41
圖 3.4 : 氣象及空氣污染記錄的欄位篩選	49

## 表目錄

表 2.1：全台空氣品質監測站分布	11
表 2.2：「門診處方及治療明細檔」說明	15
表 2.3：「醫事機構基本主檔」說明	16
表 2.4：學者對 Data Mining 的定義	23
續表 2.4：學者對 Data Mining 的定義 (2)	24
表 2.5：學者將資料萃取技術應用於各領域	29
續表 2.5：學者將資料萃取技術應用於各領域 (2)	30
續表 2.5：學者將資料萃取技術應用於各領域 (3)	31
表 3.1：研究中氣象資料的收集	42
續表 3.1：研究中氣象資料的收集 (2)	43
表 3.2：彙整後的 Database Table(當日版)	45
表 3.3：彙整後的 Database Table(時序版)	46
表 3.4：知識探索成果 (大台北地區) (單位：%)	51
表 4.1：第一階段實驗結果比對 (單位：%)	53
續表 4.1：第一階段實驗結果比對 (單位：%) (2)	54
表 4.2：第二階段實驗結果比對-台北、高雄 (單位：%)	55
表 4.3：第二階段其餘 14 縣市實驗結果 (單位：%)	56
續表 4.3：第二階段其餘 14 縣市實驗結果 (單位：%) (2)	57
表 4.4：第三階段實驗結果(挑選 8 日比對)-台北	58
續表 4.4：第三階段實驗結果 (挑選 8 日比對) -台北 (2)	59



## 謝辭

本文作者要[按一下這裡並鍵入謝辭]

## 辭彙

文字. [按一下這裡並鍵入定義]

# 第一章緒論

## 1.1 研究動機

近百年來科技發達，創造了許許多多的物資及產品，來改善人們的物質生活，但也同時製造了大量的污染物質，讓我們賴以維生的地球遭到破壞，因此人們得到一些以前從未聽聞或是提早面臨的病症，也就是一般人所稱的『文明病』。這些病症或許有藥可治或者是無藥可救，並且這些病的成因也讓專家覺得頭疼，因為這都是文明科技所帶來的副產品，是我們拋不開也丟不掉的。所幸，目前已有為數眾多的專家學者致力於污染的防治及病症的醫療研究；可是，這些的成果或許不是五年十年就能有所成就的，所以在他們的進行研究之前，我們希望先為他們找尋一些蛛絲馬跡，讓他們可以減少走錯路的時間，或許可以找到一些知識，作為專家找出疾病的因果關係甚至是治癒方法的指引。

當前，專家領域的知識可以確定：空氣污染及天氣現象確實會影響氣喘的發作，但影響的因子甚多（如：降雨量、氣溫、相對濕度、風速、風向、二氧化碳、一氧化碳、臭氧...），且兩者間的關係又有著微妙的複雜關聯性，（這些影響因子若以統計/分析的方法，一個個地將影響的程度釐清，是需要花上很多時間及精神！），這樣的複雜關係可借助類神經網路的協助，在神經元相互連接的關係中即便可以找出影響的程度，所以本文需要仰賴類神經網路建立起預測模型，期能真實地展現兩者的趨勢及兩者的相互影響程度，以維持預測氣喘病發作的準確度，讓研究模型趨近真實的狀況，以提高可信度。

## 1.2 研究目的

本研究在於建立起『疾病記錄』、『天氣記錄』、『空氣品質記錄』的資料庫，利用資料萃取的技術及類神經網路的預測功能，以期能找出病症發作/惡化的導火線，並且提出預測模型以供參考。因此，本研究預期朝向幾個目標努力：

1. 以實際的各地/每日氣喘病看診人次，計算出氣喘病發作的可能性。
2. 探討研究中相關背景知識，以找出三個領域間互相影響的因子。
3. 探尋類神經網路 BPN 模型及建置過程。
4. 訓練出氣喘病發作的預測模型。
5. 驗證資料萃取技術加上 BPN 模型，於此三個資料庫的結合中是否可行。
6. 找尋未知的內隱知識供專家做為後續研究。

## 1.3 研究範圍與限制

本研究將以一年份的中央氣象局的氣象觀測記錄、環保署空氣污染監測記錄、健保局的健保記錄（氣喘病部份）為研究分析對象，重點在於使用類神經網路中的 BPN 模型的理論基礎，來建構出全台的氣喘病發作可能性模型。

因此，本研究中的研究限制範圍如下：

1. 因為健保記錄僅有一年份的資料，為了維持資料的一致性，所以其它資料庫的記錄也只能取相同年度的資料，所以無法建立長程的模型。
2. 環保署的空污記錄時有“無效值/缺漏值”出現，且無法由別的記錄推估到真實的記錄值，再則由於樣本數筆數已經很少，所以不能把無效的記錄去除，本研究中僅能以平

- 均值填補這些無效值/缺漏值的部份。再則於氣象記錄方面亦有缺憾，由於氣象測站點的分佈不均，所以有些縣市沒有三等以上的氣象站，故未能提供逐日各項監測記錄，所以本研究中無法取得這些縣市的資料，僅能以臨近縣市的記錄經過插補( interpolation )手續後值模擬真實記錄值
3. 由於資料提供單位僅出售或提供測站逐日的記錄資料，所以無法得知病患當時是處在怎樣的環境中發病的，以逐日記錄來模擬逐時的變化，難免反應會比較不敏銳或失真。
  4. 由於類神經網路的 BPN 模式未能將“ 時序 ”納入考量，因而研究中考量“ 向前推 ( 數 ) 日 ”的資料記錄，以補強 BPN 模型預測的準確率。
  5. 健保記錄中，無法辨別氣喘病患是否是單純的複診領藥或是病發送醫；並且在觀察每日的記錄後，發現數值最多僅有二位數字，面對各縣人口皆有百萬以上的人口，氣喘病人數若為單純複診取藥的話，數字應遠遠大於兩位數字，所以我們假設研究中每筆記錄皆當成是病發就醫的狀況。
  6. 由於病人基本資料已由健保局加密，故無法推得病人居住地，學者王台平等[9]引用 Javagi, Rao, Thomas[64]的研究結論，指出病人選擇求診醫院的考量中第一個就是考慮離家近的；並且歸納出：有近六成的人選擇看診的醫院會找離家近的，因此本研究推測病人可能的就診醫院與居住地相去不遠。
  7. 氣喘病屬慢性病的一種，但發作起來若不即時處理，恐有生命的危險！因此當前氣喘病患身上都會隨身攜帶急救藥劑，因而降低求診的次數。

## 1.4 論文結構

本研究共分為五章，其結構如下：

第一章 緒論：說明本研究的研究動機、研究目的、研究流程及研究限制範圍。

第二章 相關研究：介紹資料萃取技術、氣象、各類空污、健保（氣喘病）及類神經網路相關領域研究。

第三章 研究方法：依據研究目的及相關研究中的理論基礎，描述研究方法及類神經網路 BPN 模型建置預測模型的步驟。

第四章 實驗結果：說明本研究的研究成果。

第五章 結論與未來研究方向：本研究結果討論及後續研究的方向說明。

以下是本研究的研究架構圖：

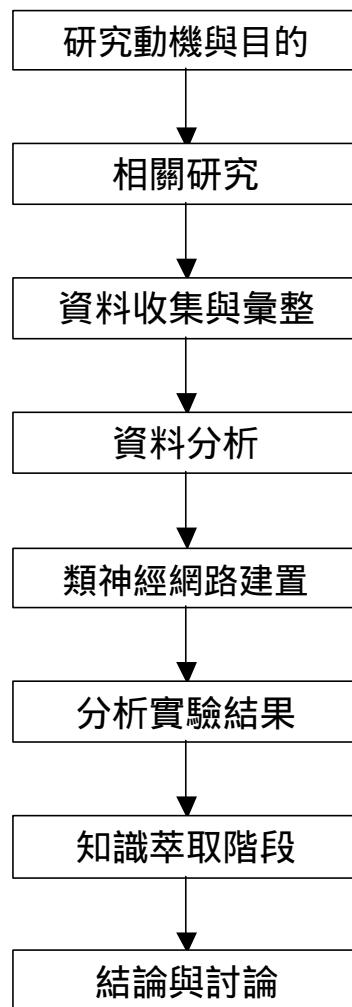


圖 1.1：研究流程

本研究架構說明如下：

1. 研究動機與目的：說明本研究的源起及預期的目標。
2. 相關研究：詳讀學者們的相關研究及氣象、空污、健保(氣喘病)三領域的背景知識。
3. 資料的收集與匯整：天氣、空污、健保記錄的取得，之後預先做前處理，然後再將三個資料庫的記錄匯進同一個資料表，以利網路模型的建置。
4. 資料分析：根據文獻探討中的相關知識，以便挑選出具影響研究模型建置的欄位屬性。
5. 類神經網路建置：本階段在於設定網路的參數值(包括層數、神經元數...)、訓練樣本的訓練過程及測試樣本的測試過程。
6. 分析實驗結果：對研究所得到的數據結果作個說明、討論。
7. 知識萃取階段：根據上一步的結果，整理出未知、內隱的知識。
8. 結論與討論：說明本研究的成果及後續研究的方向。

## 第二章相關研究

首先，先探討氣象、空污及氣喘病方面的相關研究，以說明氣象、空污與氣喘病的關係；其次，為什麼研究中需要以類神經網路來建置預測模型，再探討類神經網路的相關背景；最後探討資料萃取技術的相關研究及整理出研究設計方向。

### 2.1 氣象知識

由[11]可知到天氣氣象原始資料的欄位有：

測站號碼 時間 氣壓 氣溫 濕度 露點 雲狀 雲量 風 降水  
日照時數 能見度 天氣狀況 地面狀況 天氣及視線 地中溫度 降水紀錄 錯誤記錄。

在以往的研究報告中，發現幾個氣候與空污間的相關訊息 [12, 21, 23]：

1. 在同時考慮溫室氣體的增溫作用與氣溶膠冷卻作用的情況下，所有的氣候模式皆推估出—台灣鄰近地區的平均氣溫將持續上升。在二氧化碳增為 1.9 倍時，溫度將上升 0.8-2.4 C，四季增溫的程度無差異。
2. 溫室氣體增加的同時，大氣的懸浮微粒也會相對地增加，因為這樣的緣故也造成冷卻的作用。
3. 降水量的增加有效降低揚塵，可使得懸浮微粒大量地減少。

再則，與專家討論過後得到幾個影響的因素：

1. 若有高壓盤旋該地將使空氣污染物停滯，造成污染的加劇。
2. 空氣中的懸浮微粒可以幫助霧氣凝結，形成雨雲。



3. 飽和濕空氣中的絕對濕度是隨著溫度而有顯著變化的。溫度每上升 10 度 c，絕對濕度約可以增加近一倍[30]。
4. 日造量充足將有效地幫助光化學反應，繼而產生光化學煙霧[2]，對人體健康產生危害。

所以，本研究中經初步決策分析後，定義出天氣氣候中會對氣喘及空污的主要影響因素如下：

風速 氣溫 溫差 相對濕度 降水量 氣壓 日照率<sup>1</sup>

- 風

風速應以 m/s(米/秒)或 kn(節)為單位，並以整數報告。將風吹來的方向定義為風向，並以從地理正北起始順時鐘向量度。氣象局以風程計來計算單位時間風走的行程，若在一小時內風程計的記錄值是 60000m，則風速即為 60000m/1hr=60km/hr。

- 氣溫

氣溫係以溫度計量測之，指在距地面 1.25 - 2.00 公尺間流動，而不受太陽直達輻射影響之空氣溫度而言。

- 濕度

濕度係大氣中水份含量多少之表示。可分為絕對濕度及相對溼度。而本研究中採用『相對溼度』計量，因為對於人體的感覺最為直接。

相對濕度：即空氣中實際含有之水汽量，與相同溫度下可含最大水汽量之百分比。空氣在完全飽和狀態時相對濕度為 100%；如空氣中所含水汽量僅為當時溫度下所含最大水汽量之一半時，

---

<sup>1</sup>見 2.2 光化學反應的催化及植物花粉

則相對濕度為 50%。對人體而言，空氣之相對濕度在 40 60% 間時，令人最感舒適。

- 降水量

降水量係指在一定時間內之降水，儲積在一平面上，在無蒸發、流失或滲透等損耗情況下，其儲積量之深度謂之。

- 氣壓

何謂氣壓?氣壓乃靜止時大氣之壓力。

在地面上，氣壓即單位面積氣柱之垂直重量，亦即單位面積所受力之大小 ( $P=f/A$ )。

#### 大氣壓力的單位

國際單位制(SI)的壓力單位是 Pa ( 帕 ,  $\text{Dyne/cm}^2$  ) 或  $\text{Newton/m}^2$ 。但是這個單位太小，在氣象上用起來不甚方便。氣象上報告壓力用的單位是 hPa ( 百帕 )。1hpa 等於 100Pa，所有氣象用氣壓表的標尺應採用 hPa 為單位來刻度，這與過去常用的單位 mb (  $10^{-3}\text{bar}$  , 毫巴 ,  $1 \text{ bar} = 1 \text{ Newton/m}^2$  ) 是相等的。

以下為氣象地面觀測站的分布位置圖。



圖 2.1：氣象地面觀測站分布圖<sup>2</sup>

## 2.2 空氣污染的產生及危害

一般而言，空污可分為一次污染物、二次污染物[11]。

**一次污染物：**一次污染物是指直接從污染源排放的污染物質，如二氧化硫、一氧化氮、一氧化碳、顆粒物等。它們又可分為反應物和非反應物，前者不穩定，在大氣環境中常與其他物質發生化學反應，或者作催化劑促進其他污染物之間的反應，後者則不發生反應或反應速度緩慢。

**二次污染物：**二次污染物是指由一次污染物在大氣中互相作用經化學反應或光化學反應形成的與一次污染物的物理、化學性質完全不同的大氣污染物，其毒性比一次污染物還強。最常

<sup>2</sup> 資料來源：中央氣象局

見的二次污染物如硫酸及硫酸鹽氣溶膠、硝酸及硝酸鹽氣溶膠、臭氧、光化學氧化劑  $O_x$ ，以及許多不同壽命的活性中間物（又稱自由基），如  $HO_2$ 、 $HO$  等。

由環保署[11]所提供的 7 項污染物指標中，抽取出 5 項：一氧化碳 臭氧 氮氧化物 二氧化硫 懸浮微粒。以下將逐項介紹各污染物的成因及危害：

- **懸浮微粒 (PM10)**

係指粒徑在 10 微米以下之粒子，又稱浮游塵。主要來源包括道路揚塵、車輛排放廢氣、露天燃燒、營建施工等或由空氣污染物轉化成之二次污染物，由於粒徑小於 10 微米以下，能深入人體肺部深處，如該粒子附著其他污染物，則將加深對呼吸系統之危害。

- **二氧化硫 ( $SO_2$ )**

除自然界產生外，一般為燃料中硫份燃燒與空氣中之氧結合者，為一具刺激臭味之無色氣體，易溶於水，與水反應為亞硫酸；於空氣中可氧化成硫酸鹽，為引起酸雨的主要物質。長期吸入二氧化硫，會影響黏膜或細胞的正常功能。它常與懸浮微粒狼狽為奸，會使氣管炎、氣喘、肺氣腫的病患增加及病情惡化。

- **氮氧化物 ( $NO_x$ )**

氮氧化物主要包括一氧化氮 ( $NO$ ) 及二氧化氮 ( $NO_2$ )，其生成原因係來自燃燒過程中，空氣中氮或燃料中氮化物氧化而成，一氧化氮為無色無味氣體稍溶於水，燃燒過程生成之氮氧化物以一氧化氮為主要成份，光化學反應中可反應成二氧化氮。二氧化氮為具刺激味道之赤褐色氣體，易溶於水，與水反應為亞硝

酸及硝酸，參與光化學反應，吸收陽光後分解成一氧化氮及氧，在空氣中可氧化成硝酸鹽，亦是造成雨水酸化原因之一。

二氧化氮具有刺激性，會刺激眼、鼻及肺部，而引起氣管、炎肺炎等症狀。甚至會降低呼吸器官的抵抗力，使其容易受到感染。

一氧化氮容易與血紅素結合，影響氧氣運輸的正常功能；吸入濃度過高時，會造成體內缺氧，使人的中樞神經受損。

- 一氧化碳 (CO)

除森林火災、甲烷氧化及生物活動等自然現象產生外，主要來自石化等燃料之不完全燃料產生，無色無味，比空氣輕，由於一氧化碳對血紅素的親和力比氧氣大得多，因此，可能造成人體及動物血液和組織中氧氣過低，而產生中毒現象。

- 臭氧 (O<sub>3</sub>)

係一種由氮氧化物、反應性碳氫化合物及日光照射後產生之二次污染物。具強氧化力，對呼吸系統具刺激性，能引起咳嗽、氣喘、頭痛、疲倦及肺部之傷害，特別是對小孩、老人、病人或戶外運動者有較大影響，同時對於植物，包括農作物有不良影響，對於人造材料，諸如橡膠（輪胎等）及油漆等，均能造成危害。

以下是各縣市空氣污染測站配置狀況。

表 2.1：全台空氣品質監測站分布<sup>34</sup>

---

<sup>3</sup> 資料來源：行政院環保署

<sup>4</sup> [交通站無臭氧測值](#)，[工業站無臭氧與一氧化碳測值](#)

縣市	站名
基隆市	仁愛
台北市	士林、中山、松山、萬華、古亭、 <u>大同</u> 、陽明
台北縣	淡水、林口、菜寮、汐止、新莊、板橋、土城、新店、 <u>三重</u> 、 <u>永和</u> 、萬里
桃園縣	桃園、大園、五權、龍潭、 <u>中壢</u> 、觀音
新竹市	新竹
新竹縣	湖口、竹東
苗栗縣	苗栗、 <u>頭份</u> 、三義
台中市	西屯、忠明
台中縣	豐原、沙鹿、大里
彰化縣	彰化、二林、 <u>線西</u>
南投縣	南投、埔里、竹山
雲林縣	崙背、斗六、台西
嘉義市	嘉義
嘉義縣	新港、朴子
台南市	安南、後甲
台南縣	新營、善化
高雄市	楠梓、左營、三民、前金、小港、 <u>復興</u> 、 <u>前鎮</u>
高雄縣	美濃、仁武、鳳山、大寮、林園橋頭
屏東縣	屏東、潮州、恒春
宜蘭縣	宜蘭、冬山
花蓮縣	花蓮
台東縣	台東

## 2.3 健保及氣喘病

在過去的研究中指出[6]，對高屏地區 11-16 歲國中學生作調查，分析室內外環境因素對於氣喘的相關性，在室外的環境因素以空氣污染對氣喘有顯著的相關。在進一步校正干擾因子後，發現以懸浮微粒、二氧化氮、一氧化碳、臭氧的相關性最為顯著。並且 Warlaw [68]的研究中也指出：

1. 臭氧及二氧化硫可能是氣管收縮劑，會刺激呼吸道做直接的反應。
2. 吸入低濃度的污染物會引起呼吸道發炎，高濃度的污染物則是直接對呼吸道上皮造成毒性的作用。
3. 空氣污染會改變免疫系統的反應，因此易引起氣喘的過敏反應。

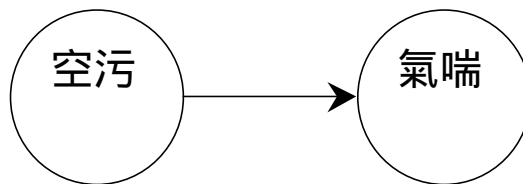


圖 2.2：空氣污染與氣喘的關連性

然而，引起氣喘病發作因素不僅於天氣污染，還有氣象的因素..等等，以下列舉幾項與天氣變化有關係的做為參考[8, 47]：

1. 春夏季傍晚間氣喘會加重，這是因為與病人吸入花粉有關。因為空氣中的花粉經過白天的曝曬後隨風飄散，晚間的氣溫下降而使得氣流向下，花粉也因此下沉，此時花粉在空氣中的濃度高，所以病人發病的機率就大。
2. 秋冬季天氣變化無常，病人吸入冷空氣，會讓呼吸道急遽收縮，引起氣喘症狀。
3. 夏季溫度升高，再加上濕度加大，使得細菌也容易繁殖，吸入後容易讓呼吸道產生炎症，導致氣喘的發作。

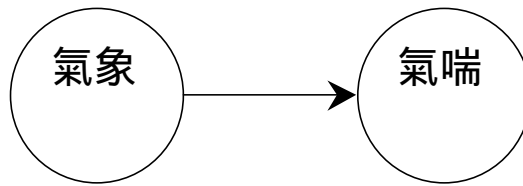


圖 2.3：氣象與氣喘的關連性

而空氣的污染也讓我們居住的環境遭受到破壞，比如：酸雨 [22, 12, 23] 是由二氧化硫及二氧化氮所造成；二氧化碳 [11, 30] 則是造成溫室效應的主要原因，進而使得全球都遭到熱浪的襲擊，並且因為氣候改變造成沙漠化的形成 [53]；再則是失去臭氧層 [27] 的保護，將使地球生物圈暴露在更多的輻射線下，可能會造成：皮膚癌、白內障等疾病罹患率的增加。經由以上的說明，可以知道天氣、空氣污染對氣喘病的發作存在相依性，於是便可以藉著天氣氣象、空氣污染與健保的記錄提供作預測分析，了解相關要素的變化 [1, 2]。

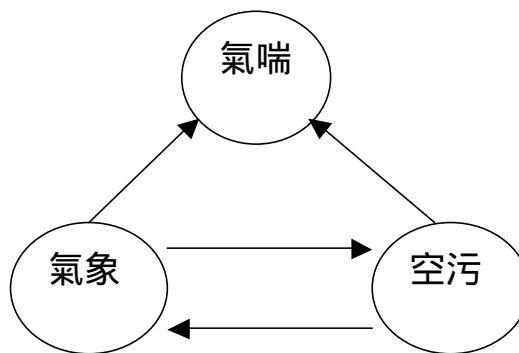


圖 2.4：氣象、空污、氣喘病間關聯性

在健保局 [43] 的健保資料庫中的「門診處方及治療明細檔」裡面一共有 37 個欄位，從中抽取出跟本研究有關的欄位，分別是：醫事機構代碼 (再跟醫事機構基本檔做串檔動作後，可以找出該醫事機構位於哪個縣市) 就醫日期 國際疾病分類號。



表 2.2：「門診處方及治療明細檔」說明<sup>5</sup>

序號	中文欄位	長度	序號	中文欄位	長度
1	費用年月	6	20	國際疾病分類號二	5
2	申報類別	1	21	國際疾病分類號三	5
3	醫事機構代號	10	22	主手術代碼	4
4	申報日期	8	23	給藥日份	2
5	案件分類	2	24	處方調劑方式	1
6	流水號	6	25	醫師代號	10
7	特定治療項目代號 (一)	2	26	藥師代號	10
8	特定治療項目代號 (二)	2	27	用藥明細金額小計	8
9	特定治療項目代號 (三)	2	28	診療明細金額小計	8
10	特定治療項目代號 (四)	2	29	診察費項目代號	12
11	就醫科別	2	30	診察費	8
12	就醫日期	8	31	藥事服務費項目代號	12
13	治療結束日期	8	32	藥事服務費	8
14	出生日期	8	33	DRG 參考碼	2
15	身份證統一編號	10	34	合計金額	8
16	健保卡就醫序號	2	35	部分負擔金額	8
17	給付類別	1	36	申請金額	8
18	部分負擔代號	3	37	性別	1
19	國際疾病分類號一	5			

<sup>5</sup> 資料來源：健保局

表 2.3：「醫事機構基本主檔」說明<sup>6</sup>

序號	中文欄位	長度	序號	中文欄位	長度
1	醫事機構代號	10	15	原評鑑等級生效日	8
2	特約類別	1	16	縣市區碼	4
3	特約類別起日	8	17	負責醫師 ID	10
4	特約類別迄日	8	18	執照核發日	8
5	型態別	2	19	特約狀況碼	1
6	型態別起日	8	20	合約起日	8
7	型態別迄日	8	21	合約訖日	8
8	教學成本註記	1	22	續約起日	8
9	教學成本起日	8	23	續約迄日	8
10	教學成本迄日	8	24	停約起日	8
11	評鑑等級	2	25	停約迄日	8
12	評鑑等級起日	8	26	休診起日	8
13	評鑑等級迄日	8	27	休診迄日	8
14	原評鑑等級	2	28	舊醫事機構	10

參照以上三個資料庫的區域劃分，空氣品質觀測是將各個縣市都區分開，以下有數個不等的自動測站；健保記錄也是以縣市做為基本單位，之後再細分到各鄉鎮市鎮；而氣象資料的部份，雖然全省各地自動測站（四等以上的）加起來有 24 站，但分布位置沒有平均散落各地，在台北的附近就設了台北、鞍部、竹子湖、淡水等 4 站，而桃園、雲林、彰化等地卻沒有測站的設置；所以，本研究中將以縣市為劃分的單位，以利進行資料庫的整合。

<sup>6</sup> 資料來源：健保局

根據先前三小節的介紹，本研究發現要釐清氣象、空氣污染因子對氣喘病的影響性，非僅僅利用統計/分析方法就能得到答案的，其困難點在於：

1. 氣象與空污間具有相互影響關係。
2. 目前無法以線性方程式表明其間的影響程度。
3. 地域的條件限制，每個縣市的氣喘病模型不一定相同。

並且，類神經網路相對於統計方法有以下的優點[45]：

1. 可以建構非線性的模型，其準確性較佳。
2. 可以表達輸入變數間的交互作用。
3. 可以接受邏輯、數值……等等的變數做為輸入，適應性強。
4. 可以用於函數映射、數列預測、樣本分類等問題，應用範圍廣。

因此研究中需要仰賴類神經網路建立起氣喘病的預測模型，因為它可以藉著訓練中權值的調整，找出三者間的相關程度以便趨近真實的狀況。

## 2.4 類神經網路

以往類神經網路在各領域有著許多的應用，特別是在機器學習及預測[58, 59]上有著顯著的成效。而類神經網路的架構一般而言有三種[3]：迴歸式網路(recurrent net)、前授式網路(feed-forward net)、動態式網路(dynamic net)。

### 1. 迴歸型網路

迴歸式網路中的神經元彼此相連接，對於每個神經元而言，它的輸出連接至所有的其它神經元，而它的輸入來自所有其它的神經元輸出，如圖 2.2 所示。

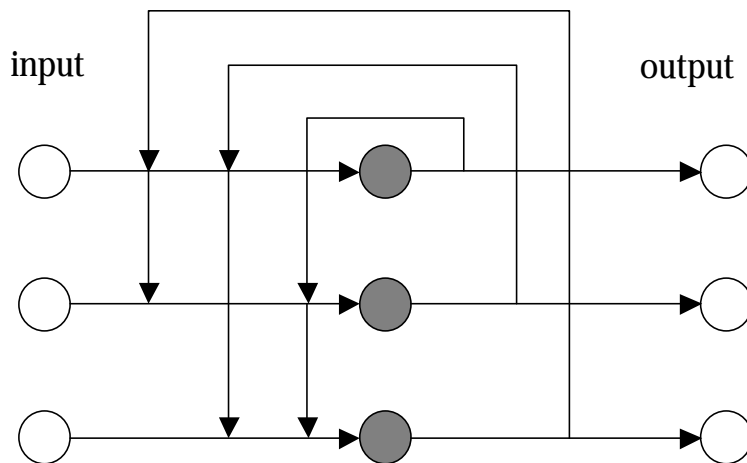


圖 2.2：迴歸式網路  
(資料來源：本研究整理)

## 2. 前授式網路

前授式網路屬於一種階層式的網路，包括了三個層次：輸入層、隱藏層、輸出層，如圖 2.3 所示。

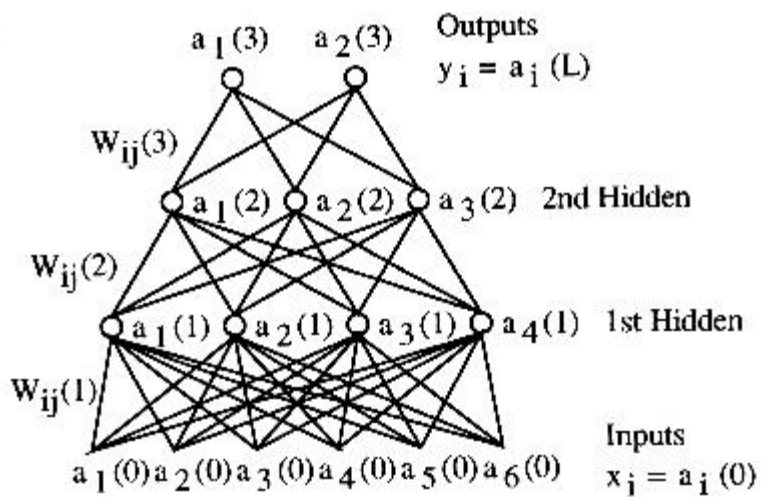


圖 2.3：前授式網路<sup>7</sup>

<sup>7</sup> 資料來源：類神經網路與模糊控制理論入門-王進德、蕭大全

每一層皆由數個神經元組成，而同層間的神經元並不相互連接，不同層的神經元則可以連接。信號的傳遞僅為單方向，由輸入層一步步傳送到輸出層，也就是訊號從最下層輸入，經過多層神經元的傳導，最後到達輸出層。而本類的網路，最具代表性的就是倒退式網路（Back-propagation Neural Network）。

### 3. 動態式網路

類神經網路的架構是為了能夠完全的學習，如果設計的架構與問題本身不相合，則無法正確學習。當類神經網路模型太過簡單時，就無法學習到所有的輸入及輸出的映對；即便加入大量的節點或是連接權值，也依然可能效果不好。所以此時就可以使用動態式的類神經網路架構。

這類的類神經網路一開始僅有輸入層及輸出層。當單層網路無法良好地學習時，就會增加一個隱藏節點（hidden node）。這個節點會自動連接所有的輸入神經元後開始學習，學習完後就如同之前的神經元一樣會與輸出層相連接。如果效果仍是不佳，將會一直重覆先前的動作直到學習效果良好為止。如圖 2.4 所示。

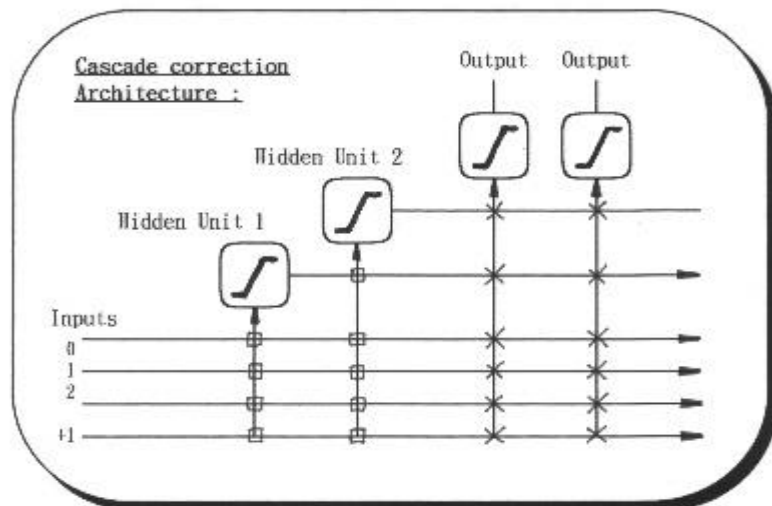


圖 2.4：動態式網路<sup>8</sup>

<sup>8</sup> 資料來源：類神經網路與模糊控制理論入門-王進德、蕭大全

再則，一般而言，類神經網路會有下列幾項特性：

1.

類神經網路採用大量的平行計算，經由許多不同的神經元作運算處理，這樣的平行處理主要是引導電腦運算程序可轉化成指令，儲存在記憶體中，再按照指令逐一執行的線性過程。

2. 錯誤容許度

倘若資料中摻雜著些許的雜訊干擾，仍不會影響類神經網路運作的正確性，即使有部份神經元失去效用，類神經網路仍能有效運作。

3.

在迴歸式類神經網路中沒有資料記憶區，但網路卻可以記住訓練範例，並且爾後若須對輸入的資料作運算，整個網路可以藉運算程序來聯想相對應的輸出值。

4. 解決最佳化的問題

我們可以在問題領域中，找到一組設計變數值，使其在滿足設計限制下，使整個設計目標達到最佳化的狀態。

而網路間各層的運作方式可分為「學習過程」及「回想過程」[45, 32]。

1. 學習過程：

依照學習的習性來區分，可以分做監督式學習、無監督式學習及聯想式學習，這三種學習過程皆是從「能量函數」推導出來的，而衡量網路學習效果的指標函數則稱為「誤差函數」，透過網路學習可使「誤差函數」的值達到最小化。

A. 監督式學習：

其學習的過程是為了降低網路輸出的目標值與預測值間的差異，其「誤差函數」通常定義如下：

$$E = \frac{1}{2} \sum_i (T_j - A_j)^2 \dots\dots\dots(2.1)$$

其中：

$T_j$  = 網路輸出目標值

$A_j$  = 網路輸出推論值

B. 無監督式學習：

此學習過程中的每一個輸入單元皆代表一個叢聚，而學習的目的在於使網路中優勝單元的連接權值與輸入向量間的距離最小化，使得每一個輸出單元的連結權值可以代表輸入範例在樣本空間中的叢聚形心，其「誤差函數」通常定義如下：

$$E = \frac{1}{2} \sum_i (X_j - W_{ij})^2 \dots\dots\dots(2.2)$$

其中：

$X_j$  = 網路輸入向量

$W_{ij}$  = 優勝單元的連接加權值

C. 聯想式學習：

其過程是以處理單元的組成表示一個樣式，其學習過程在於使經過「聯想」疊代所得的最終狀態變數與網路記憶的樣式之一相同或近似。

2. 回想過程：

回想過程為網路經由學習過程後，輸入變數經網路計算至輸出的過程，其與學習過程不同點在於計算過程中不再使用「誤差

函數」，也不再對加權值加以修正，只是單純地對網路執行輸入及輸出的工作。

而本研究中為何會挑選 BPN 模型做為建置預測模型的本體，主要是考量幾點因素：

1. 無法以線性方程式表明其間的關聯性。
2. 資料中攙雜些許的雜訊，即無效值與缺漏值。
3. 需要聯想記憶。
4. 本研究的預測模型需要根據“每日的氣喘發病可能性”，做監督式的學習。
5. 目前相關領域的研究中，發現以 BPN 模型的效果最好。

所以，在研究中選擇以類神經網路的 BPN 模型來做為氣喘病發作預測模型的本體。以下將介紹本研究中找尋內隱、未知知識的方法 - 資料萃取技術。

## 2.5 資料萃取技術

### ● 源起

在外部環境的變遷下，使得資訊科技的發展一日千里，於是企業需要重新制定新的經營策略，來改善經營模式，期望能為企業帶來競爭優勢。因而，引進最新的資訊技術，替企業創造及儲藏大量的交易資料，然而面對如此眾多的資料時，卻遇到資料分析的瓶頸--無法有效的由資料中發掘重要的資訊，因此資料萃取（Data Mining，DM）便孕育而生。它是一項利用電腦快速計算的能力，可以從大量的資料中不斷探索，以發掘隱藏在繁雜資料中的知識與規則的過程。

資料萃取的工作，是近年來資料庫應用領域中相當熱門的議題。不過，它的起源可追溯至二次世界大戰前，美國政府在人口



普查以及軍事方面使用這些技術。並且資料萃取所使用的分析方法，如預測模型（迴歸、時間數列）、資料庫分割（Database Segmentation）、連接分析（Link Analysis）、偏差偵測（Deviation Detection）等，都是在 1970 年代就已經開始研究了。但是資訊科技的進展超乎想像，新工具的出現，例如關聯式資料庫、物件導向資料庫、柔性計算理論（包括 Neural networks、Fuzzy theories、Genetic Algorithms、Rough Sets 等）、人工智慧的應用（如知識工程、專家系統），以及網路通訊技術的發展，使專家可以從資料堆中挖掘到寶藏，因此資料萃取技術成為不可或缺的一環。

- 定義

根據先前學者的研究對 Data Mining 的定義，本研究彙整成表格，如下表所示。

表 2.4：學者對 Data Mining 的定義

學者姓名	對 Data Mining 所下的定義	出處
朱慧德 周卓定	從資料中發掘資訊或知識。利用資訊科技的輔助分析資料以得到有用的資訊，並建立模擬真實世界的模型，並利用模型描繪出資料的特徵與關係。	[10]
林文修	從資料中發掘有用的資訊及知識（有人稱為 Knowledge Discovery in Database, KDD），也有人稱為『資料考古學』	[26]
Cabena Hadjinian Stadler Verhees Zamasi	將先前是未知的，有效的資訊從龐大的資料庫中萃取出來的程序，並且將萃取出的資訊提供給主官（管）做為決策的輔助。	[60]
Linoff	面對大量的資料，使用自動或半自動的方法進行分析動作，藉以發掘有意義性的關係或是法則	[65]

續表 2.4：學者對 Data Mining 的定義（2）

Grupe Owring	對於現在有的資料找出新的事實及發掘出專家尚未發覺的新關聯性	[65]
Yongjian	在大型資料庫中，辯視出有意義的樣式之過程	[65]
Fayyad Smyth	屬於 KDD 的一個過程，可以用於某些計算技術上，在容許的運算效率限制下，產生些可以建立於資料之上的特定樣式。	[65]
林信忠	其主體由資料（庫）、相關技術所構成的，目標在於發掘新的、未知的法則或樣式	[24]
張瑋倫	將資料庫中看似無用的資料轉化成有用的資訊，並找出隱藏的關聯法則。	[33]
李紹倫 陳富中	利用分類、關聯性、群聚分析、序列分析、行走路徑分析、機器學習及統計方法，從雜亂且資料量龐大的資料庫中找出隱藏的、未知的、卻又對企業經營非常有用的資訊。	[14]
Akeel Al-Attar	可以讓複雜的企業程序成為易瞭解且可重新設計，並發掘出過去營運模式中一些有意義的樣式，而藉著發展這些模式，就可以用來改善經營績效及避免企業經營危機。	[61]
郭展盛	需要特別的演算法輔助，才能抽取出特定型式資料集中有意義的樣式或規則。	[28]
Chung Gray Mannino	認為知識發掘的重點是在資料中發掘出有用的知識，而資料萃取主要著重在一些為了從資料截取知識所用的演算法應用上面，因此知識發掘較具學術研究的意味,資料萃取則為其操作過程。	[3]

綜合前述學者對資料萃取的定義可知，資料萃取主體由資料或資料庫、相關技術（偏重於資訊科技）所構成，其目標則是為

了發掘出新的、未知的規則或樣式。這些樣式有兩種用處；第一，瞭解資料的特徵與關係可以提供做決策時所需要的資訊。第二，資料的特徵可以幫助做預測。

所以，先前提到的類神經網路是一項可以符合資料萃取技術要求的技術，因為它不但可以協助發掘新的、未知的規則或樣式，並且可以依資料的特徵來建置預測的模型。

- 特性

在[24]將資料庫知識發掘分為“資料選擇”、“資料前處理”與“轉換”、“資料萃取”、“詮釋/評估”這四個階段。

1. 資料選擇：主要目的在於建立目標資料集（target data set），經過考慮後決定需要萃取何種資料？
2. 資料前處理與轉換：資料萃取出來後，對於不適用的資料（比如：重覆或 null）預先作處理，再將資料轉換成工具能讀取的資料格式。
3. 資料萃取：藉由資料萃取軟體找出隱藏在資料背後的規則、特性、樣式。
4. 詮釋/評估：分析與解釋實驗中產生的結果。

並且“資料萃取”這個階段還可以建立六種模式[14，25，56]：**Classification**（分類）、**Regression**（迴歸）、**Time Series**（時序）、**Clustering**（分群）、**Association**（關聯）以及 **Sequence**（循序）。

針對這些模式說明如下：

1. **Classification**：是根據一些變數的數值做計算，再依照結果作分類。常被用來處理對象篩選的問題，我們會用一些

已經分類的資料來研究它們的特徵，然後再根據這些特徵對其他未經分類或是新的資料做預測，這些用來尋找特徵的已分類資料可能是來自現有的歷史性資料，或是將一個完整資料庫做部份取樣，再經由實際的運作來測試。

2. **Regression**：是使用一系列的現有數值經過統計與分析後，來預測一個連續數值的可能值。
3. **Time-Series**：Time-Series 與 Regression 很像，不同點在於它所分析的數值都與時間有關，利用現有的數值來預測未來的數值。Time-Series 的工具可以處理有關時間的一些特性，例如時間的階層性、季節性、節日、以及其它的一些特別因素，如過去與未來的關聯性有多少。
4. **Clustering**：將資料分為幾組，其目的是要將組與組之間的差異找出來，同時也要將一個組之中的成員的相似性找出來。
5. **Association**：是要找出在某一事件或是資料中會同時出現的東西，也就是如果 A 是某一事件的一部份，則 B 也出現在該事件中的機率有 X %。
6. **Sequence**：與 Association 很類似，所不同的是 Sequence 中有關連的資料是以時間來區分先後關係，例如：如果做了 X 手術，則 Y 病菌在手術後感染的機率是 45%。

在[26]中提到 Glymour 等人的研究，建議資料萃取過程的進行步驟：

1. 理解資料與進行的工作；
2. 獲取相關知識與技術（Acquisition）；
3. 融合與查核資料（Integration and checking）；
4. 去除錯誤或不一致的資料（Data cleaning）；
5. 發展模式與假設（Model and hypothesis development）；
6. 實際資料萃取工作；

7. 測試與檢核所挖掘的資料 ( Testing and verification ) ；
8. 解釋與使用資料 ( Interpretation and use ) 。

從八個步驟來看，資料萃取牽涉大量的規劃與準備，而從其他文獻得知，專家聲稱高達 80% 的過程花在準備資料階段，資料萃取只是知識發掘過程中的一個步驟而已，而達到這個步驟前還有許許多多的工作要完成。

### ● 應用

資料萃取技術應用的領域十分廣泛，在行銷、醫學、製造、財務金融與行為偵測 ( behavior detection ) 方面已有莫大的貢獻。特別是在行為偵測方面，IBM 公司利用模糊模式和統計技術，來分析醫療業詐欺索賠的案件，而澳洲醫療保險機構則利用類神經網路和統計方式，協助找出詐欺與濫用醫療資源的案件，執行的結果十分良好[63]。

在學者陳文地[37]的研究中提到，在資料庫中發現知識即是現在所知的資料萃取資料庫中可以發現有用的特徵 ( patterns ) 和規則 ( rules ) 。而這些特徵及規則在決定組織的作為時是極有用的。因此近來資料萃取已在各種不同領域的專家中獲得極大的注意，如機器學習 ( machine learning ) 、統計學 ( statistics ) 、人工智慧 ( artificial intelligence ) 和專家系統 ( expert system ) 等。

英國 Safeway 公司使用 Intelligent Miner 從資料庫中萃取商業知識。這家公司兩要知道哪些種類的客戶買了哪些種類的產品比及購買的頻率，以建立 “ 個人導向的市場 ” 。而透過 Intelligent Miner ，他們找出超過人類概念範圍的關聯性。例如：某一種乳酪產品的銷受量是排行第 209 名，但是消費額最高的客戶中有 25% 都常購買這項產品，若以傳統的分析方法的話，一定會把這項產

品停賣，但事實上這項產品是相當重要的。由此可知，透過資料萃取可以發掘出許多以往分析方法所未能發現的重要事實。

在學者魏志平、張勳騰、李彥賢[57]的研究中提及，由於資訊科技的進步，企業收集累積大量的交易記錄。傳統上，企業將這些資料當成是過去歷史記錄來應用，並未善加利用這些企業資料；事實上，這些資料的價值不僅是過去歷史記錄的呈現，更重要的是這些資料記錄了消費者的決策結果。因此，若能由這些資料中，找出顯著且有效的消費者行為或模式，將能賦予企業更多的經營智慧，而此一過程稱之為資料萃取。其目的是從大量資料中，尋找事前未知、有效且可以付諸行動的規則或知識，以輔助企業決策、提升決策品質。

目前企業界把資料萃取應用在許多領域。例如，行銷、財務、銀行、製造廠、通訊等。並且產學合作下，發展出許多實用的系統。

在[26]記錄了一個應用在行銷領域的例子：經由記錄客戶的消費記錄與採購路線，超級市場可以設計出更吸引顧客購買的環境。根據資料萃取的特別資訊，因此現在超級市場的廚房用品，是按照女性的視線高度來擺放。根據研究指出：美國婦女的視線高度是 150 公分左右，男性是 163 公分左右，而最舒適的視線角度是視線高度以下 15 度左右，所以最好的貨品陳列位置是在 130 至 135 公分之間。

以下是過去學者將資料萃取應用於各領域的彙整，本研究並加註該研究使用資料萃取中何種模式（分類：1，迴歸：2，時序：3，分群：4，關聯：5，循序：6）。

表 2.5：學者將資料萃取技術應用於各領域

應用領域	以往學者研究	所屬模式	備註
健康保險	「資料萃取法在健保費用查核之研究」	1、2、4	[24]
健康保險	「結合模糊集合理論與貝氏分類法之資料探勘技術」	1、2、5	[56]
健康保險	「健保醫療費用審查自動化之研究」	1、5	[7]
醫療	「應用資料採礦於基因體之重複序列資料庫」	1、5	[16]
醫療	「跨平台資料探勘工具之設計與建立」	1、2	[20]
醫療	「資料探索在醫學資料庫之應用」	1、2、5	[13]
醫療	「以資料採礦技術發展臨床路徑之研究」	5、6	[19]
醫療	「運用資料挖掘技術預測救護車服務量」	1、2	[36]
關係行銷	「資料挖掘與顧客關係管理之研究」	2、4	[33]
關係行銷	「以資料發掘經由知識管理建構顧客關係管理模式之研究」	2、4、5	[25]
目標行銷	「資料挖掘之關聯性法則架構以零售業目標行銷為例」	1、2、5	[15]
學習系統	「網路資訊與學習系統之中文全文探勘工具」	1、2、5	[29]
學習系統	「以資料發掘技術規劃技職校院課程查詢網站之研究」	5	[42]

續表 2.5：學者將資料萃取技術應用於各領域（2）

學習系統	「以資料發掘分析政大通識課程」	4、5	[49]
學習系統	「資料挖掘應用於找尋瀏覽網頁之型樣」	5	[46]
學習系統	「在全球資訊網上有效挖掘瀏覽路徑之研究」	1、5	[17]
機器學習	「發掘時間序列資料之知識」	1、2、3	[4]
影像處理	「以資料挖礦法挖掘多屬性序列式資料規則之研究」	1、5	[39]
電子商務	「資料提煉技術在虛擬商店之應用」	4、5	[48]
金融	「時間序列資料之查詢與資料發掘—以台灣股市為例」	1、2、3	[5]
金融	「資料探勘應用於股市股價趨勢之研究」	1、5	[55]
銀行	「資料發掘在信用卡之應用」	1、2、5	[14]
銀行	「應用資料挖掘技術研究銀行顧客消費行為」	4、5	[50]
電信業	「通信資料庫之資料探勘：目標行銷之應用」	4、5	[57]
警政	「網路犯罪模式分析及偵防對策之研究」	4、5	[38]
警政	「資料探勘應用於犯罪分析模式建構之研究」	2、4	[37]
企業	「資料探勘之應用—會員消費特徵之發掘」	4、5	[41]



續表 2.5：學者將資料萃取技術應用於各領域（3）

企業	「在 DAG 中挖掘家族特徵規則」	3、5、6	[54]
電機工程	「偵錯與診斷之設計與應用」	1、6	[35]
音樂	「音樂資料庫中重複及循序特徵探勘之研究」	1、5、6	[28]
地政	「空間資料萃取模式之建立—以土地估價為例」	1、5	[31]

手邊整理到的文獻有三篇是健保與資料萃取相結合的研究，恰好都是拿健保費用的查核來做研究的。林信忠[24]藉類神經網路的 BPN 模式、迴歸分析、區別分析等三種方法；藍中賢[56]藉貝式分類法；王復中[7]藉分類及關聯法則等三種方法，都是想找出申報費用是否有浮報的狀況。在三篇論文中，雖然使用的模式不同，但都能得到顯著的成效，故證明資料萃取技術在此能發揮效用。

並且，在以上的相關研究中發現使用類神經網路 - BPN 模型的比例相當高，也都有相當好的成效，所以本研究中選擇以類神經網路的 BPN 模式是其來有自。

# 第三章研究方法

## 3.1 BPN 模型說明

本研究中所使用類神經網路中的 BPN 模型[18, 32, 34]，是一種具有學習能力的多層式前授網路，該網路是由 Rumelhart、McClelland 在 1985 年所創立的。BPN 的運作分成“學習”與“回想”兩個部份；學習過程中是以監督式學習的方式，將問題領域中得到訓練樣式及目標輸出，再將其導入網路中，反覆調整權重值；回想過程是分類跟預測過程，可以讓網路告訴我們最可能的輸出結果[3]。利用資料萃取技術在大型資料庫中定義其演算法與分類規則的項目[66, 67]，並切以類神經網路的探勘技術，監督式學習的特性，提供輸入/出的相關特徵因子，在動態或相互抑制的神經網路下，得到有相關的神經元。之後再加以分類/預測，即可得到快速一套預測分類/預測的模型系統。

以下將對於研究主題的類神經網路 BPN 模式做細部的描述。BPN 網路架構如下圖所示，包括：

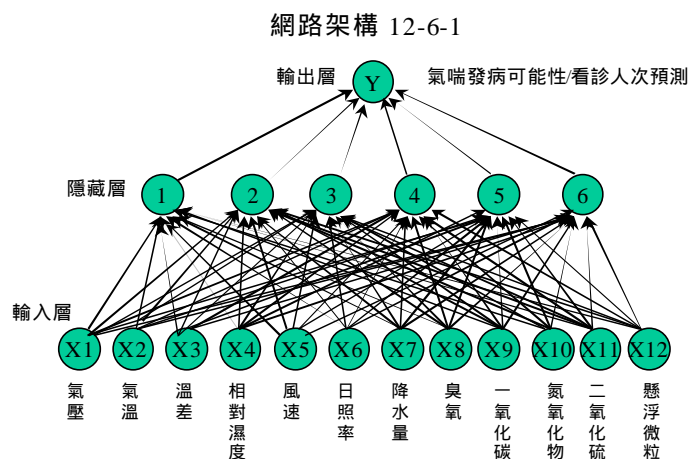


圖 3.1：BPN 網路架構

**輸入層(Input Layer)：**用以展現網路的輸入變數，處理單元的數量依問題而定，非固定模式。然而不同的輸入變數間有時因為尺度的差異過大，而影響網路的學習效果，所以在輸入時必須進行正規化的步驟，使輸入值落於[0,1]。

**隱藏層(Hidden Layer)：**用以展現輸入處理單元間的交互作用，其處理單元的數量亦無標準方法可以決定，所以常以試誤法 (try error) 來決定最佳的數量。隱藏層的層數可以不只一層，亦可無隱藏層。然而在增加處理單元與層數的優先次序上，應以先增加處理單元數目再考慮增加層數，因為有時增加層數會讓網路過於複雜，使得學習過程中收斂於局部最小值。

**輸出層(Output Layer)：**用以展現網路的輸出變數，處理單元的數量也是依問題而定。此層的主要工作為落於[0,1]間的值經過適當的轉換成為有意義的輸出值。而 BPN 模式的可貴就在於隱藏層，它可以表現輸入變數間的交互作用，其對於複雜的非線性系統能有較佳的模擬能力，且在轉換函數方面，BPN 使用平滑可微分的函數，所以可應用如攀峰法之類的演算法推導出網路加權值，以大幅提昇 BPN 的模擬能力。

令  $X = (c_1, c_2, \dots, c_i)$ ，在本研究中  $i=12$ ， $c_1 \sim c_7$  為氣象變數， $c_8 \sim c_{12}$  為空污變數 ( 時序版中  $c_{1+n} \sim c_{12+n}$ ， $n=0 \sim 3$ ，如： $c_1$  是原始值， $c_2$  是考慮前數日變化的最大值， $c_3$  是考慮前數日變化的最小值， $c_4$  是前數日變化的差值 <最大值 - 最小值> )。Y 為輸出變數。

處理單元其輸出值與輸入值的關係式，一般可用輸入值的加權乘積和函數來表示：

$$Y_j = f(\text{net}_j) \dots \dots \dots (3-1)$$

$$\text{net}_j = \sum W_{ij} X_i - \theta_j \dots \dots \dots (3-2)$$

其中

$Y_j$ =輸出變數，是模型模仿生物神經元的運作所產生的輸出訊號。

$f$ =轉換函數，是模型模仿生物神經元的非線性處理機能，是可以從其它處理單元輸入的輸入值的加權相乘積和，轉換成處理單元輸出的數學式。

$W_{ij}$ =連結加權值，模型模仿生物神經元的突觸強度。

$X_i$ =輸入變數，模型模仿生物神經元的輸入訊號。

$\theta_j$ =門限值(閾值)，模型模仿生物神經元的閾值，又稱偏權值。

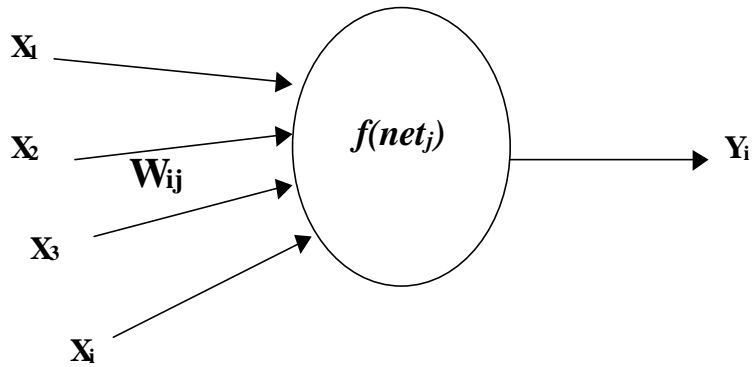


圖 3.2：人工神經元圖<sup>9</sup>

### 3.2 BPN 演算法

以下將介紹 BPN 演算法的推導[45]：

若  $net_j^n$  定義為第  $n$  層之第  $j$  個處理單元的集合函數值，其定義式如下：

$$net_j^n = \sum_i (W_{ij} A_i^{n-1} - q_j) \dots \dots \dots (3-3)$$

其中：

---

<sup>9</sup> 資料來源：應用類神經網路-葉怡成

$W_{ij}$  = 第 n-1 層第 i 個處理單元與第 n 層第 j 個處理單元間的加權值。

$A_j^n$  = 第 n 層第 j 個處理單元的輸出值，而且

$$A_j^n = f(\text{net}_j^n) \dots \dots \dots (3-4)$$

其中：

$q_j$  = 計算集合函數值，經由亂數產生的偏權值。

$f$  = 轉換函數。

$E$  = 輸出層的誤差函數值，且

$$E = \frac{1}{2} \sum_i (T_j - A_j)^2 \dots \dots \dots (3-5)$$

其中：

$T_j$  = 輸出層第 j 個處理單元的目標輸出值

$h$  = 學習速率

網路的學習過程是為了讓誤差函數能達到最小化的過程，以最陡坡降法(Steepest Descent Method)使誤差函數最小化，其權值的調整幅度與誤差函數對該權值的敏感程度成正比，也就是：

$$\Delta W_{ij} = -h \frac{\partial E}{\partial W_{ij}} \dots \dots \dots (3-6)$$

利用連鎖率可得

$$\frac{\partial E}{\partial W_{ij}} = \left( \frac{\partial E}{\partial A_j^n} \right) \left( \frac{\partial A_j^n}{\partial \text{net}_j^n} \right) \left( \frac{\partial \text{net}_j^n}{\partial W_{ij}} \right) \dots \dots \dots (3-7)$$

(3-6) 式中的第一項可以分做兩種可能來做討論

1. 第 n 層為網路的輸出層

將(3-5)式  $E = \frac{1}{2} \sum_i (T_j - A_j)^2$  代入可得

$$\frac{\partial E}{\partial A_j^n} = \frac{\partial}{\partial A_j^n} \left( \frac{1}{2} \sum_j (T_j - A_j)^2 \right) = -(T_j - A_j^n) \dots \dots \dots (3-8)$$

2. 第 n 層為網路的隱藏層之一

利用連鎖率可得

$$\frac{\partial E}{\partial A_j^n} = \sum_k \left( \frac{\partial E}{\partial net_k^{n+1}} \right) \left( \frac{\partial net_k^{n+1}}{\partial A_j^n} \right) \dots\dots\dots(3-9)$$

$$\text{定義} \left( \frac{\partial E}{\partial net_k^{n+1}} \right) = -\mathbf{d}_k^n \dots\dots\dots(3-10)$$

將(3-3)式代入(3-15)式第 2 項中可得

$$\left( \frac{\partial net_k^{n+1}}{\partial A_j^n} \right) = W_{ij} \dots\dots\dots(3-11)$$

再將(3-10)式及(3-11)式代入(3-9)式中可得

$$\frac{\partial E}{\partial A_j^n} = - \sum_k \mathbf{d}_k^n W_{ij} \dots\dots\dots(3-12)$$

接著把(3-8)式中的第 2 項以(3-4)式代入可得

$$\left( \frac{\partial A_j^n}{\partial net_j^n} \right) = \frac{\partial}{\partial net_j^n} f(net_j^n) = f'(net_j^n) \dots\dots\dots(3-13)$$

(3-8)式中的第 3 項將(3-3)代入可得

$$\frac{\partial net_j^n}{\partial W_{ij}} = \frac{\partial}{\partial W_{ij}} \left( \sum_i (W_{ij} A_i^{n-1} - \mathbf{q}_j) \right) = A_i^{n-1} \dots\dots\dots(3-14)$$

以上之推導過程中總結可以得到兩種不同的情況：

1.  $W_{ij}$  處在隱藏層與輸出層之間

將(3-8)、(3-11)及(3-14)式代入(3-7)可得

$$\frac{\partial E}{\partial W_{ij}} = -(T_j - A_j^n) f'(net_j^n) A_i^{n-1} \dots\dots\dots(3-15)$$

又將(3-10) (3-14)代入(3-7)式可得

$$\frac{\partial E}{\partial W_{ij}} = -\mathbf{d}_j^n A_i^{n-1} \dots\dots\dots(3-16)$$

比較(3-15)及(3-16)式得知

$$\mathbf{d}_j^n = (T_j - A_j^n) f'(net_j^n) \dots\dots\dots(3-17)$$

2.  $W_{ij}$  處在隱藏層與隱藏層之間

將(3-12)、(3-13)及(3-14)代入(3-7)式中可得

$$\frac{\partial E}{\partial W_{ij}} = -\mathbf{d}_j^n A_i^{n-1} \dots\dots\dots (3-18)$$

比較(3-18)及(3-15)式得知

$$\mathbf{d}_j^n = \left( \sum_k \mathbf{d}_k^{n+1} W_{kj} \right) f'(net_j^n) \dots\dots\dots (3-19)$$

從以上的推導可知  $\frac{\partial E}{\partial W_{ij}}$  的通式為

$$\frac{\partial E}{\partial W_{ij}} = -\mathbf{d}_j^n A_i^{n-1} \dots\dots\dots (3-20)$$

其中如果  $W_{ij}$  處於隱藏層與輸出層之間

$$\mathbf{d}_j^n = (T_j - A_j^n) f'(net_j^n)$$

否則

$$\mathbf{d}_j^n = \left( \sum_k \mathbf{d}_k^{n+1} W_{kj} \right) f'(net_j^n)$$

最後將(3-20)代入(3-6)中可得

$$\Delta W_{ij} = \mathbf{h} \mathbf{d}_j^n A_i^{n-1} \dots\dots\dots (3-21)$$

以上為 BPN 演算法的關鍵公式，以下再介紹實際進行演算之計算步驟。

學習過程：

以 sigmoid function： $f(net_j^n) = \frac{1}{(1 + e^{-net_j^n})}$  為例

1. 以亂數產生器給定  $W_{ij}$  及  $\mathbf{q}_i$  的初始值
2. 給定輸入向量 X 以及目標輸出向量 T
3. 計算推論輸出向量 Y
  - A. 計算隱藏層輸出向量 H

$$net_h = \sum_i (W_{ih})(X_i) - \mathbf{q}_h$$

$$H_h = f(\text{net}_j^n) = \frac{1}{(1 + e^{-\text{net}_h})}$$

B. 計算輸出層向量  $Y$

$$\text{net}_j = \sum_{ij} (W_{ij})(H_h) - q_j$$

$$Y_j = f(\text{net}_j^n) = \frac{1}{(1 + e^{-\text{net}_h})}$$

4. 計算差距量  $d$

A. 計算輸出層差距量  $d_j$

$$d_j = Y_j(1 - Y_j)(T_j - Y_j)$$

B. 計算隱藏層差距量  $d_h$

$$d_h = H_h(1 - H_h) \sum_j W_{hj} d_j$$

5. 計算權值與偏權值的修正量  $\Delta W$  及  $\Delta q$

A. 計算輸出層的修正量

$$\Delta W_{hj} = \mathbf{hd}_h X_h$$

$$\Delta q_j = -\mathbf{hd}_j$$

B. 計算隱藏層的修正量

$$\Delta W_{ih} = \mathbf{hd}_h X_i$$

$$\Delta q_h = -\mathbf{hd}_h$$

6. 更新權值  $W$  偏權值  $q$

A. 更新輸出層的權值與偏權值

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

$$q_h = q_h + \Delta q_h$$

B. 更新隱藏層的權值與偏權值

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

$$q_h = q_h + \Delta q_h$$

重覆以上 3 至 6 的步驟，直到誤差函數已經小到可以接受的範圍，而回想過程只計算推論輸出向量  $Y$  即可，在學習的過程中



學習速率  $h$ 、隱藏層數目及處理單元數目，可視問題的複雜度不同以試誤法找出最佳的組合。

### 3.3 研究設計

使用類神經網路模式的目的是，在於獲得一個氣喘病的模型，以利於氣喘病發作的預測。在整個類神經網路的建置，本研究中會嘗試以「一層隱藏層」、「二層隱藏層」的方式做對照，以找出本研究是需要使用何種型式的類神經網路，來建置整個氣喘病預測模型。

研究中採用類神經網路中，最常用且有效的 BPN 模式來建立模型，利用一年份密集的資料數據，採 2:1 的隨機方式分為訓練樣本及測試樣本，並以隨機方式挑選。

隱藏層的選擇，先選用一層來做實驗，之後再與二層隱藏層實驗結果做比較，。

而隱藏層神經元數量的選擇上，依據[45]的建議是：依問題的難易程度界定，簡單問題採用(輸入單元個數+輸出單元個數)/2，普通問題採用(輸入單元個數+輸出單元個數)，困難問題採用(輸入單元個數+輸出單元個數)\*2，所以我們會嘗試前兩種方式來建置網路模型。

本研究所用的網路架構，係屬於監督式學習的應用，就是從問題領域間取得訓練範例，並從中學習輸入變數與輸出變數的內在對應規則，以應用於新的樣式，也就是輸入變數值，而推論出輸出變數值。

網路輸入層的輸入單元將會整合天氣及空污的各項屬性，一共是 12 個單元。隱藏層處理單元數，分別以簡單問題及普通問題的考量作比較後決定[45]。輸出層的輸出單元僅考慮“氣喘病發作可能性”一個單元。層與層之間採完全連結，採用標準 sigmoid 函數作為轉換函數，輸出值域[0,1]。之後會考慮時序的因素，訓練出向前推一日、二日、三日的模型，當然神經元的數量也會做因應的改變。

氣喘病的發病可能性為輸出層的目標值，研究中定義以每日氣喘門診人次/年度氣喘門診人次為氣喘病發作可能性。

本研究中考慮時序的作法，是為了找出病人會延遲幾日就診的習性，所以除了原本樣本的記錄外再加上向前推（一、二、三）日內的最大、最小值及差值（最大值-最小值），期能從中找出各地延遲就診天數的樣式。

### 3.4 研究建置步驟

本研究將研究設計分為下面幾個建立步驟：

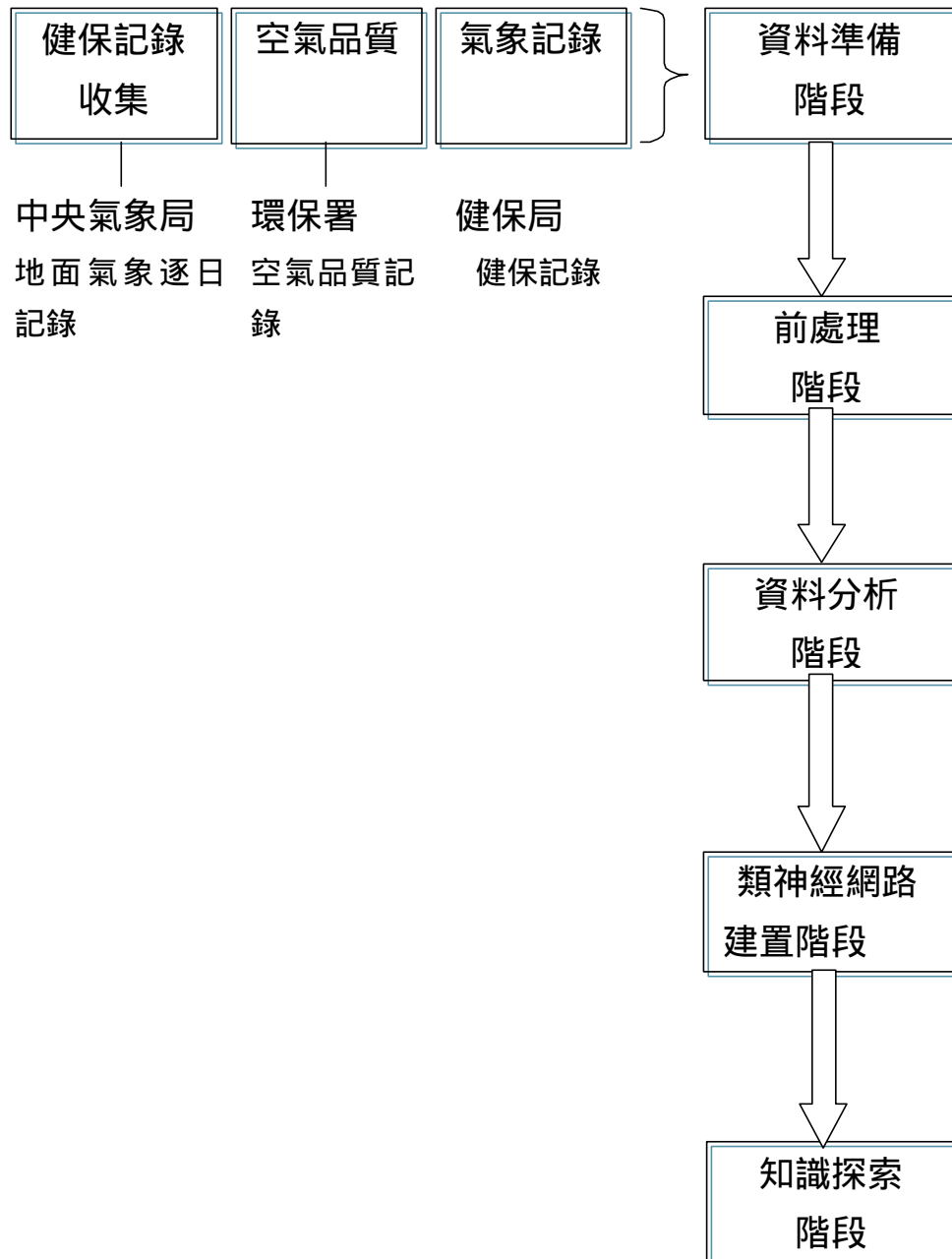


圖 3.3：研究設計圖

整個分類/預測系統的開發原則，是以天氣氣象及空氣品質的條件，來 5 等級劃分[0,1]的機率，可以對每日外在環境評分及預測氣喘病的發作。其建置步驟如下：

## 1. 資料準備階段：

建構一套預測模型需要有輸入變數及輸出變數；而這些輸入變數就是天氣、空污的屬性值，輸出變數則需要健保記錄的每日看診人數，統計出氣喘病發作可能性以作為學習的目標值，因而研究進行的第一個步驟就是收集這些資料。以下將介紹這三個資料庫資料的來源及收集方法。

### A. 氣象資料的收集：

中央氣象局的職掌在於觀測天氣的微妙變化，之後予以記錄成為一個完整的氣象資料庫，並執行每日預測天氣狀況的職掌。而氣象局如何進行這樣的預測動作呢？主要是憑著多年的經驗、氣象學的原理及掌握關鍵的屬性記錄，為了能達成建構研究中需要的預測模型，故需準備此部份的資料庫。所以本研究向中央氣象局資料組洽購逐日地面氣象資料記錄，取得一年份的觀測記錄，並與健保記錄、空污記錄相應。

本研究中收集到全省 15 個氣象站點的逐日記錄，如表 3.1 所示。但因部份縣市未能有氣象測站，故本研究中以鄰近縣市的資料插補，資料的不完整為研究中的缺憾。所以本研究中各縣市均有 365 日的逐日氣象記錄，共 5475 筆記錄。

表 3.1：研究中氣象資料的收集

縣市名稱	氣象測站
台北縣市	淡水、台北、竹子湖
桃園縣	
新竹縣市	新竹
苗栗縣	
台中縣市	台中

續表 3.1：研究中氣象資料的收集（2）

彰化縣	
雲林縣	
嘉義縣市	嘉義
台南縣市	台南
高雄縣市	高雄
屏東縣	恆春
宜蘭縣	宜蘭
花蓮縣	花蓮
台東縣	台東、大武、成功
基隆縣市	基隆
南投縣	

B. 空污記錄的收集：

環保署的職掌在於監測每日的空氣品質狀況，並提供公眾一個客觀的數據參考。依據之前文獻探討中所提及的各項污染物質的危害，所以取得這些資料記錄將有助於預測氣喘病發作的模型建置。為此本研究花費許多的時間在環保署網頁上，下載與其它資料庫同年度的空氣品質記錄。研究中收集到空氣品質記錄共 74 個測站，每站各 365 筆（請參閱表 2.3）的逐日記錄，共 27010 筆記錄。

C. 健保記錄的收集：

本研究中需要健保記錄的原因，是為了要定義出氣喘病各地每日的發作可能性，以作為預測模型中的目標值。並取得與上述兩個資料庫年份相應的記錄，所以健保方面的逐日記錄亦有 365 日（42528 筆記錄）。

在這三個資料庫的組合下，氣象記錄將與空污記錄做配對的動作，即便是將各縣的氣象記錄（一個測站）與各縣的空污記錄（數個測站）以 joint 方式產生樣本的 input pattern：至於 output pattern 就由健保記錄中的看診人次經過統計計算後產生“氣喘發病可能性”的值來填入。研究中的樣本數共為 20710 筆（各縣樣本數不一定，視空污測站的點數多寡）。

## 2. 前處理階段：

資料在取得之後需先經過前處理才能彙整，因為在資料中存在許多的雜訊（缺漏值與無效值），這對於模型的學習產出是很大的影響。面對這樣的問題，有兩種策略可以選擇！不是丟棄就是用平均值回補，但本研究所獲得的資料筆數不多，故難以用丟棄的方式處理，所以本研究中選擇以測站的月平均值填補，相信對模型而言這樣的處理可以把失真的程度降低。

在氣象記錄的前處理，須對『氣溫』欄位做文字轉換成數字的處理，可由氣象局所提供的逐日記錄 A 表中找出對應關係。對於未提供逐日記錄的縣市，以鄰近縣市的記錄作插補以模擬實際值。

在空污記錄的前處理，須先找出備註欄中註明為“無效值”、及數值為“0”、“9999”的記錄或 null、空白值的“缺漏值”記錄，對於以上皆以當月該站平均值回填。

## 3. 資料分析階段：

在先前的文獻探討中找到許多文獻的智慧，省去許多作實驗的時間跟成本。所以根據前人的研究，在解讀資料欄位及轉換資料後，選擇研究中各資料中必要的欄位匯進資料庫作整合，而成為一個資料表。

- A. 在氣象記錄方面，挑選出以下所屬的欄位：  
**a 氣壓 b 氣溫 c 溫差 d 相對濕度 e 風速 f 日照率**  
**g 降水量**
- B. 在空氣品質記錄方面，挑選出以下所屬的欄位：  
**h 一氧化碳 i 臭氧 j 氮氧化物 k 二氧化硫 l 懸浮微粒**  
 以上兩個資料庫是輸入層的輸入的部份。
- C. 在健保記錄方面，挑選出以下所屬的欄位：  
**醫事機構代碼 就醫日期 國際疾病分類號**。根據**醫事機構代碼**與**醫事機構基本檔**做串檔動作後，可以找出該醫事機構位於哪個縣市，得到**縣市區碼**。利用**就醫日期**及**縣市區碼**可以得到各地每日的氣喘病看診人次，透過先前定義的氣喘發作可能性計算出**m 每日的氣喘可能性**，做為輸出層的目標值。

將這三資料庫的資料庫，分別匯進同一個資料表中，成為完整的資料庫。

如下表 3.2 及 3.3 所示：

表 3.2：彙整後的 Database Table(當日版)

天氣							空污					Target
a	B	c	d	e	f	g	h	i	j	k	L	m
7 個欄位							5 個欄位					

表 3.3：彙整後的 Database Table(時序版)

天氣							空污					Target
a	A1	a2	a3	...	...	g3	h	h1	h2	...	l3	m
由 a a1 a2 a3....共 28 個欄位							由 h1.....l3 共 20 個欄位					

時序版考慮到前 n 日天氣、空污的變化，所以在原本欄位後加上（連本日在內）前 n 日的最大值、最小值、及差值（最大值-最小值），因此由原本 12 個欄位的輸入值增加到為 48 個欄位的輸入值，目標值的部份仍是氣喘發病可能性。

#### 4. 類神經網路建置階段：

在一次清蔚科技的研討會中，見識到 Matlab 對數學運算的能力強大(矩陣運算)，對於研究中處理大量的資料將有助於處理效率的提昇，並且 Matlab 提供許多的函式庫及工具箱，對於程式的撰寫上可以節省很多的功夫，所以本研究挑選 Matlab 作為撰寫 BPN 的語言開發環境。

再將先前所提到的“網路層數”、“神經元的數量”.....等細部性的屬性值/資料庫，及“訓練樣本的資料庫”匯入 Matlab 的矩陣中，透過 Matlab 所撰寫的 BPN 程式來訓練出我們的模型，在程式中定義清楚這些屬性及節點數/層數，並依據實驗結果的比較來調整屬性值/網路層數。

依據 Matlab 所提供的類神經網路工具箱，撰寫相關程式  
其中：



模型建置：以縣市為區隔，各縣市皆有一個獨立的模型，如此可以避開地理上的限制，使同個地區的基本條件差異為最小。

輸入層的輸入：是 Database Table 中的  $a \sim l$  欄位。

輸出層的目標值：是 Database Table 中的  $m$  欄位。

正規化：輸入各項變數的最大值及最小值後，Matlab 會自動將數值以 sigmoid 的方式轉換到[0,1]的值域中。

隱藏層數的選擇：一層或兩層

網路架構：一層隱藏層時，簡單問題 12-6-1（時序版 48-24-1），普通問題 12-12-1（時序版 48-48-1）。二層隱藏層時，簡單問題 12-6-3-1（時序版 48-24-12-1），普通問題 12-12-6-1（時序版 48-48-24-1）。

樣本：以各地原始樣本中，以隨機方式選取全部樣本中的  $2/3$  為訓練樣本、 $1/3$  為測試樣本（測試樣本個數控制在 100 個以上）來進行訓練。

亂數種子：固定為一個常數值（0.5），此數值的差異對訓練的結果的影響很小，在[45]裡可以得到驗證。

學習速率：使用 default 值（1.04），研究中曾以 0.99 及 1.05 試做比較，但效果不明顯故使用 default 值。

學習循環數目：在試作台北的模型時，在 MSE 的收斂圖形中，看出收斂的趨勢在 5000 次以後可以降到誤差  $e^{-2}$ ，在 10000 次可以降到  $e^{-3}$ ，在 15000 次時略為提昇，由此可知訓練次數應在 10000 至 15000 次間可以得到最小值，但本研究曾嘗試以 12000、12500、13000 次的訓練，仍無法找出最低點，故本研究中以 5000、10000、15000 次訓練結果作為比較。

其它未定義的參數值：皆按照 Matlab 類神經工具箱中的 default 值，不另加說明。

#### 5. 知識探索階段：

模型產生的輸出值，是模型經過學習、回想過程後對未知樣式產生的預測值，數值可以告訴我們這樣的條件會對氣喘病發作的可能性為何。並對其作分類及預測的實驗，記錄下所有過程結果彙整，以便找出內隱的知識及規則，如：台北人的慣性/趨勢是延遲幾日就診。

根據研究的結果加以分析探討，得到結論及內隱的知識及規則。

對於以上的研究設計，需要以下的軟/硬體設備的支援：

**硬體：**

Pentium II 266 以上等級之硬體相關設備

RAM 128MB 以上

**作業系統：**

Windows 98/98NT、2000 作業系統

**套裝軟體：**

OFFICE 2000 (Excel、Word)

**程式語言開發環境：**

MATLAB 5.2/5.3/6.0

**資料庫：**

SQL 7.0

### 3.5 實例說明

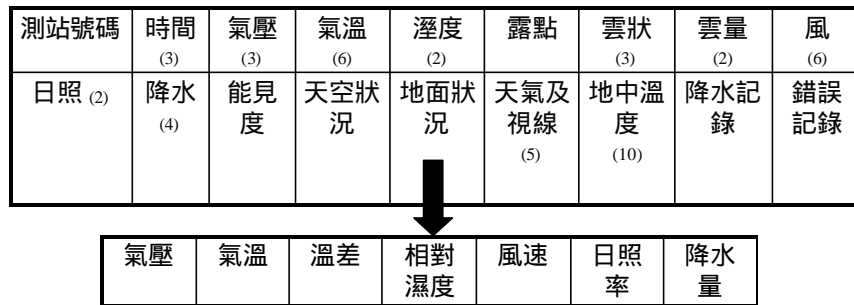
以下僅以建置大台北地區氣喘病模型為例，描述其細部各個過程。

取得中央氣象局的氣象記錄後，再依 3.4 的資料收集過程的步驟，將原本 18 個類別的欄位記錄，剔除相關因素較低的部份欄位，只留下研究中需要的 7 個欄位值。

取得環保署的空污記錄後，再依 3.4 的資料收集過程的步驟，將原本 7 個欄位的記錄，剔除相關因素較低的部份欄位，只留下研究中需要的 5 個欄位。

如下圖所示：

#### 氣象的欄位篩選



#### 空氣污染的欄位篩選

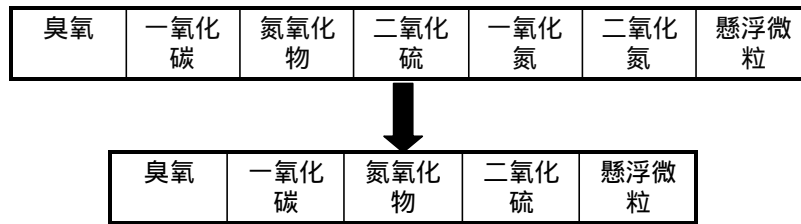


圖 3.4：氣象及空氣污染記錄的欄位篩選

在取得健保局的健保記錄檔後，從“門診處方及治療明細檔”中，找到醫事機構代碼、就醫日期、國際疾病分類號。再利用「醫事機構代碼」這個欄位對應到“醫事機構基本檔”就可以找出縣市區碼。根據就醫日期、國際疾病分類號、縣市區碼，即可統計出每日各地看氣喘病的人次。再將每日的統計人次除以年度看診人次，就可以得到每日的氣喘病發病可能性。得到此值後，就可以做為輸出層（監督式學習）的目標值。

將上述三個資料庫匯整成一個資料表，如表 3.2 所示。

彙整之後的資料表，依照測站地點、時間的先後次序排序，以便日後作延遲就診的“前一日”、“前二日”及“前三日”的樣本，如表 3.3 所示。

接下來類神經網路建置部份，選出大台北地區的樣本，以隨機方式採 2：1 的比例分成訓練樣本及測試樣本。

1. 輸入層的輸入變數是氣壓、氣溫 ..... 等 12 個變數（時序版的，在每個變數之後多了幾日天氣變化的最大值、最小值、差值）。
2. 輸出層的目標值是氣喘發病的可能性/氣喘並看診人次。
3. 資料的正規化動作就交給 Matlab 處理，我們只要輸入各項變數中的最大值及最小值，Matlab 會以 sigmoid 的方式將數值轉換到[0,1]的值域中。
4. 隱藏層的數量依當時訓練的需求給予一層或是兩層。
5. 網路架構的給定在選定隱藏層的層數後，若是選定一層隱藏層時，則以 12-6-1（時序 48-24-1）與 12-12-1（時序 48-48-1）做比較；若是選定兩層隱藏層時，則以 12-6-3-1（時序 48-24-12-1）與 12-12-6-1（時序 48-48-24-1）做比較。
6. 亂數種子則給定 0.5。
7. 學習速率依 Matlab 類神經網路工具箱的 default 值，為 1.04。
8. 學習目標給定為  $e^{-4}$ 。
9. 學習循環數目則給定 5000 次、10000 次、15000 次，三者做比較。
10. 訓練模式皆以 sigmoid 的方式。

之後將訓練所得的結果，以 Excel 比對目標值與預測值的差異，藉以取得誤判率的數值，再將結果彙整成表格資料。

最後是知識探索的階段，我們從表格資料中歸納出大台北地區的民眾有延遲 2 日看診的習性。如下表所示。（\*為最佳值）

表 3.4：知識探索成果（大台北地區）（單位：%）

縣市	訓練 Epoch	12-6-1(時序版 48-24-1)				12-12-1(時序版 48-48-1)			
		當日	前一	前二	前三	當日	前一	前二	前三
台北	5000	16.6	14.68	15.01	14.74	16.44	14.74	15.45	15.01
	10000	16.6	14.79	<b>13.64*</b>	14.25	16	14.47	14.58	14.08
	15000	16.99	14.63	14.03	13.75	16.99	14.25	<b>13.42*</b>	14.03

縣市	訓練 Epoch	12-6-3-1 (時序版 48-24-12-1)				12-12-6-1 (時序版 48-48-24-1)			
		當日	前一	前二	前三	當日	前一	前二	前三
台北	5000	16.33	14.74	13.15	14.68	16.93	17.21	14.58	14.19
	10000	14.74	14.85	13.64	13.81	15.34	14.08	14.14	13.53
	15000	14.85	13.48	<b>12.82*</b>	13.21	15.07	13.48	<b>12.66*</b>	13.64

## 第四章實驗結果

本研究實驗共分三階段；第一階段以三層網路模式(輸入層、一層隱藏層、輸出層)，網路架構分別以 12-6-1 ( 時序版 48-24-1 ) 及 12-12-1 ( 時序版 48-48-1 ) 兩種方式做比較，輸出層以氣喘病可能性為其目標值；第二階段以四層網路模式(輸入層、二層隱藏層、輸出層)，網路架構分別以 12-6-3-1( 時序版 48-24-12-1 )及 12-12-6-1 ( 時序版 48-48-24-1 ) 兩種方式做比較，輸出層以氣喘病可能性為其目標值；前兩階段的實驗以分類結果的誤判率為其衡量標準，而第三階段接續前兩階段的成果，以四層網路模式(輸入層、二層隱藏層、輸出層)，網路架構 12-12-6-1 ( 時序版 48-48-24-1 )，輸出層以實際的氣喘病門診人數為其目標值，目的在於預測每日氣喘看診的人次。

以上三階段的實驗，都將當日、前一日、前二日、前三日等時序因素考慮進去，而產生“當日”、“前一”、“前二”、“前三”的實驗結果。

研究中將氣喘病的發作可能性[0,1]共分成 5 級，0-0.2(綠燈)、0.2-0.4(黃綠燈)、0.4-0.6(黃燈)、0.6-0.8(橙紅燈)、0.8-1(紅燈)，以下列出模型誤判率的值，作為準確性衡量的準則。

「誤判」意指原應屬某類別，而類神經網路卻分類到其它類別。誤判率=誤判值加總/樣本數。

**第一階段實驗結果如下：(\*為最佳值)**

表 4.1：第一階段實驗結果比對（單位：％）

縣市	訓練 Epoch	12-6-1(時序版 48-24-1)				12-12-1(時序版 48-48-1)			
		當日	前一	前二	前三	當日	前一	前二	前三
台北	5000	16.6	14.68	15.01	14.74	16.44	14.74	15.45	15.01
	10000	16.6	14.79	<b>13.64*</b>	14.25	16	14.47	14.58	14.08
	15000	16.99	14.63	14.03	13.75	16.99	14.25	<b>13.42*</b>	14.03
桃園	5000	17.04	18.47	19.01	18.74	15.78	18.74	18.96	18.52
	10000	<b>16.93*</b>	18.52	19.4	18.79	15.23	18.58	19.67	19.67
	15000	<b>16.93*</b>	18.03	20.71	18.47	<b>15.12*</b>	18.96	19.34	19.62
新竹	5000	15.29	14.3	12.88	14.03	16.99	13.48	16.33	12.60
	10000	15.18	12.6	11.67	10.03	17.04	11.73	16.16	10.58
	15000	15.29	10.63	10.58	<b>7.56*</b>	17.10	11.23	16.22	<b>7.78*</b>
苗栗	5000	<b>19.45*</b>	20.16	19.78	20.11	19.73	19.95	20.66	20.44
	10000	19.84	22.63	19.73	21.1	20.49	21.81	20.55	20.49
	15000	20.22	24.88	20.44	21.7	<b>19.12*</b>	22.19	19.89	21.15
台中	5000	15.73	14.85	14.47	14.14	14.90	15.01	133.7	13.59
	10000	15.23	14.08	13.21	12.44	14.58	12.49	11.34	12.77
	15000	14.47	12.93	<b>12.11*</b>	<b>12.11*</b>	13.64	10.79	<b>9.92*</b>	11.34
彰化	5000	15.73	15.23	15.4	15.4	16.71	15.51	15.01	15.34
	10000	15.56	13.48	15.34	14.63	16.60	14.25	14.79	13.70
	15000	15.29	<b>13.04*</b>	13.32	13.37	16.38	<b>12.53*</b>	13.32	13.32
雲林	5000	16.44	15.18	14.14	16.44	15.95	14.74	14.25	14.74
	10000	16.49	14.41	12.38	16.49	15.84	14.30	13.37	13.97
	15000	16.49	13.32	<b>10.79*</b>	16.49	15.40	13.59	<b>13.10*</b>	13.44
嘉義	5000	14.96	13.92	13.81	13.53	14.25	13.59	13.97	13.21
	10000	11.07	12.77	11.73	12.44	13.64	13.37	12.93	12.71
	15000	13.48	11.34	<b>10.85*</b>	11.67	13.59	12.77	<b>11.62*</b>	11.77

續表 4.1：第一階段實驗結果比對（單位：%）（2）

台南	5000	17.04	14.3	13.81	14.30	15.56	14.41	14.58	13.42
	10000	17.04	12.71	12.71	13.15	14.74	13.21	12.38	11.89
	15000	16.99	11.56	<b>10.85*</b>	12.00	14.47	12.71	<b>9.97*</b>	11.18
高雄	5000	15.51	15.56	15.01	16.11	17.37	15.12	15.89	15.40
	10000	14.19	13.59	13.48	14.79	16.11	13.64	14.90	16.55
	15000	13.21	13.26	<b>11.01*</b>	13.81	15.73	11.56	<b>10.36*</b>	14.96
屏東	5000	19.51	16.6	17.04	16.77	18.47	16.22	16.88	16.60
	10000	19.29	14.47	15.23	15.12	18.41	14.90	15.67	15.56
	15000	19.4	14.08	14.74	<b>13.53*</b>	18.52	13.21	14.90	<b>12.97*</b>
台東	5000	14.93	16.72	16.23	18.03	16.23	15.57	16.89	15.41
	10000	<b>14.43*</b>	17.38	17.21	17.54	15.90	17.05	17.05	15.25
	15000	<b>14.43*</b>	16.72	17.87	20.16	<b>15.07*</b>	16.56	16.89	16.72
花蓮	5000	17.54	18.69	16.72	16.23	18.69	16.89	<b>14.92*</b>	17.70
	10000	<b>16.56*</b>	18.52	17.54	18.52	19.34	16.72	17.70	18.03
	15000	<b>16.56*</b>	20	19.02	20.16	18.85	16.72	19.51	20.33
宜蘭	5000	15.89	14.30	15.67	15.45	16.11	14.58	15.78	15.18
	10000	16	10.36	12.38	13.86	14.63	12.99	13.64	12.60
	15000	15.84	<b>9.37*</b>	11.34	13.1	13.97	<b>11.45*</b>	12.60	11.85
基隆	5000	<b>17.05*</b>	22.95	20.66	20	<b>17.70*</b>	21.64	20.33	22.79
	10000	17.70	22.13	21.64	21.15	18.03	22.13	21.15	23.44
	15000	18.03	23.93	21.64	23.11	18.52	24.75	22.13	25.25
南投	5000	16.79	15.31	16.05	16.63	16.54	16.54	17.12	16.30
	10000	16.71	<b>14.9*</b>	15.88	16.05	16.87	<b>15.64*</b>	16.87	15.80
	15000	16.79	15.39	15.47	15.39	17.12	15.72	16.54	15.99

就以上數據顯示（表格中加入灰階網底打\*的黑體字，是 4 種版本中誤判率最低的值），時序效應都是相合的，僅有花蓮部份



有所出入，但後者的數值比前者為低，故推論前者可能限於區域最小值。

在神經元選定的策略上，兩種版本的效果差異性不大，但若為節省時間的情況下，網路架構可以選擇 12-6-1( 時序版 48-24-1 ) 的模式進行實驗；但若追求準確率最大化時，網路架構的選擇策略應參考上表，選擇誤差率小的作為適當之網路架構。

**第二階段實驗結果如下：**

由於四層式網路的訓練十分耗費時間，本階段的實驗以網路架構 12-6-3-1 ( 時序版 48-24-12-1 ) 與網路架構 12-12-6-1 ( 時序版 48-48-24-1 ) 先做比較，取高雄及台北兩地做實驗，根據比對的結果再決定以何者為研究中優先考量的型式。

表 4.2：第二階段實驗結果比對-台北、高雄（單位：%）

縣市	訓練 Epoch	12-6-3-1 ( 時序版 48-24-12-1 )				12-12-6-1 ( 時序版 48-48-24-1 )			
		當日	前一	前二	前三	當日	前一	前二	前三
台北	5000	16.33	14.74	13.15	14.68	16.93	17.21	14.58	14.19
	10000	14.74	14.85	13.64	13.81	15.34	14.08	14.14	13.53
	15000	14.85	13.48	<b>12.82*</b>	13.21	15.07	13.48	<b>12.66*</b>	13.64
高雄	5000	16.23	15.75	14.9	15.2	15.18	15.1	13.53	14.14
	10000	14.11	14.55	14.74	13.42	13.59	12.44	13.04	12.6
	15000	12.99	13.01	12.73	<b>12.02*</b>	12.11	10.79	11.12	<b>9.7*</b>

根據以上兩組實驗結果的比對，發現不管是台北或是高雄在神經元 12-12-6-1 ( 時序版 48-48-24-1 ) 大都有較低的誤判率，所

以將以神經元 12-12-6-1 ( 時序版 48-48-24-1 ) 的型式完成其餘的部份。

表 4.3：第二階段其餘 14 縣市實驗結果 ( 單位：%)

縣市	訓練	實驗結果			
	Epoch	當日	前一	前二	前三
桃園	5000	15.07	18.79	19.01	18.47
	10000	13.21	19.12	19.40	19.45
	15000	<b>12.33*</b>	19.67	19.62	20.33
新竹	5000	12.16	12.66	11.45	12.49
	10000	11.73	9.21	9.81	6.08
	15000	8	7.73	<b>5.7*</b>	7.62
苗栗	5000	<b>19.01*</b>	<b>19.01*</b>	19.12	20.93
	10000	19.45	20.55	20.05	20.93
	15000	19.95	21.10	20.22	21.86
台中	5000	13.26	14.08	13.70	12.60
	10000	11.73	12.38	11.62	10.25
	15000	9.37	9.59	<b>9.23*</b>	9.92
彰化	5000	15.18	13.75	15.18	13.21
	10000	14.96	12.33	13.75	12.11
	15000	14.03	<b>11.16*</b>	12.44	11.62
雲林	5000	15.18	14.47	14.14	14.30
	10000	14.79	14.30	12.38	13.42
	15000	14.79	13.37	<b>10.79*</b>	11.78
嘉義	5000	13.04	13.32	12.27	14.52
	10000	12.77	11.40	11.01	14.36
	15000	9.97	9.1	<b>7.73*</b>	14.30

續表 4.3：第二階段其餘 14 縣市實驗結果（單位：%）（

台南	5000	12.33	13.70	13.53	14.14
	10000	12.77	12.81	12.22	11.01
	15000	9.7	11.12	<b>6.58*</b>	8.05
屏東	5000	15.95	15.56	14.85	15.62
	10000	14.08	12.99	12.33	12.44
	15000	15.01	10.03	12.82	<b>8.11*</b>
台東	5000	<b>13.74*</b>	17.21	16.89	15.90
	10000	<b>13.74*</b>	18.85	16.23	17.87
	15000	17.05	19.02	17.54	19.51
花蓮	5000	18.85	17.87	<b>16.89*</b>	18.20
	10000	18.52	18.52	23.61	17.70
	15000	19.67	20.16	21.31	21.48
宜蘭	5000	14.03	13.37	12.93	13.97
	10000	11.12	10.68	9.15	11.67
	15000	9.75	<b>4.99*</b>	7.45	8.49
基隆	5000	<b>18.69*</b>	20.82	20.33	20.66
	10000	20	20.66	23.11	25.08
	15000	21.64	22.30	22.79	22.03
南投	5000	15.88	15.80	17.04	16.21
	10000	15.64	14.57	16.30	15.56
	15000	16.21	<b>14.40*</b>	16.05	15.47

此階段的實驗結果與上一階段的結果比較，可以發現大多數縣市的誤判率都降低了。由此可見，本研究中的預測模型需要以四層網路模式，且網路架構的選擇應以 12-12-6-1（時序版 48-48-24-1）的型式來建置。故下階段的氣喘病看診人次預測模型，將以 12-12-6-1（時序版 48-48-24-1）的型式來進行，跟本階

段的不同點僅為目標值由“氣喘病發作可能性”改為“氣喘病看診人次”。

第三階段實驗結果如下：

此階段在於驗證，模型是否能有效地預測出該日氣喘病看診的人次，以維持模型的可信度。（\*為最大值，\*\*為最小值，加框者為最理想值）

下表所列結果為模型大台北地區中可信度較高的數值，其餘縣市模型的預測結果請詳見附錄。

表 4.4：第三階段實驗結果(挑選 8 日比對)-台北

縣市 名稱	訓練 Epoch	Target	當日	前一	前二	前三
			output	output	output	output
台北	5000	12	12.51	13.1	13.1	<b>12.14</b>
	10000		12.48	<b>14.9*</b>	<b>14.9*</b>	12.38
	15000		12.46	12.49	12.49	<b>11.95**</b>
台北	5000	13	12.77	11.99	11.99	11.84
	10000		12.89	<b>10.44**</b>	<b>10.44**</b>	12.81
	15000		<b>12.91*</b>	11.86	11.86	11.84
台北	5000	15	12.80	13.03	13.03	11.70
	10000		13.16	12.19	12.19	11.06
	15000		<b>13.28*</b>	12.23	12.23	<b>9.90**</b>
台北	5000	12	11.71	11.41	11.41	<b>12.46*</b>
	10000		11.16	<b>10.76**</b>	<b>10.76**</b>	11.82
	15000		11.05	11.03	11.03	<b>11.89</b>
台北	5000	12	<b>12.27*</b>	11.28	11.28	8.33
	10000		<b>11.97</b>	10.44	10.44	8.68
	15000		11.90	11.10	11.10	<b>8.21**</b>

續表 4.4：第三階段實驗結果（挑選 8 日比對）-台北（2）

台北	5000	12	12.81	12.42	12.42	12.61
	10000		13.08	<b>8.72**</b>	<b>8.72**</b>	12.58
	15000		<b>13.13*</b>	<b>12.11</b>	<b>12.11</b>	12.81
台北	5000	12	<b>12.74*</b>	11.30	11.30	10.45
	10000		12.69	<b>11.88</b>	<b>11.88</b>	<b>9.37**</b>
	15000		12.64	11.43	11.43	9.46
台北	5000	12	12.81	<b>12.14</b>	<b>12.14</b>	12.76
	10000		<b>13.08*</b>	10.50	10.50	11.14
	15000		13.12	12.3	12.30	<b>8.72**</b>

在本研究中對於每個縣市皆獨立訓練個別模型，所以表格中都會有（當日版、前一日、前二日、前三日）四種不同時序，及分別施以 5000、10000、15000 次的訓練，所以每個 Target 都會有 12 個預測值可做為比對，而『最大值』、『最小值』及『最理想值』的標示，是用來評估模型的預測能力；『最大值』是模型對預測人次的高估，『最小值』是模型對預測人次的低估，倘若高估或低估的值都處在容許的誤差範圍下，那就表示模型的可信度很高，而以上 16 個縣市的氣喘病看診人次預測模型，其誤差人次最大皆不超過 5 人次，所以本研究的預測模型的可信度應很高。

## 第五章 結論與未來研究方向

### 5.1 結論

在本研究中建立起全台灣 16 個縣市地區的氣喘病模型，嘗試以類神經網路與資料萃取的技術，找出天氣氣象、空氣污染、健保記錄（氣喘病）的關聯。

首先，在第一階段的部份實驗結果中可以看出，網路架構 12-6-1（時序版 48-24-1）跟網路架構 12-12-1（時序版 48-48-1）的時序效應是幾乎吻合的，並且在網路架構 12-12-1（時序版 48-48-1）中大體上都比神經元 12-6-1（時序版 48-24-1）的誤判率低些；若在研究時間不充裕下，可以採神經元 12-6-1（時序版 48-24-1）方式進行模型的訓練，若在時間充裕的情況下，可以挑選最佳方式進行模型的訓練。

再則，在第二階段的部分，一開始先拿高雄、台北做網路架構 12-6-3-1（時序版 48-24-12-1）與網路架構 12-12-6-1（時序版 48-48-24-1）的比較，發現以網路架構 12-12-6-1（時序版 48-48-24-1）的誤判率比較低，所以之後的 14 個縣市均以網路架構 12-12-6-1（時序版 48-48-24-1）為準，並且本階段實驗中的誤判率也比前一階段的要來的低，所以可見本研究需要以二層隱藏層的 BPN 模型，來建置研究中的預測模型。

先前兩階段的結果皆為分類性的結果，所以可以用誤判率來衡量其優劣程度，但之後第三階段的實驗結果是預測每日氣喘病發病的人次，所以就不能用誤判率來衡量，因此本研究中僅以測試樣本中，幾組實驗數據的展現做為成果比較。

根據以上的實驗結果，我們可以看出各地區民眾延遲就診的習性，以及四層式 BPN 模型可以有效地做分類及預測的任務。在民眾延遲就診方面，或許可以從醫療院所的分布狀況加以改進，使得民眾就診的往來時間縮短，將有助於病情延誤的可能性，並加強宣導健康觀念，使得民眾正視自己的健康問題。對於訓練模型，在透過控制變因下調整可變因子，將有助於找出臨界值及相互影響的關係，將可以幫助學者進一步找出天氣、空污、氣喘病間的關連性。

## 5.2 未來研究方向

以上的結果仍有需要努力的方向，比如：

1. 應該把劃分的地區縮小到各市鎮，但目前天氣氣象方面無法獲得這樣細部的資料；或是把研究中的逐日記錄尺度降低為逐時記錄，但這在實務上需要各單位將這些資料下放，否則模型的誤判率就無法降低，並且不能把氣喘病盛行的區域劃分出來。
2. 本系統雖考慮到時序的建置，但未來可以改找一些模型可考量時序因素的，如 HMM（隱藏式馬可夫模型），如此可以考慮到長期的趨勢，及先前發生的事件對於該日的影響。
3. 本研究僅能獲得一年份的記錄，這在抽樣及訓練上實屬不足，希望後續研究可以取得至少 5 年以上的資料來建置模型。
4. 本研究中未能找尋到可以獲得最佳解的網路參數，此乃受於硬體運算能力限制，希望日後能在資訊科技獲得提昇後，於實驗時以試誤法找出最佳解的狀態。

5. 本研究中僅對氣喘病發作建置出模型,希望日後學者能朝心臟病、皮膚癌.....等等的疾病進行研究。
6. 本研究的成果未來可以建置在網站上,提供一般公眾可以做線上查詢,並由系統自動更新資料記錄,以維持系統的正確性及建立起長期趨勢。



## 參考書目

1. 中央氣象局網站 <http://www.cwb.gov.tw/>，氣象小百科。
2. 王紹漢，「環境污染與健康」，科技圖書，1994 年 8 月。
3. 王進德、蕭大全，「類神經網路與模糊控制理論入門」，全華科技圖書，pp23-58，1994 年 9 月。
4. 王亦凡，「發掘時間序列資料之知識」，淡江大學資訊工程學系碩士論文，1997 年 6 月。
5. 王皓正，「時間序列資料之查詢與資料發掘-以台灣股市為例」，台灣大學資訊管理研究所碩士論文，1999 年 6 月。
6. 王姿乃，「氣喘之環境與遺傳因素研究」，高雄醫學大學醫學研究所博士論文，2000 年 6 月。
7. 王復中，「健保醫療費用審查自動化之研究」，政治大學資訊管理學系碩士論文，2000 年 6 月。
8. 王旭丹，「氣喘中西醫雙效自療手冊」，旺文股份有限公司，pp.25-38，2000 年 9 月。
9. 王台平、葉家榮、陳普曼、曾宗賢、陳慶帆，「應用層級分析程序之就醫選擇智慧代理人」，2001 年資訊管理學術暨實務研討會論文集，2001 年 3 月。
10. 朱慧德、周卓定，「以遠端資料挖掘建構虛擬資料倉儲」，中華管理評論 Vol. 3, No.2 pp.69-81，2000 年 5 月。
11. 行政院環保署網站 <http://www.epa.gov.tw/>，空氣污染的危害。
12. 吳明進，「臺灣北部地區降雨氣候之研究」，國立臺灣大學大氣科學研究所碩士論文，1994 年 6 月。
13. 吳國禎，「資料探索在醫學資料庫之應用」，中原大學醫學工程學系碩士論文，1999 年 6 月。
14. 李紹倫、陳富中，「資料發掘在信用卡之應用」，1999 中華民國科技管理論文集，1999 年 10 月。

15. 沈維明，「資料挖掘之關聯性法則架構以零售業目標行銷為例」，1999 年中華民國科技管理論文集，1999 年 10 月。
16. 卓文福，「應用資料採礦於基因體之重複序列資料庫」，中央大學資訊工程研究所碩士論文，2000 年 6 月。
17. 周卓定，「在全球資訊網上有效挖掘瀏覽路徑之研究」，國防管理學院國防資訊管理所碩士論文，1999 年 6 月。
18. 周政宏，「神經網路-理論與實務」，松崗電腦圖書，pp43-69，1996 年二月。
19. 周賢昭，「以資料採礦技術發展臨床路徑之研究」，中山大學資訊管理學系碩士論文，1998 年 6 月。
20. 林伊蓉，「資料探勘工具之設計與建立」，陽明大學公共衛生研究所碩士論文，1998 年 6 月。
21. 林煜棋，「鋪面道路車行揚塵特性與排放係數之建立」，國立中興大學環境工程學系碩士論文，1998 年 6 月。
22. 林能暉、彭啟明、陳進煌、陳靖沅，「臺灣酸雨之研究：源與受體關係」，第六屆全國大氣科學學術研討會論文集，1999 年。
23. 周昌宏、許晃雄、陳正達、柯文雄、鄒治華，「臺灣環境變遷與全球氣候變遷衝擊之評析-氣候(Environmental Change of Taiwan and The Assessment of Vulnerability to the Impacts of Global Change-Climate)」，行政院國家科學委員會專題研究計劃成果報告，1999 年。
24. 林信忠，「資料萃取法在健保費用查核之研究」，88 年全國管理碩士論文集，1999 年 12 月。
25. 邱顯貴，「以資料發掘經由知識管理建構顧客關係管理模式之研究」，89 年中山大學資管學系研討會論文集，2000 年 5 月。
26. 林文修，「Data Mining 探索(上)(下)」，決策論壇第 13、14 期，  
<http://www.gss.com.tw/eis/13/mining.htm>、  
<http://www.gss.com.tw/eis/14/mining2.htm>。

27. 空氣品質保護與噪音管制處網站  
[http://www.epa.gov.tw/airpollution/空保處\\_0205/default.htm](http://www.epa.gov.tw/airpollution/空保處_0205/default.htm)。
28. 俞大智，「音樂資料庫中重複及循序特徵探勘之研究」，台灣師範大學資訊教育研究所碩士論文，1998年6月。
29. 施建新，「網路資訊與學習系統之中文全文探勘工具」，中央大學資訊工程研究所碩士論文，2000年6月。
30. 張家誠、王立，「氣候變化四問」，明文書局，1994年。
31. 許芝芬，「空間資料萃取模式之建立-以土地估價為例」，逢甲大學土地管理研究所碩士論文，1996年6月。
32. 徐新益，「結合類神經網路與空氣污染預報模式應用於核災應變系統之研究」，成功大學環境工程學系碩士論文，1999年6月。
33. 張瑋倫，「資料挖掘與顧客關係管理之研究」，1999 中華民國科技管理論文集，1999年10月。
34. 斯華齡，「電腦人腦化 神經網路-第六代計算機」，松崗電腦圖書，pp.79-92，1993年8月。
35. 陳建偉，「偵錯與診斷之設計與應用」，中原大學機械工程學系碩士論文，1998年6月。
36. 曾詠淑，「運用資料挖掘技術預測救護車服務量」，成功大學工程科學系碩士論文，1998年6月。
37. 陳文地，「資料探勘應用於犯罪分析模式建構之研究」，88年世新大學資訊管理系研討會論文，pp.95-102，1999年6月。
38. 陳文地，「網路犯罪模式分析及偵防對策之研究」，警察大學資訊管理研究所碩士論文，1999年6月。
39. 陳智宗，「以資料挖礦法挖掘多屬性序列式資料規則之研究」，中央大學資訊管理研究所碩士論文，1999年6月。
40. 郭展盛，「數量化資料模糊挖掘模式之研究」，88年全國管理碩士論文集，1999年12月。

- 41.黃彥文，「資料探勘之應用-會員消費特徵之發掘」，屏東科技大學資訊管理系碩士論文，1999年6月。
- 42.黃錚田，「以資料發掘技術歸劃技職校院課程查詢網站之研究」，台灣師範大學工業教育研究所碩士論文，1999年6月。
- 43.國家衛生研究院網站-全民健康保險學術研究資料庫  
<http://www.nhri.org.tw/nhird/>。
- 44.連宏銘，「終身癌症壽險保費之訂定標準」，政治大學風險管理與保險學系碩士論文，2000年7月。
- 45.葉怡成，「應用類神經網路」，儒林圖書有限公司，1999年4月。
- 46.楊昇宏，「資料挖掘應用於找尋瀏覽網頁之型樣」，逢甲大學資訊工程學系碩士論文，1999年6月。
- 47.喬安 克洛著，陳芳智譯，「過敏 305 問」，原水文化，pp45-91，2000年12月。
- 48.劉慧瑜，「資料提煉技術在虛擬商店之應用」，中興大學應用數學系碩士論文，1999年6月。
- 49.樓玉玲，「以資料發掘分析政大通識課程」，政治大學資訊管理學系碩士論文，1997年6月。
- 50.蔡永恆，「應用資料挖掘技術研究銀行顧客消費行為」，靜宜大學資訊管理學系碩士論文，1999年6月。
- 51.鄭喜文，「資料倉儲的科技與現況」，台大慶齡工業研究中心海外學人邀請講座教材，1997年。
- 52.鄭啟仲、嚴紀中、黃興進，「企業導入資料倉儲之探討」，第10屆國際資訊管理學術研討會論文集，1999年6月。
- 53.燃 燒 的 地 球 網 站  
<http://www.netvigator.com.tw/~a621027/first.htm>。
- 54.盧靜婷，「在 DAG 中挖掘家族特徵規則」，中央大學資訊管理研究所碩士論文，1999年6月。

- 55.蕭正南，「資料探勘應用於股市股價趨勢之研究」，輔仁大學資訊管理學系碩士論文，1998年6月。
- 56.藍中賢，「結合模糊集合理論與貝式分類法之資料探勘技術」，89年中山大學資管學系研討會論文集，2000年5月。
- 57.魏志平、張勳騰、李彥賢，「通信資料庫之資料探勘：目標行銷之應用」，89年中山大學資訊管理學系研討會論文集，2000年5月。
- 58.A. Vellido, P.J.G Lisboa, J. Vaughan "Neural Networks in Business: a survey of applications (1992-1998)" Expert Systems with Application 17 (1999), pp.51-70 .
- 59.Bo K. Wong, Thomas A. Bodnovich, Yakup Selvi "Neural Network applications in the literature (1988-95) Decision Support Systems 19 (1997), pp.301-320.
- 60.Cabena , P.P.O. Hadjinian , R. Stadler , DR. J. Verhees , and A. Zanasi (1997), *Dicovering Data Mining from Concept to Implementation* , Prentice Hall , p.12 .
- 61.Dr. Akeel Al-Attar (1998), "Data mining - Beyond Algorithms", <http://www.attar.com/tutor/mining.htm>
- 62.H. Michael Chung, Michael Manninion, "Introduction to Data Mining and Knowledge Discovery", 1998.
- 63.Hall, C. (1996), "Intelligent Data Mining at IBM: New Products and Applications. ", *Intelligent Software Strategies*, Vol. XII, No.5 pp.1-11.
- 64.Javagi, Rao, Thomas "Choosing a hospital: Analysis of Consumer Tradeoffs." *The Analytic Hierarchy Process* pp.21.
- 65.Michael J. A. Berry Gordon Linoff "Data Mining A Hands-on Approach for Business Professionals". *Data Mining*.
- 66.Nicolas Pasquier, Yves Bastide, Rafik Taouil and Lotfi Lakhal "Efficient Mining of association Rules Using closed Itemset Lattices" *Information Systems* 1999 Vol. 24 No.1, pp.25-46.

67. Rakesh Agrawal , Tomasi Imielincki , Aran Swami "Mining Association Rules Between Sets of Item in Large Database", ACM 1993.
68. Wardlaw AJ. "The role of air pollution in asthma." Clin Exp Allergy 1993, vol.23, pp.81-96.

## 附錄

以下表格數據為本研究中第三階段之結果，此階段在於驗證，模型是否能有效地預測出該日氣喘病看診的人次，以維持模型的可信度。（\*為最大值，\*\*為最小值，加框者為最理想值）

下列附表 1 至 16 所列結果為模型中可信度較高的數值。

附表 1：第三階段實驗結果（挑選 8 日比對）-台北

縣市 名稱	訓練 Epoch	Target	當日	前一	前二	前三
			output	output	output	output
台北	5000	12	12.51	13.1	13.1	<b>12.14</b>
	10000		12.48	<b>14.9*</b>	<b>14.90*</b>	12.38
	15000		12.46	12.49	12.49	<b>11.95**</b>
台北	5000	13	12.77	11.99	11.99	11.84
	10000		12.89	<b>10.44**</b>	<b>10.44**</b>	12.81
	15000		<b>12.91*</b>	11.86	11.86	11.84
台北	5000	15	12.80	13.03	13.03	11.70
	10000		13.16	12.19	12.19	11.06
	15000		<b>13.28*</b>	12.23	12.23	<b>9.90**</b>
台北	5000	12	11.71	11.41	11.41	<b>12.46*</b>
	10000		11.16	<b>10.76**</b>	<b>10.76**</b>	11.82
	15000		11.05	11.03	11.03	<b>11.89</b>
台北	5000	12	<b>12.27*</b>	11.28	11.28	8.33
	10000		<b>11.97</b>	10.44	10.44	8.68
	15000		11.90	11.10	11.10	<b>8.21**</b>
台北	5000	12	12.81	12.42	12.42	12.61
	10000		13.08	<b>8.72**</b>	<b>8.72**</b>	12.58
	15000		<b>13.13*</b>	<b>12.11</b>	<b>12.11</b>	12.81
台北	5000	12	<b>12.74*</b>	11.30	11.30	10.45
	10000		12.69	<b>11.88</b>	<b>11.88</b>	<b>9.37**</b>
	15000		12.64	11.43	11.43	9.46
台北	5000	12	12.81	<b>12.14</b>	<b>12.14</b>	12.76
	10000		<b>13.08*</b>	10.50	10.50	11.14
	15000		13.12	12.3	12.30	<b>8.72**</b>



附表 2：第三階段實驗結果（挑選 8 日比對）-桃園

縣市 名稱	訓練 Epoch	Target	當日	前一	前二	前三
			output	output	output	output
桃園	5000	7	6.53	7.59	6.54	6.44
	10000		7.75	7.79	<b>6.07**</b>	6.29
	15000		<b>8.49*</b>	7.80	6.65	<b>6.76</b>
桃園	5000	10	<b>9.92*</b>	9.1	9.24	9.63
	10000		9.79	8.72	8.95	8.18
	15000		9.76	<b>7.84**</b>	9.67	8.18
桃園	5000	11	10.13	10.22	10.61	10.24
	10000		9.79	10.20	10.05	<b>11.21</b>
	15000		9.76	<b>9.57**</b>	<b>10.01</b>	<b>11.81*</b>
桃園	5000	9	10.78	11.10	9.81	10.13
	10000		10.95	11.72	10.33	10.90
	15000		11.32	<b>12.26*</b>	<b>9.63**</b>	10.67
桃園	5000	10	9.89	9.77	10.64	9.48
	10000		9.78	<b>10.04</b>	<b>11.08*</b>	8.58
	15000		9.76	10.61	11.01	<b>8.29**</b>
桃園	5000	10	10.78	7.41	9.46	7.94
	10000		10.95	6.79	9.80	8.65
	15000		<b>11.32*</b>	<b>6.67**</b>	<b>9.99</b>	7.87
桃園	5000	4	4.44	5.54	6.28	<b>4.25</b>
	10000		3.24	5.55	5.7	4.90
	15000		<b>2.98**</b>	5.08	<b>7.22*</b>	4.89
桃園	5000	8	7.33	<b>8.36*</b>	5.63	<b>5.0**</b>
	10000		<b>7.92</b>	8.26	5.89	5.45
	15000		7.89	8.14	6.06	5.52

附表 3：第三階段實驗結果(挑選 8 日比對)-新竹

縣市 名稱	訓練 Epoch	Target	當日	前一	前二	前三
			output	output	output	output
新竹	5000	5	4.59	4.44	<b>4.22**</b>	4.74
	10000		4.73	4.76	<b>5.08*</b>	4.95
	15000		4.26	5.07	<b>4.96</b>	5.06
新竹	5000	4	<b>3.95</b>	4.32	<b>4.53*</b>	<b>3.69**</b>
	10000		3.84	4.31	4.12	<b>3.95</b>
	15000		4.45	4.44	3.71	3.91
新竹	5000	3	<b>3.71*</b>	2.73	2.92	3.27
	10000		3.62	2.80	3.63	<b>2.71**</b>
	15000		3.09	3.31	<b>3.04</b>	2.87
新竹	5000	4	4.07	3.77	<b>2.85**</b>	3.53
	10000		3.65	4.46	3.16	3.90
	15000		<b>4.02</b>	<b>4.56*</b>	3.46	4.16
新竹	5000	4	<b>5.00*</b>	4.6	4.12	3.57
	10000		4.97	<b>3.99</b>	4.05	3.92
	15000		4.98	<b>3.39**</b>	4.06	3.83
新竹	5000	3	3.71	<b>2.94</b>	2.52	2.57
	10000		3.64	2.79	3.09	<b>2.39**</b>
	15000		<b>3.79*</b>	2.71	3.26	2.57
新竹	5000	4	5.00	4.59	4.24	4.52
	10000		4.97	<b>3.83</b>	<b>3.49**</b>	3.59
	15000		<b>5.32*</b>	3.73	4.32	3.93
新竹	5000	5	<b>4.96</b>	3.88	4.40	4.09
	10000		4.83	3.16	4.74	4.70
	15000		4.61	<b>2.77**</b>	<b>5.08*</b>	4.73

附表 4：第三階段實驗結果(挑選 8 日比對)-苗栗

縣市 名稱	訓練 Epoch	Target	當日	前一	前二	前三
			output	output	output	output
苗栗	5000	4	<b>3.65**</b>	3.8	3.87	<b>4.83*</b>
	10000		4.29	3.66	3.96	4.75
	15000		<b>4.01</b>	4.41	4.34	4.73
苗栗	5000	5	4.24	4.58	<b>5.78*</b>	5.1
	10000		5.4	4.19	<b>4.98</b>	5.58
	15000		4.93	4.38	5.45	<b>4.08**</b>
苗栗	5000	4	4.06	4.15	3.79	<b>3.36**</b>
	10000		3.94	3.94	3.83	3.83
	15000		<b>3.95</b>	<b>4.63*</b>	3.54	4.26
苗栗	5000	3	2.45	<b>2.08**</b>	2.53	3.23
	10000		2.77	2.29	2.53	3.38
	15000		2.77	<b>3.40*</b>	2.77	<b>3.1</b>
苗栗	5000	4	3.94	<b>4.78*</b>	4.37	<b>3.99</b>
	10000		4.40	4.49	4.48	<b>3.7**</b>
	15000		3.93	4.58	3.92	3.87
苗栗	5000	3	3.26	3.4	2.5	3.49
	10000		3.25	3.69	<b>2.46**</b>	3.57
	15000		3.58	<b>3.13</b>	2.86	<b>3.7*</b>
苗栗	5000	4	4.1	3.75	<b>3.27**</b>	4.41
	10000		<b>4.66*</b>	4.24	3.63	4.54
	15000		4.5	<b>4.09</b>	4.18	4.20
苗栗	5000	1	<b>1.93*</b>	<b>0.47**</b>	1.09	1.33
	10000		1.1	1.32	1.29	1.2
	15000		1.15	<b>1.05</b>	1.13	1.14

附表 5：第三階段實驗結果(挑選 8 日比對)-台中

縣市 名稱	訓練 Epoch	Target	當日	前一	前二	前三
			output	output	output	output
台中	5000	15	15.82	<b>12.12**</b>	<b>15.33</b>	13.92
	10000		16.38	13.26	14.31	13.69
	15000		<b>17.25*</b>	12.98	13.90	13.43
台中	5000	11	<b>11.98**</b>	16.88	16.58	16.67
	10000		13.84	17.03	16.84	17.26
	15000		15.91	16.89	16.98	<b>17.50*</b>
台中	5000	18	18.72	16.88	16.58	16.67
	10000		<b>15.06**</b>	17.06	16.84	17.26
	15000		<b>19.44*</b>	<b>17.50</b>	16.98	17.48
台中	5000	16	15.19	16.88	<b>16.58</b>	16.67
	10000		<b>11.90**</b>	17.03	16.84	17.26
	15000		<b>18.60*</b>	16.89	16.98	17.50
台中	5000	19	<b>18.12*</b>	16.88	16.58	16.67
	10000		<b>16.16**</b>	17.05	16.84	17.26
	15000		17.12	17.95	16.98	17.50
台中	5000	17	17.81	16.88	16.58	16.67
	10000		16.54	<b>17.05</b>	16.84	17.25
	15000		<b>15.60**</b>	<b>18.12*</b>	16.98	17.42
台中	5000	20	18.19	16.88	<b>16.58**</b>	16.67
	10000		19.23	17.05	<b>16.58**</b>	17.25
	15000		<b>19.69*</b>	18.12	16.98	17.39
台中	5000	17	16.67	16.86	16.39	15.32
	10000		<b>19.09*</b>	<b>17.05</b>	15.82	14.81
	15000		17.64	16.89	16.13	<b>14.60**</b>

附表 6：第三階段實驗結果(挑選 8 日比對)-彰化

縣市 名稱	訓練 Epoch	Target	當日	前一	前二	前三
			output	output	output	output
彰化	5000	8	9.45	9.53	10.77	8.27
	10000		8.69	9.44	<b>11.55*</b>	6.53
	15000		<b>8.03</b>	9.38	9.66	<b>6.07**</b>
彰化	5000	10	9.47	10.25	9.50	9.53
	10000		10.22	8.40	<b>9.97</b>	8.66
	15000		10.53	<b>7.76**</b>	<b>10.64*</b>	10.11
彰化	5000	9	9.44	10.14	9.97	10.54
	10000		8.66	10.70	7.36	<b>8.86</b>
	15000		8.07	<b>10.84*</b>	<b>7.05**</b>	8.17
彰化	5000	8	8.84	8.77	8.47	<b>10.13*</b>
	10000		9.28	<b>7.97</b>	9.14	7.83
	15000		9.88	7.75	9.84	<b>7.44**</b>
彰化	5000	10	9.48	9.14	9.32	8.77
	10000		<b>10.23</b>	9.32	9.19	8.96
	15000		<b>10.53*</b>	8.65	9.73	<b>8.61**</b>
彰化	5000	10	9.23	10.05	10.07	<b>10.44*</b>
	10000		9.69	<b>10.01</b>	7.94	8.84
	15000		<b>7.90**</b>	9.46	8.26	9.62
彰化	5000	10	<b>9.48**</b>	10.40	11.38	11.17
	10000		<b>10.11</b>	10.61	12.06	<b>12.21*</b>
	15000		10.42	10.46	11.80	12.02
彰化	5000	8	9.33	8.54	8.32	<b>7.99</b>
	10000		8.81	8.89	<b>6.93**</b>	8.07
	15000		8.12	<b>9.34*</b>	7.07	8.88

附表 7：第三階段實驗結果(挑選 8 日比對)-雲林

縣市 名稱	訓練 Epoch	Target	當日	前一	前二	前三
			output	output	output	output
雲林	5000	5	<b>5.04</b>	5.42	<b>5.04</b>	5.18
	10000		5.05	<b>5.86*</b>	<b>4.89**</b>	5.41
	15000		5.05	5.4	5.43	5.47
雲林	5000	5	<b>5.04</b>	4.49	5.41	<b>4.46**</b>
	10000		5.05	<b>5.07*</b>	4.84	5.23
	15000		5.05	4.85	4.56	5.45
雲林	5000	5	<b>5.04</b>	<b>7.66*</b>	5.16	6.41
	10000		5.05	7.46	5.33	5.28
	15000		5.05	7.45	<b>4.63**</b>	5.54
雲林	5000	6	<b>5.04**</b>	5.45	5.56	5.56
	10000		5.05	<b>6.27*</b>	6.20	<b>5.84</b>
	15000		5.05	<b>6.27*</b>	5.57	5.60
雲林	5000	6	5.04	4.09	4.33	<b>3.61**</b>
	10000		5.05	3.88	4.42	4.07
	15000		5.05	<b>3.61**</b>	<b>5.2*</b>	4.32
雲林	5000	4	3.79	3.48	3.52	3.79
	10000		3.79	3.78	<b>3.97</b>	3.24
	15000		3.79	3.9	<b>4.75*</b>	<b>3.03**</b>
雲林	5000	3	3.79	3.24	3.34	<b>3.06</b>
	10000		3.79	3.28	3.43	<b>2.94**</b>
	15000		3.79	3.63	<b>4.13*</b>	3.13
雲林	5000	4	3.79	3.82	<b>3.35**</b>	<b>4.61*</b>
	10000		3.79	4.23	3.72	4.14
	15000		3.79	<b>4.04</b>	3.91	3.44

附表 8：第三階段實驗結果(挑選 8 日比對)-嘉義

縣市 名稱	訓練 Epoch	Target	當日	前一	前二	前三
			output	output	output	output
嘉義	5000	7	5.97	6.92	<b>5.67**</b>	<b>7.3*</b>
	10000		6.26	<b>7.03</b>	5.75	6.9
	15000		6.29	7.22	<b>5.67**</b>	6.58
嘉義	5000	6	6.29	6.47	6.29	5.27
	10000		7.37	7.36	6.16	5.29
	15000		<b>7.83*</b>	7.16	<b>6.11</b>	<b>5.03**</b>
嘉義	5000	5	5.97	6.13	6.58	5.39
	10000		5.81	5.66	6.63	<b>5.38**</b>
	15000		5.98	6.04	<b>6.68*</b>	5.6
嘉義	5000	7	6.01	<b>7.12</b>	<b>5.50**</b>	6.41
	10000		6.26	7.23	5.74	<b>7.47*</b>
	15000		6.34	6.73	5.74	6.8
嘉義	5000	8	<b>6.02**</b>	6.84	6.22	7.36
	10000		6.25	7.55	6.28	7.22
	15000		6.23	<b>7.69*</b>	6.19	7.34
嘉義	5000	6	5.97	6.20	6.46	<b>5.48**</b>
	10000		5.81	5.61	<b>6.5*</b>	6.39
	15000		<b>5.98</b>	5.72	<b>6.5*</b>	6.35
嘉義	5000	5	6.09	5.41	6.35	<b>6.71*</b>
	10000		5.81	<b>4.81**</b>	6.26	6.01
	15000		5.32	<b>5.04</b>	6.31	6.13
嘉義	5000	4	4.18	4.66	5.75	4.89
	10000		4.34	4.31	<b>5.87*</b>	4.95
	15000		4.41	<b>3.89**</b>	5.8	4.33

附表 9：第三階段實驗結果(挑選 8 日比對)-台南

縣市 名稱	訓練 Epoch	Target	當日	前一	前二	前三
			output	output	output	output
台南	5000	13	12.86	12.61	11.11	11.84
	10000		<b>12.97*</b>	12.32	<b>10.89**</b>	12.73
	15000		12.96	12.93	11.24	12.58
台南	5000	12	12.86	12.94	<b>12.15**</b>	<b>14.49*</b>
	10000		12.97	13.39	12.81	14.34
	15000		12.96	13.11	13.04	12.65
台南	5000	12	13.54	13.03	13.84	14.40
	10000		13.72	14.53	14.52	<b>15.34*</b>
	15000		13.35	14.92	15.14	<b>12.68**</b>
台南	5000	17	15.52	<b>12.66**</b>	12.95	13.66
	10000		15.82	12.92	13.58	14.68
	15000		<b>15.83*</b>	13.63	13.66	14.96
台南	5000	13	<b>11.58**</b>	12.40	11.95	13.79
	10000		<b>11.58**</b>	12.57	12.25	15.14
	15000		11.62	<b>12.81</b>	12.65	<b>15.56*</b>
台南	5000	16	15.52	<b>13.11**</b>	13.86	14.59
	10000		15.82	14.44	14.55	16.48
	15000		<b>15.83</b>	14.80	14.69	<b>17.12*</b>
台南	5000	14	15.52	<b>13.09**</b>	13.23	14.56
	10000		15.82	14.90	<b>13.90</b>	15.19
	15000		<b>15.83*</b>	15.44	14.44	15.56
台南	5000	10	11.58	11.62	12.14	11.38
	10000		11.58	11.28	12.15	11.98
	15000		11.62	<b>11.05**</b>	11.89	<b>12.41*</b>



附表 10：第三階段實驗結果(挑選 8 日比對)-高雄

縣市 名稱	訓練 Epoch	Target	當日	前一	前二	前三
			output	output	output	output
高雄	5000	14	14.71	15.57	14.30	13.83
	10000		14.92	17.45	14.49	<b>13.94</b>
	15000		14.93	<b>18.34*</b>	15.23	<b>13.46**</b>
高雄	5000	15	14.87	14.91	15.10	14.76
	10000		<b>14.94</b>	13.77	13.41	13.88
	15000		<b>14.94</b>	13.88	<b>12.18**</b>	<b>15.16*</b>
高雄	5000	16	14.84	15.42	<b>13.83**</b>	15.07
	10000		14.94	16.21	<b>16.04</b>	15.72
	15000		14.94	16.29	<b>17.76*</b>	15.83
高雄	5000	10	10.93	11.90	<b>13.57*</b>	11.45
	10000		10.65	10.60	11.55	10.62
	15000		10.66	10.81	<b>9.84</b>	<b>9.75**</b>
高雄	5000	11	11.17	<b>15.46*</b>	13.26	<b>11</b>
	10000		10.67	14.08	12.87	10.14
	15000		10.66	13.65	11.89	<b>9.28**</b>
高雄	5000	13	14.87	15.57	15.38	14.64
	10000		14.94	15.79	15.46	14.44
	15000		14.94	15.70	<b>16.12*</b>	<b>13.74**</b>
高雄	5000	14	14.87	15.57	15.37	<b>14.80**</b>
	10000		14.94	15.77	15.72	15.82
	15000		14.94	16.15	<b>16.42*</b>	15.59
高雄	5000	12	<b>14.87**</b>	14.88	15.31	14.97
	10000		14.94	15.77	16.32	16.00
	15000		14.94	16.11	<b>17.53*</b>	16.56

附表 11：第三階段實驗結果(挑選 8 日比對)-屏東

縣市 名稱	訓練 Epoch	Target	當日	前一	前二	前三
			output	output	output	output
屏東	5000	6	5.26	6.20	5.67	<b>6.76*</b>
	10000		5.22	<b>5.93</b>	<b>3.63**</b>	<b>6.07</b>
	15000		5.22	5.69	6.59	5.60
屏東	5000	7	<b>6.90</b>	6.30	7.31	6.77
	10000		<b>6.90</b>	<b>6.17**</b>	6.76	6.78
	15000		<b>6.90</b>	6.36	<b>8.36*</b>	6.56
屏東	5000	8	<b>6.90**</b>	7.52	7.58	7.19
	10000		<b>6.90**</b>	<b>7.82</b>	7.68	7.63
	15000		<b>6.90**</b>	7.54	<b>8.38*</b>	7.15
屏東	5000	4	5.26	6.11	4.79	<b>6.15*</b>
	10000		5.22	5.39	<b>3.99</b>	4.78
	15000		5.22	4.58	3.79	<b>3.51**</b>
屏東	5000	7	6.90	6.46	6.13	6.61
	10000		6.90	6.13	<b>7.98*</b>	<b>5.26**</b>
	15000		6.90	6.48	<b>6.94</b>	5.61
屏東	5000	4	6.90	6.68	7.21	7.20
	10000		6.90	6.42	<b>7.76*</b>	7.56
	15000		6.90	<b>5.73**</b>	7.54	5.92
屏東	5000	7	6.61	6.43	6.66	6.14
	10000		6.76	6.10	<b>7.1*</b>	7.04
	15000		6.90	<b>6.09**</b>	6.67	<b>7.01</b>
屏東	5000	6	6.90	5.7	7.22	7.26
	10000		6.90	<b>4.61**</b>	7.61	<b>8.61*</b>
	15000		6.90	<b>5.17</b>	7.1	6.92

附表 12：第三階段實驗結果(挑選 8 日比對)-台東

縣市 名稱	訓練 Epoch	Target	當日	前一	前二	前三
			output	output	output	output
台東	5000	1	1.15	1.25	<b>0.74**</b>	1.78
	10000		1.13	1.14	<b>1.04</b>	1.34
	15000		1.47	<b>2.41*</b>	1.23	<b>1.04</b>
台東	5000	1	1.55	1.49	1.48	<b>0.86</b>
	10000		<b>2.34*</b>	1.42	1.59	0.85
	15000		1.7	1.56	1.51	<b>0.67**</b>
台東	5000	4	2.39	1.84	2.09	2.40
	10000		3.02	1.59	2.55	2.22
	15000		<b>3.49*</b>	<b>1.40**</b>	2.68	2.19
台東	5000	3	2.23	<b>0.95**</b>	1.69	2.39
	10000		1.47	1.95	1.34	<b>2.71*</b>
	15000		1.40	1.57	1.83	2.22
台東	5000	0	0.46	0.92	0.85	<b>0.33**</b>
	10000		0.92	0.49	<b>0.42</b>	0.60
	15000		<b>0.94*</b>	0.46	0.61	0.78
台東	5000	0	0.43	<b>0.94*</b>	0.2	0.92
	10000		0.24	0.85	0.33	0.5
	15000		0.24	0.78	<b>0.14**</b>	0.54
台東	5000	1	0.57	1.28	<b>0.17**</b>	1.08
	10000		0.24	1.40	0.47	1.28
	15000		0.76	1.27	<b>0.95</b>	<b>1.51*</b>
台東	5000	1	<b>0.49**</b>	1.46	1.24	1.73
	10000		1.88	1.55	1.53	2.08
	15000		1.66	<b>0.92</b>	<b>2.39*</b>	1.87

附表 13：第三階段實驗結果(挑選 8 日比對)-花蓮

縣市 名稱	訓練 Epoch	Target	當日	前一	前二	前三
			output	output	output	output
花蓮	5000	1	1.52	1.63	1.26	1.32
	10000		<b>0.90**</b>	1.29	1.35	1.14
	15000		<b>1.02</b>	1.59	1.15	<b>1.89*</b>
花蓮	5000	2	2.73	2.49	2.95	2.86
	10000		2.36	2.55	2.43	<b>2.12**</b>
	15000		<b>3.61*</b>	2.81	2.95	2.91
花蓮	5000	3	<b>2.70</b>	3.38	2.25	2.47
	10000		3.44	<b>4.1*</b>	<b>2.23**</b>	2.59
	15000		3.62	3.90	2.36	2.64
花蓮	5000	3	2.72	<b>2.87</b>	3.39	<b>3.98*</b>
	10000		3.36	<b>1.95**</b>	3.69	3.61
	15000		3.16	2.5	3.57	3.29
花蓮	5000	2	2.58	2.28	2.41	2.45
	10000		2.85	2.59	2.74	3.21
	15000		2.23	2.71	<b>2.06**</b>	<b>3.44*</b>
花蓮	5000	2	2.28	<b>2.01**</b>	2.29	3.26
	10000		3.18	2.89	2.39	3.08
	15000		2.19	2.94	2.1	<b>3.6*</b>
花蓮	5000	3	2.5	2.45	2.7	<b>2.98</b>
	10000		3.16	2.83	<b>2.14**</b>	2.43
	15000		<b>3.54*</b>	2.91	2.26	3.22
花蓮	5000	3	2.78	2.82	2.82	2.43
	10000		3.52	<b>3.01</b>	3.26	2.49
	15000		<b>3.65*</b>	2.91	2.83	<b>2.42**</b>

附表 14：第三階段實驗結果(挑選 8 日比對)-宜蘭

縣市 名稱	訓練 Epoch	Target	當日	前一	前二	前三
			output	output	output	output
宜蘭	5000	5	<b>5.11</b>	<b>5.94*</b>	4.64	4.4
	10000		5.15	5.22	4.43	4.72
	15000		5.12	4.63	4.59	<b>4.1**</b>
宜蘭	5000	3	3.16	3.25	3.39	<b>3.92*</b>
	10000		3.37	3.08	<b>3.02**</b>	3.52
	15000		3.34	3.03	3.58	3.43
宜蘭	5000	3	3.02	3.65	3.14	3.34
	10000		<b>3**</b>	3.25	<b>3.88*</b>	<b>3**</b>
	15000		3.35	3.62	3.46	3.01
宜蘭	5000	2	2.75	2.45	2.26	2.7
	10000		2.66	2.13	<b>1.92**</b>	<b>2.88*</b>
	15000		2.46	2.64	<b>2</b>	2.38
宜蘭	5000	5	<b>5.01</b>	<b>4.67**</b>	<b>5.68*</b>	5.02
	10000		5.11	4.74	5.56	5.11
	15000		5.57	4.95	5.3	5.24
宜蘭	5000	4	<b>4.19**</b>	4.76	4.9	<b>5.24*</b>
	10000		4.73	4.20	4.48	5.05
	15000		4.55	4.89	4.69	4.92
宜蘭	5000	6	5.19	4.92	5.29	5.90
	10000		5.46	<b>4.74**</b>	<b>6.02*</b>	5.52
	15000		5.65	5.53	5.96	5.15
宜蘭	5000	7	6.21	<b>6.72</b>	6.39	5.7
	10000		<b>5.68**</b>	7.34	6.02	5.85
	15000		6.64	<b>7.51*</b>	6.83	6.74

附表 15：第三階段實驗結果(挑選 8 日比對)-基隆

縣市 名稱	訓練 Epoch	Target	當日	前一	前二	前三
			output	output	output	output
基隆	5000	4	<b>3.04**</b>	3.55	3.26	3.25
	10000		3.31	<b>3.84*</b>	3.41	3.58
	15000		3.32	3.4	3.57	3.15
基隆	5000	3	2.34	<b>2.98</b>	<b>2.24**</b>	3.39
	10000		2.44	<b>3.69*</b>	2.96	2.7
	15000		2.25	3.23	2.96	2.88
基隆	5000	2	2.34	1.75	1.7	<b>2.52*</b>
	10000		1.79	<b>1.62**</b>	2.41	1.82
	15000		<b>1.97</b>	2.11	2.2	2.30
基隆	5000	2	2.29	<b>2.91*</b>	2.44	2.28
	10000		2.17	<b>1.9**</b>	2.34	2.56
	15000		2.12	2.53	2.49	2.15
基隆	5000	4	<b>4.02</b>	3.85	3.93	3.44
	10000		3.84	3.83	3.44	<b>3.25**</b>
	15000		3.96	3.83	<b>4.11*</b>	<b>4.02</b>
基隆	5000	3	2.80	<b>3.08</b>	3.34	3.18
	10000		<b>2.61**</b>	3.51	3.1	3.5
	15000		2.7	<b>3.72*</b>	3.25	2.84
基隆	5000	0	0.47	0.66	0.48	<b>0.75*</b>
	10000		0.21	<b>0.04**</b>	0.26	0.58
	15000		0.09	0.57	0.06	0.66
基隆	5000	1	<b>1.68*</b>	1.46	1.6	1.13
	10000		<b>1</b>	1.05	1.35	1.38
	15000		<b>0.82**</b>	0.91	1.17	0.84

附表 16：第三階段實驗結果(挑選 8 日比對)-南投

縣市 名稱	訓練 Epoch	Target	當日	前一	前二	前三
			output	output	output	output
南投	5000	4	3.89	<b>3.8**</b>	4.11	<b>3.95</b>
	10000		4.16	4.06	<b>4.43*</b>	3.93
	15000		4.16	4.07	3.79	4.08
南投	5000	4	3.59	3.6	3.72	<b>3.55**</b>
	10000		4.16	3.64	4.29	4.16
	15000		<b>4.59*</b>	4.35	4.2	<b>4.04</b>
南投	5000	3	<b>2.53**</b>	2.84	<b>3.93*</b>	2.84
	10000		<b>2.93</b>	3.62	2.86	2.65
	15000		3.11	3.35	2.91	3.18
南投	5000	1	1.37	1.17	1.39	1.22
	10000		<b>1.42*</b>	0.94	1.14	0.95
	15000		0.96	<b>0.81**</b>	1.04	<b>1</b>
南投	5000	2	2.57	<b>2.58*</b>	1.82	2.46
	10000		<b>1.70**</b>	1.75	2.34	2.06
	15000		<b>1.96</b>	2.32	1.95	1.95
南投	5000	3	2.53	<b>2.46**</b>	2.49	<b>3.27*</b>
	10000		3.09	2.51	2.63	2.88
	15000		3.09	2.89	3.26	<b>3.06</b>
南投	5000	5	4.89	<b>4.75**</b>	4.9	4.86
	10000		5.16	5.06	5.2	4.96
	15000		5.16	<b>5.01</b>	5.2	<b>5.21*</b>
南投	5000	0	0.37	0.21	0.6	<b>0.85*</b>
	10000		0.42	0.16	0.35	0.75
	15000		0.21	<b>0.14**</b>	0.16	0.31

