

國立臺灣大學管理學院資訊管理學研究所

碩士論文

Department of Information Management

College of Management


National Taiwan University

Master Thesis

IDTBN 方法應用於整合專家意見之實證研究  
——以電力長期負載預測為例

An Empirical Study on IDTBN Applied to  
the Integration of Expert Opinions

-- The Case of Long-term Electric Load Prediction



楊琇珊

Yang, Hsiu-Shan

指導教授：曹承礎 博士

Advisor: Chou Seng-Cho, Ph.D.

中華民國 97 年 7 月

July, 2008

IDTBN 方法應用於整合專家意見之實證研究  
——以電力長期負載預測為例  
An Empirical Study on IDTBN Applied to  
the Integration of Expert Opinions  
-- The Case of Long-term Electric Load Prediction



本論文係提交國立台灣大學  
資訊管理學研究所作為完成碩(博)士  
學位所需條件之一部份

研究生：楊琇珊 撰

中華民國九十七年七月



## 誌謝辭

如果論文是兩年研究所生活的總結，那麼誌謝辭就是用以感謝這一路上陪我度過的人們。論文與口試得以順利完成，並非單靠個人努力完成，首先要感謝我的指導教授曹承礎老師，在研究方向上給我許多建議，點出許多論文撰寫時應注意的問題，讓我得以發現論文進行中的盲點；也要特別感謝台電的陳鳳惠學姐，提供許多關於研究案例的資料，給予我寶貴意見並協助我擬定研究架構，也不斷鼓勵我不要放棄，持續努力把事情作到最好；在論文寫作期間，數次寄信詢問味亭學姐一些研究時遭遇的問題，她也非常熱心為我解答，非常感謝。

研究所生活中少不了一群陪著我成長的研究室夥伴們，我們有最堅強的研究陣容：貝瑜、立穎、宇頌、亭光、鼎鈞、Roger（志成）、彥一（志光）、子林，在課業上相互切磋，論文後期大家彼此間鼓勵加油的話語，也一路伴隨我到口試結束。兩年裡常有聚餐和唱歌等活動，也因為有你們，為我的生活妝點了不少活潑色彩，這些都是難以忘懷的回憶點滴。研究室的學長姐和學弟妹，也都很親切熱心，儘管大家一起出現在研究室的時間不多，但有大家一起在研究室就會感覺到格外溫暖和氣氛融洽，貼心的學弟妹在學期末為我們策畫的送舊烤肉聚會，讓我們真的很捨不得離開這個大家庭，以後有空會回來看看大家的。此外，也要感謝我的大學好友聆寧，總是在我面臨論文瓶頸時，以即時通訊軟體陪我聊天，分享許多新鮮的訊息給我，稍稍舒緩我的論文壓力，並給予我論文摘要潤飾和文法上的建議，非常感謝。最後，還有感謝陪我走過三年半的男友，雖然無法常相見，但睡前的手機通話總是陪伴著我入夢，給予我心靈上莫大的支持。

我也要特別感謝我的家人，感謝一路支持栽培我到研究所的父母，養育我們的辛苦從不掛在嘴邊，並常在我低潮或忙碌於課業時，關心並叮嚀我要注意身體健康，回到家看到妹妹的笑容和她們貼心的問候，便遺忘生活中所遇到的一切不愉快，也深刻認同「家是唯一的避風港」這句話。再次感謝所有在成長路上給予我幫助及關懷的朋友，願將這份喜悅與你們分享。

# 論文摘要

論文題目：IDTBN 方法應用於整合專家意見之實證研究——以電力長期負載預測為例

作者：楊琇珊

民國九十七年七月

指導教授：曹承礎 博士

企業在進行研發及生產等活動皆須仰賴大量資訊，因此，如何將資訊轉化為輔助企業決策的知識，即成為資訊超載時代之下所面對的首要課題。而在獲取知識的途徑中，通常針對特定專業領域，萃取並整合專家意見的作法是為較常見的模式。

若能有效統整專家意見並將其以結構化模型呈現，則可提供日後分析與應用，並進一步支援相關領域之決策制定。本研究先擷取專家意見與討論中所提及的外生變數及對預測值之可能影響程度，再整合決策樹（Decision Tree）分析與貝氏網路（Bayesian Network）方法，將所得之意見資料建立成完整的專家知識脈絡。使原本僅能得到單一分析者觀點的預測值，現在能保存集合多元專家預測觀點的推論架構，進而作為決策支援系統預測值之微調參考。

由於電力需求與供給面的整合，可大幅提升電力資源之使用績效，電力負載預測也進而成為重要課題。而透過決策支援系統（Decision Support System, DSS）的輔助，已能達到以長期歷史資料佐以專家意見來進行電力負載量之預測，再由專家根據決策支援系統的預測值作微調，然而此傳統作法並無保存專家意見背後據以判斷的龐大知識架構。本研究乃選定電力產業為實例應用研究之標的，以研究中所提出之 IDTBN（Integrated Decision Tree and Bayesian Network）方法，期能整合專家諮詢會議中所提出之意見並建立出分析模型，使專家對決策支援系統之結果值進行微調時，能以此反映出未來發展趨勢的模型並作為參考，使電力負載預測結果更具準確性。

**關鍵字：**貝氏網路、決策樹、資料探勘、專家意見整合、電力長期負載預測

# THESIS ABSTRACT

An Empirical Study on IDTBN Applied to the Integration of Expert Opinions  
-- The Case of Long-term Electric Load Prediction

By Hsiu-Shan Yang

MASTER DEGREE OF BUSSINESS ADMINISTRATION

DEPARTMENT OF IMFORMATION MANAGEMENT

NATIONAL TAIWAN UNIVERSITY

JULY 2008

ADVISER : Dr. Seng-Cho Chou

The enterprise rely on a lot of information to lead production, research and development activities; therefore, how to transfer the information into knowledge which assists enterprises in making decision, becomes one of the most important issue in the age of information overload. Usually, extracting and integrating expert opinions of specific fields are the common rule to acquire knowledge.

Effectively converging expert opinions and showing the structure of the knowledge model can be used for analysis and application in the future and the decision-making of related areas. This thesis aims to extract the extraneous variables and their impact degree toward predicting value. The goal is to integrate Decision Tree analysis and Bayesian Network and construct a complete knowledge profile of experts with those variables and data. Not only will the predicting value of single analyst's view but also an inferring structure including multiple views of experts be acquired.

Because the integration of demand and supply side for electric power could enhance the utility performance of electric power resource, the prediction of power load becomes more and more important. With the aid of Decision Support System (DSS), the power load could be predicted by making use of historical records and

experts' opinions and slightly adjusting the predicting value of DSS by experts, but the great knowledge structure used for judgment was not kept. In this study, we select the power industry as our research target, and use IDTBN (Integrated Decision Tree and Bayesian Network) method presented in this paper, to integrate the opinions given in the expert meeting and build an analysis model. As a result, it can improve the prediction of electric power load and cope with the changing trend of the future.

**Keywrds:** Bayesian Network, Decision Tree, Data Mining, Integration of Expert Opinions, Long-term Electric Load Prediction



# 目錄

誌謝辭.....	I
論文摘要.....	II
THESIS ABSTRACT.....	III
目錄.....	V
圖目錄.....	VII
表目錄.....	VIII
第一章、緒論.....	1
第一節 研究動機.....	1
第二節 研究目的.....	2
第三節 研究步驟與流程.....	4
第四節 論文章節安排.....	6
第二章、文獻探討.....	8
第一節 決策樹方法.....	8
2.1.1 決策樹分析介紹.....	9
2.1.2 決策樹之構建與修剪.....	10
第二節 貝氏網路模式.....	11
2.2.1 建構貝氏網路.....	12
2.2.2 貝氏網路之數學模式與架構.....	13
2.2.3 貝氏網路之學習模式.....	17
第三節 類神經網路.....	19
2.3.1 類神經網路理論及架構.....	19
2.3.2 倒傳遞類神經網路.....	22
第四節 案例式推理.....	24
2.4.1 案例式推理概念.....	25
第三章、研究方法.....	28
第一節 應用實例—台電公司電力長期負載預測.....	28
第二節 IDTBN 模式.....	32
第三節 實證研究方法與設計.....	35
第四節 研究架構.....	37
第五節 專家意見內容過錄.....	37
第四章、研究過程與結果.....	39
第一節 資料前置處理.....	39
第二節 效益衡量指標.....	42
第三節 決策樹方法結果.....	43
第四節 貝氏網路方法結果.....	47
第五節 IDTBN 方法應用結果.....	51



第六節 IDTBN 方法效益分析 .....	55
第五章、結論與建議 .....	60
第一節 結論 .....	60
第二節 建議 .....	62
5.2.1 研究限制 .....	62
5.2.2 未來研究方向 .....	63
附錄一：專家諮詢會議資料過錄表（前 5 筆資料） .....	64
附錄二：資料過錄格式表 .....	70
附錄三：原始資料檔案 .....	72
參考文獻 .....	74



## 圖目錄

圖 一-1	本論文研究流程 .....	6
圖 二-1	DAG 圖形 .....	13
圖 二-2	貝氏網路範例：family-out 問題.....	15
圖 二-3	貝氏網路範例：信用卡盜刷問題 .....	16
圖 二-4	類神經網路單元之模型 .....	20
圖 二-5	類神經網路的範例.....	21
圖 二-6	標準倒傳遞類神經網路.....	22
圖 二-7	案例表達方式 .....	25
圖 二-8	CBR 系統推論流程.....	26
圖 二-9	案例式推理系統之範例.....	27
圖 三-1	本論文研究架構 .....	37
圖 四-1	WEKA 交叉確認模式設定 .....	44
圖 四-2	WEKA J48 參數設定介面.....	45
圖 四-3	WEKA J48 決策樹模型.....	46
圖 四-4	WEKA BayesNet 參數設定介面 .....	48
圖 四-5	WEKA BayesNet Simple Estimator 參數設定介面 .....	49
圖 四-6	WEKA BayesNet search algorithm : K2 參數設定介面 .....	49
圖 四-7	WEKA BayesNet 貝氏網路模型 .....	50
圖 四-8	參考決策樹模型之兩大類資料 .....	51
圖 四-9	WEKA BayesNet 貝氏網路模型：分類一 .....	52
圖 四-10	WEKA BayesNet 貝氏網路模型：去除錯誤資料之分類一 .....	53
圖 四-11	WEKA BayesNet 貝氏網路模型：分類二.....	54
圖 四-12	WEKA BayesNet 貝氏網路模型：去除錯誤資料之分類一 與其部分 分機率分布表 .....	57
圖 四-13	WEKA BayesNet 貝氏網路模型：分類二 與其部分機率分布表 .....	58
圖 四-14	各階段方法之 Accuracy 比較圖 .....	59

## 表目錄

表 三-1	常用支援進行長期電力負載預測之計量方法比較.....	30
表 三-2	個案公司電力用電預測結果—時間序列分析模式.....	31
表 四-1	資料分析所選取之變數名稱與代表意義.....	40
表 四-2	專家意見過錄對照表.....	41
表 四-3	confusion matrix 示意圖.....	42
表 四-4	WEKA J48 決策樹之參數設定.....	44
表 四-5	WEKA J48 分類結果衡量指標.....	46
表 四-6	WEKA J48 之 confusion matrix.....	46
表 四-7	WEKA BayesNet 貝氏網路之參數設定.....	48
表 四-8	WEKA BayesNet 分類結果衡量指標.....	50
表 四-9	WEKA BayesNet 之 confusion matrix.....	50
表 四-10	WEKA BayesNet 分類結果衡量指標：分類一.....	52
表 四-11	WEKA BayesNet 之 confusion matrix：分類一.....	52
表 四-12	WEKA BayesNet 分類結果衡量指標：去除錯誤資料之分類一.....	53
表 四-13	WEKA BayesNet 之 confusion matrix：去除錯誤資料之分類一.....	54
表 四-14	WEKA BayesNet 分類結果衡量指標：分類二.....	55
表 四-15	WEKA BayesNet 之 confusion matrix：分類二.....	55
表 四-16	各階段方法之 F-measure 指標比較表.....	59

# 第一章、緒論

本章將說明研究動機、研究目的、研究步驟與流程，以及論文章節安排。

## 第一節 研究動機

20 世紀是資訊社會的時代，企業依賴大量資訊進行研發、生產、銷售和服務的活動，因此，如何消化處理大量資訊即是以資訊為導向的時代中之重要課題。而管理科學下的決策理論與數量方法等領域所著重的提供問題解決方案，在問題的解決過程中，從界定問題與描述問題開始，及至建立問題的解決架構，其中牽涉的也是如何處理資料、構建資料處理模型、形成決策、支援決策制定等過程，實則也就是「將資訊轉成問題解決方案的過程」(許士軍，1982)，進而提供決策制定之參考。

進入 21 世紀，則是面臨知識暴漲的時代，社會科學和管理科學著重的是如何將資訊轉化成知識，提供企業和組織在知識經濟的時代，保有競爭優勢，佔有一席之地(Baupin and Zreik 2000)。在知識經濟的時代下，前一世紀所著重處理大量資訊，進而將資訊轉成問題解決方案，仍屬重要，但如何從資訊到知識，或是說如何進一步將資訊延展成知識，則是知識經濟時代下另一個重要的課題。

從狹義來看，將資訊轉化成知識，或者說進一步將資訊延展成知識，可以有兩種途徑，其一是從大量資料庫中，藉由人工智慧的技術或方法，例如：應用人工類神經網路 (Neural Network)、案例式推理 (Case Based Reasoning; CBR)，或是模糊規則 (Fuzzy Rules) 等，整理分析萃取出知識。其二則是更快捷的知識獲取途徑，就是直接蒐集或詢問專家之意見，經由整理分析，而成有條理的知識呈現出來(Burnside 2005)。

在實務上，藉由整合專家意見獲取快捷知識，係知識應用常見的模式。事實上，整合專家意見內容分析出來的知識，也的確應用在很多層面，諸如：技術預測、趨勢預測、未來發展預測等，也有應用計量方法預測出來的結果，再依賴專家諮詢進行預測結果的微調，著名的德爾菲調查方法（Delphi Survey）就是專門用來在其中搭配收斂專家意見的調查法。不過，這一類研究模式主要針對發展專家意見調查，再進行調查意見之內容分析而得，其中，有關專家意見內容的分析多應用規範性研究模式(Chava et al., 2003)，進行內容分析整理，並未引用量化分析方法，分析所得資料。

至於將專家調查意見內容進行計量模式處理者，例如：應用 AHP 方法、貝氏網路模式（Bayesian Network）或是貝氏整合方法等，其資料雖經計量方法分析處理，但往往所得分析的結果，僅得資訊卻失卻當中豐富的知識脈絡。如應用 AHP 分析方法得出的是最佳化方案評選結果，應用貝氏網路模式重點在於進行新案例的預測，應用貝氏整合方法則僅止於修正原先預測目標值等。基於整合專家意見係獲取快捷知識重要的途徑，那麼如何將專家意見不僅達到內容整合的效果，倘能夠應用計量方法，達到分析的目標，又保留其中重要的知識脈絡，換言之，不只能夠處理大量資訊，又能將資訊進一步延展成知識，整理出知識的脈絡，那也許將有助於增進當代社會知識之獲取。

## 第二節 研究目的

源於上述研究背景與研究動機，本論文擬根據此一研究旨趣，選取合適的應用實例，然後依照案例資料的特性，發展 IDTBN 模式（Integrated Decision Tree and Bayesian Network），應用於整合分析專家意見，希冀能夠達到以下的效果：

- 整合專家意見內容（例如：技術預測的做法）。
- 應用計量模式處理專家意見資訊（例如：貝氏網路模式或是決策樹方法等）。

➤ 將資訊延展成知識，且進一步保留其中重要的知識脈絡。

其次，更重要的是本研究針對專家意見的蒐集並不侷限於德爾菲專家問卷調查，而是透過任何資訊平台，只要有提供社群互動功能者，均可作為資料蒐集的來源。未來只要針對專家意見內容，進一步發展自動化內容擷取與分析機制，替換本研究現階段人工資料過錄作業，即可將本研究模式與社群互動資訊平台結合，達到自動化擷取與分析專家意見內容的效果，結合當前社群應用平台功能，擴大應用範圍、有效處理大量資訊，發揮本研究之潛在研究效益，這部分將在本論文結論的地方，詳加闡述。

具體而言，本論文研究的主要重點在於：

- 一、根據本論文之研究動機與研究目的，參酌相關文獻探討結果，規劃整體研究設計、流程，以及研究架構，俾便突顯本研究之價值所在。
- 二、選取台電公司長期電力負載預測之應用實例，根據應用實例之資料特性，搭配實務上待解問題的指引，結合計量方法和技術，發展 IDTBN 模式，進行個案專家整合意見之分析與處理。
- 三、以知識應用之觀點，探討本研究所提 IDTBN 模式與其他相關研究方法，如：貝氏網路模式、決策樹方法、類神經網路或案例式推理等之間的特點異同及其中的優劣。
- 四、從個案實際應用效益的觀點，比較本研究所提 IDTBN 模式與個案其他既有相關處理模式，包括：電力長期負載預測 BAU 模式之間的優劣勢分析。
- 五、驗證 IDTBN 模式在整合專家意見與萃取知識上所具有之實質效益。
- 六、以增進潛在研究效益的觀點，探討與分析本研究與專家社群互動資訊平台結合的效益。

七、凝聚本研究的心得與結果，前瞻未來相關研究之展望。

### 第三節 研究步驟與流程

本論文的研究目的主要有兩大部分，其一是構建 IDTBN 模式，提供專家意見整合分析之另一個可能處理模式；其二是以增進決策效益之觀點，將所構建的 IDTBN 模式，實際應用於個案實例上，藉以解決實務應用之需。根據上述本論文在研究上的兩大需求，有關整體研究設計，如何滿足實務應用上之需，並且進一步驗證所構建之 IDTBN 模式具備優質效益，係本論文之研究重點，也是本論文在研究設計上需要特別費心處理的部分。

基於上述，本論文在研究步驟上主要規劃七個步驟，茲分別說明如下：

#### 一、研究動機：

從新進相關研究文獻上，挖掘出研究的新議題，試圖從中尋找本論文在研究上合適之相關研究題目及其未來研究上的可能貢獻與助益。

#### 二、研究目的：

源於研究背景與研究動機，選定本論文之研究主題與目的，作為後續研究進程之方向與依據。

#### 三、研究問題與範圍界定：

在研究目的的指引下，進一步清楚定義本研究所要解決的研究議題，並界定可行的研究範圍。

#### 四、相關文獻探討：

根據研究的目的與主要研究內涵，蒐集與本研究相關之文獻，主要文獻蒐集、整理與分析的方向，包括：貝氏網路模式、決策樹方法、類神經網路、案例式推理等相關文獻探討，希冀能夠藉此彙整出較具週延的研究架構、模

式構建與研究設計，提供本研究作為後續研究的參考依據。

#### 五、研究方法：

應用實證研究的方法蒐集相關專家意見，應用個案研究的方式進行模式與分析方法之間的比較。

#### 六、選取應用實例：

根據本論文所界定之研究目的與內容，並且考量相關資料之蒐集與整理之可能支應情形，特選取台電公司長期電力負載預測應用實例，作為本論文的實務應用案例，並於此章節介紹台電公司長期電力負載預測所使用之計量方法。

#### 七、發展 IDTBN 分析模式：

根據實例應用之需，規劃可行架構，蒐集相關所需資料，發展 IDTBN 資料分析，並探討應用模式進行個案實例之分析結果。

#### 八、IDTBN 模式驗證：

驗證 IDTBN 模式所具有之實質效益。

#### 九、結論與建議

根據本論文整體研究的過程與結果，摘述主要研究發現、價值與貢獻，並進一步探討相關研究限制，以及提出未來可行的研究方向。



本論文之具體研究流程架構，則如圖 一-1 所示。

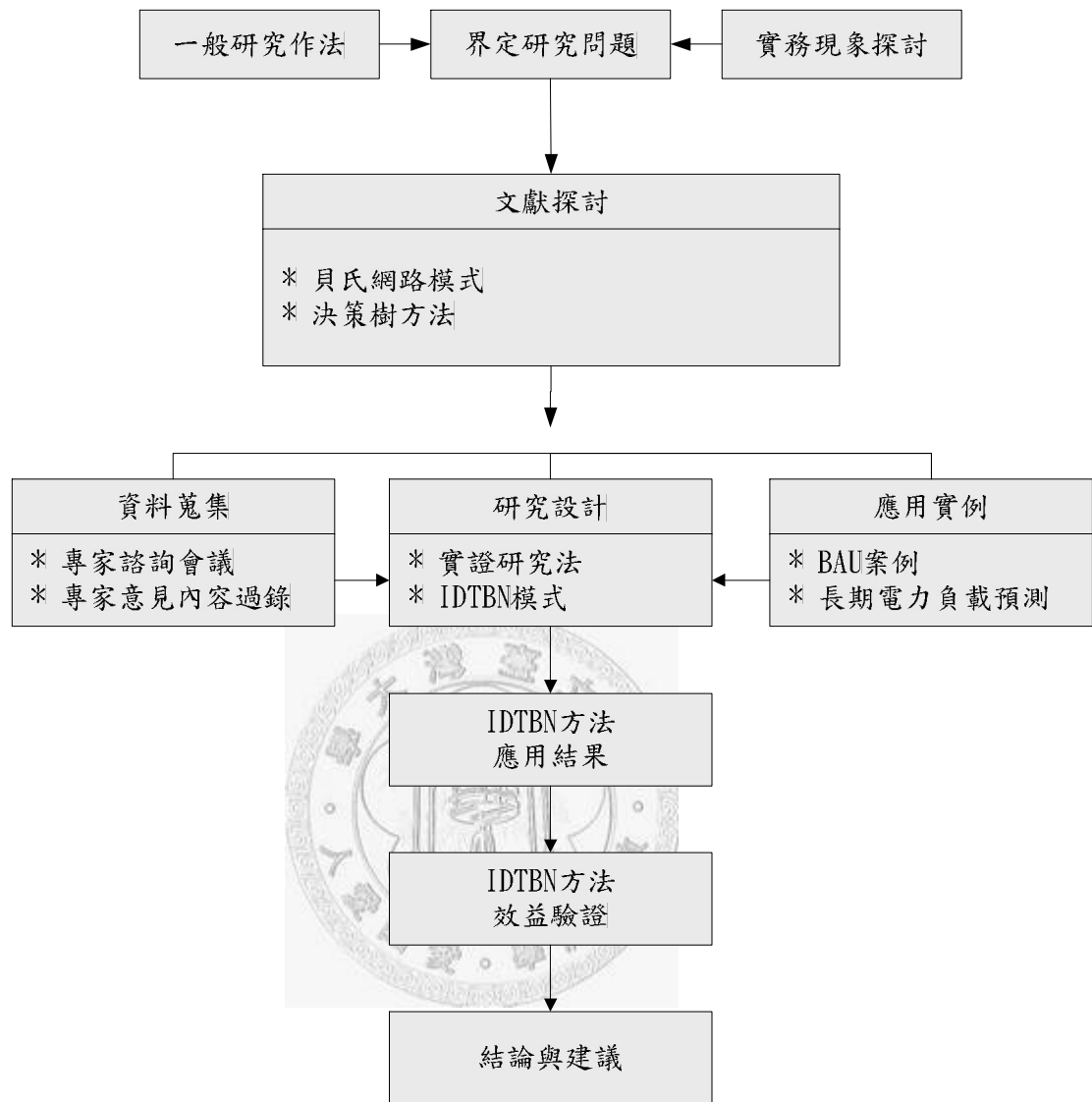


圖 一-1 本論文研究流程

(資料來源：本研究整理)

## 第四節 論文章節安排

根據前述之研究背景、研究動機，以及本論文主要研究目的與內容，本論文在章節架構之安排上共分五章，各章之大要分別為：

第一章緒論：說明研究背景、研究緣起、目的、範圍與具體研究步驟與流程。主要在於試圖表達構建 IDTBN 分析模式，對於整合專家意見、萃取知識，及其在知識經濟時代所具有的意義與效益，並說明論文整體研究方向、研究流程與可能內容。

第二章相關文獻探討：本章分就五方面領域，包括：貝氏網路模式、決策樹方法、類神經網路及案例式推理系統，進行相關文獻探討，藉此彙整出較具周延的研究架構與內容，提供本研究進行後續研究之參考依據。

第三章研究方法：本章說明本論文選取應用實例的考量、案例的具體內容、本研究所提 IDTBN 模式、實證研究方法之應用、研究設計、研究架構以及具體之專家意見蒐集與過錄的人工處理過程等。

第四章說明整體實質研究的過程與結果：先進行各別僅應用決策樹或貝氏網路分析的實驗，並進一步應用本研究所提 IDTBN 模式分析處理個案資料，各別詳述分析的過程與結果，藉此突顯 IDTBN 模式所具的效益與優勢。

第五章結論與建議：於本章中將整體論文研究的過程與結果作一綜觀，指出其中的研究價值與意義、重要研究成果、後續改進建議，並進一步探討本論文在研究上的限制以及未來可能的研究發展方向，俾有助於本論文在研究上之延伸以及供後續相關研究參考使用。

## 第二章、文獻探討

本章將對貝氏網路、決策樹、類神經網路與案例式推理等資料分析技術，進行相關文獻探討，以期對於本論文進行後續研究之參考。

第一節說明貝氏網路之構建、數學模式及其學習模式；第二節將說明決策樹之構建與修剪方法；第三節將介紹類神經網路的理論與架構，及最常使用之倒傳遞類神經網路運作方式；第四節將介紹案例式推理系統之概念及運作模式。將於第三章第二節提出本研究之 IDTBN 方法，討論在上述四個方法裡為何選擇決策樹與貝氏網路，並整合兩者之優點，使結果能優於上述方法各別作單一使用的效果。

### 第一節 決策樹方法

分類分析是從已知類別的物件實例 (Instance) 集合中，根據其屬性 (Attribute) 來建立類別 (Class) 的過程，其中物件屬性是可能影響物件類別之描述。

1. 類別：用來代表決策或所屬分類。例如，在視力檢查結果之範例中，類別可能有近視與未近視兩種，類別的所有可能值必須為有限集合的元素之一，彼此互斥。在訓練資料集中，須包含有各種類別之範例，方能獲得有效之歸納結果。
2. 屬性：用以描述範例的特徵，亦及用來分辨兩個不同的範例。屬性依其值可以分為兩類：
  - a. 類別屬性 (Categorical Attribute) 或稱名目屬性 (Nominal Attribute)：在這類屬性中，各值之間並無大小或次序間的關係。

- b. 數值屬性 (Numerical Attribute) 或稱有序屬性 (Ordered Attribute) :  
各值之間存在有大小或次序間的關係，數值屬性又分為整數屬性與實數屬性。

因此，分類分析目的是要產生一個分類模式，可用以描述物件屬性與類別之間的關係，進而對未知類別或新資料進行類別預測。而決策樹分析即為一種分類分析方法。

## 2.1.1 決策樹分析介紹

決策樹建構是利用監督式學習法，從訓練集中以適當之屬性挑選函數，從所有屬性中挑選出可用以建構決策樹根節點 (root node) 及各個內部節點 (internal nodes) 的屬性，用以建構決策樹，並對物件進行分類處理。主要優點在於可產生容易瞭解與運用的決策法則。

### 1. ID3

決策樹包括很多衍生的模型，不同模型之間除了可以判別的資料型態不同外，對於決定分類屬性先後順序的演算法也不同。ID3 為一決策樹歸納技術，在構建決策樹過程中，ID3 以資訊增益 (Information Gain) 為依據，選擇最佳屬性當成決策樹的節點，使得所建立的決策樹為一最簡單 (或接近最簡單) 的決策樹。資訊增益 (Information Gain) 是由以某一屬性為決策樹節點所產生的子決策樹之熵值 (Entropy) 與物件集合的熵值所決定 (Quinlan 1986)。ID3 選擇分類屬性的方法對於一般學習問題已經有不錯的結果，但是以資訊增益的屬性挑選準則偏好於選擇屬性值較多的屬性。倘若資料集合依據屬性  $C$  的可能值進行分類，而每個屬性  $C$  可能值的分支僅存在單一筆資料時，則該屬性  $C$  將具有最大的資訊增益；但在實際決策上，屬性  $C$  並不具決策價值。為了改進資訊增益的缺失，Quinlan 提出資訊增益率 (Gain-Ratio Measure) 的方法作為選擇適當決策屬性的根據 (Quinlan

1993)。資訊增益率方法是利用屬性資訊增益除以屬性可能值數量的正規化方法，避免決策屬性的選擇偏向於具有多個屬性值的屬性。

## 2. C4.5

但 ID3 無法處理分析性資料屬性，而 C4.5 則使用連續性數值分割方法，因此可處理分析性資料。C4.5 是目前最常用的決策樹分類法，是由學者 Quinlan 改進著名的 ID3 學習演算法而發展出的決策樹歸納學習法(Quinlan 1986)。該連續性數值分割法是：先將資料集合根據數值屬性的屬性值大小進行排序，再計算出兩個可能屬性數值的中點，稱為分割點 (Cut point)，先將資料集合分為兩部分，再以決策屬性挑選評估函數計算該分割點所得之決策屬性評估值；最後，將在眾多的分割點中選擇具有最佳屬性評估值的分割點作為該數值屬性的資料樣本區分點。

## 3. CART

能處理連續性資料的決策樹也可稱為迴歸樹 (Regression Tree)，CART (Classification And Regression Trees) 就是其中一種，對於分類屬性的先後順序，CART 採用 Gini 索引法 (Gini index)。當決策樹要預測一筆新資料時，會先從根節點進入，隨後會按照各非葉節點作分類判斷，等到分到葉節點時，則該葉節點所代表的意義，就是該筆新資料的預測結果。

## 4. CHAID

CHAID (Chi-squared Automatic Interaction Detection) 主要運用卡方檢驗 (Chi-square) 選擇具有統計顯著性區分力的屬性作為樹狀結構的內部節點。

# 2.1.2 決策樹之構建與修剪

基本的決策樹建構演算法如下：

1. 從設定決策樹的根節點 (Root Node) 開始，設定根節點為目前節點  $C$ ，此時所有訓練物件將都屬於  $C$  的物件集合。若  $C$  中所有的物件都屬於同一類別 (Class)，則將此類別設定為  $C$  之決策結果，然後停止，否則繼續執行步驟 2。
2. 對  $C$  中所有的物件，針對所有尚未出現在從根節點到目前節點路徑的所有屬性  $A_i$  (稱之為候選屬性)，利用屬性挑選函數挑選出一個具有最佳類別區分力的屬性  $A_c$  作為節點  $C$  的分類屬性。
3. 在節點  $C$ ，根據所挑選出的屬性  $A_c$  的  $m$  個屬性值，在節點  $C$  下建立子節點  $C_1$ 、 $C_2$ 、 $\dots$ 、 $C_m$ ，並根據分類屬性值將  $C$  中的所有物件分派至適當的子節點中。
4. 將每個子節點  $C_i$  當成目前節點  $C$ ，並分別由步驟 1 繼續執行決策樹的建構。

當決策樹成長完成之際，雖然能夠完整地描述訓練範例中所有的物件，但卻不夠一般化地來預測其它未觀察到的範例。因此，犧牲一些對訓練範例描述的正確性，來換取對其它眾多的未觀察範例的一般化描述，便構成決策樹修剪的原因。成本-複雜修剪法 (Cost-Complexity Pruning)、臨界值修剪法 (Critical Value Pruning)、最小誤差修剪法 (Minimum-Error Pruning)、降低誤差修剪法 (Reduced-Error Pruning)、悲觀誤差修剪法 (Pessimistic Error Pruning) 等修剪方法相互比較，其顯示前三者有較佳的結果(Mingers 1989)。

## 第二節 貝氏網路模式

貝氏網路(Bayesian Network)是一種對於不確定之事物加以描述與推論的工具 (Pearl 1988)。貝氏網路可用以表示機率之關聯性，常用於描述許多真實世界中可能面臨的問題，其中包含決策支援、問題診斷、預測、自動監控、製程控

制與資訊萃取等層面 (Heckerman, Mamdani et al. 1995)。

## 2.2.1 建構貝氏網路

貝氏網路的建立包含下列五個步驟(Enrique, Jose et al. 1997)：

1. 決定在該問題領域相關之變數的集合，如  $X = \{x_1, x_2, x_3, \dots, x_n\}$ 。
2. 將該集合中的變數做適當排列。
3. 由變數集合中選擇一變數  $x_i$  作為貝氏網路中對應的一個節點。
4. 在已建立的網路中，針對  $x_i$  尋找出集合： $Parents(x_i)$  成為  $x_i$  的『因』，並確保網路中條件獨立 (Conditional Independence) 的情況。
5. 定義出  $x_i$  的條件機率表。

貝氏網路具備下列三項優點(Heckerman 1996) (Shachter 1988)：

1. 貝氏網路結合貝氏統計方法能輕易整合領域知識 (domain knowledge) 與資料 (data)。在現實世界中，系統建構者重視與問題有關之領域知識，尤其與此問題相關的資料量很龐大或昂貴時。貝氏網路能將關於問題的知識以因果關係表示之，並以機率表示因果關係的強度。
2. 貝氏網路允許使用者對因果關係進行學習。學習能力乃貝氏網路一項極重要的能力，特別是當網路建構者對欲求解問題的領域知識不足時。
3. 貝氏網路可處理資料不完整的問題。貝氏網路在建構之時，若建構網路時資料不完整，使用者可以主觀的設定未知的參數值，如事前機率 (prior probability) 等，以彌補資料的完整性。

## 2.2.2 貝氏網路之數學模式與架構

貝氏網路包含了一組隨機變數相互之間的關連性與聯合機率分配，其中每個變數可能是有限的離散變數或是連續變數(Heckerman and Breese 1995)。貝氏網路由問題的領域知識與資料所構成，因此若以  $B$  表示貝氏網路，則  $B = (D, P)$ 。其中  $D$  表示一有方向性的非循環圖形(directed acyclic graph ; DAG)，即表示問題的領域知識，此外，網路中不允許有向性循環(directed cycles)存在。 $P$  為 DAG 中某一變數的聯合機率分配(joint probability distribution ; JPD)，即表示關於問題的資料。如圖 二-1 所示，其中每節點(node)代表一個隨機變數，節點間的連線則代表隨機變數之間所存在的因果關係，及彼此相互獨立的假設，而獨立假設可決定於隨機變數之間的機率分布，亦即所需要的機率資訊 (Charniak 1991)。

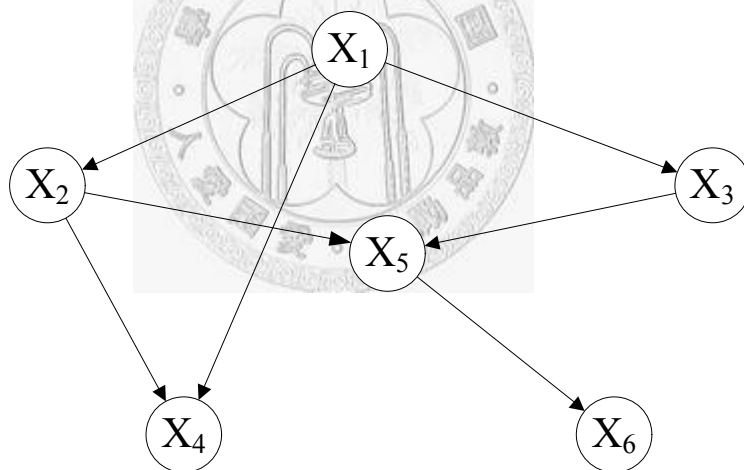


圖 二-1 DAG 圖形

貝氏網路中一個問題包含一組變數  $X = \{x_1, \dots, x_n\}$ ，以及一組條件機率分配關係，這包含一些條件相關(Conditional Dependent)與條件獨立(Conditional Independent)的假設或認知，這讓我們得以建構出  $X$  的聯合機率分配函數，根據機率的連鎖律(Chain Rule)，我們將之表示成為：

$$p(x_1, \dots, x_n) = \prod_{i=1}^n p(x_i | x_1, \dots, x_{i-1}) \quad (1)$$



對於每一個  $x_i$ ，令  $\Pi_i \subseteq \{x_1, \dots, x_{i-1}\}$  為一個表示  $\{x_1, \dots, x_{i-1}\}$  中與  $x_i$  相關的子集合，因此我們得到：

$$p(x_i | x_1, \dots, x_{i-1}) = p(x_i | \Pi_i) \quad (2)$$

其概念是將  $x_i$  的分配用  $\Pi_i$  取代  $\{x_1, \dots, x_{i-1}\}$  去描述之，如此一來問題就被簡化了。有了這些關係，貝氏網路可以將變數的關連性以一個具方向性的、非循環性的圖形去描述，其中每一個  $x_1, \dots, x_n$  分別對應到圖中的一個節點(Node)，而該節點所有父節點的集合即是  $\Pi_i$ 。在之後的論述中，我們將以  $x_i$  同時表示該節點與該節點所對應的變數。

結合式(1)與式(2)，我們得到貝氏網路聯合機率的基本表示法：

$$p(x_1, \dots, x_n) = \prod_{i=1}^n p(x_i | \Pi_i) \quad (3)$$

因此，圖 二-1 之影響程度可藉由條件機率來表達，如下式所示：

$$\begin{aligned} & P(X_1, X_2, X_3, X_4, X_5, X_6) \\ &= P(X_6 | X_5)P(X_5 | X_2, X_3)P(X_4 | X_1, X_2)P(X_3 | X_1)P(X_2 | X_1)P(X_1) \end{aligned}$$

這是最基本的架構，其中變數的順序安排必須經過適當的選擇，若選擇不當，則最後產生的結果將無法成功的利用條件獨立所帶來的簡化效果。在實作上，專家們通常可以輕易地對於變數間之因果關係與條件相關性做出假設，而我們可以直接利用這些假設去建構出貝氏網路而不用先去排列這些變數 (Heckerman 1996)。

貝氏網路中的因果關係是推測性而非決定性的。在貝氏網路中，沒有父節點的節點須指定其事前機率，而每一個擁有父節點的節點都有一個條件機率表，用以表示其直接影響的父節點之間的可能組合。

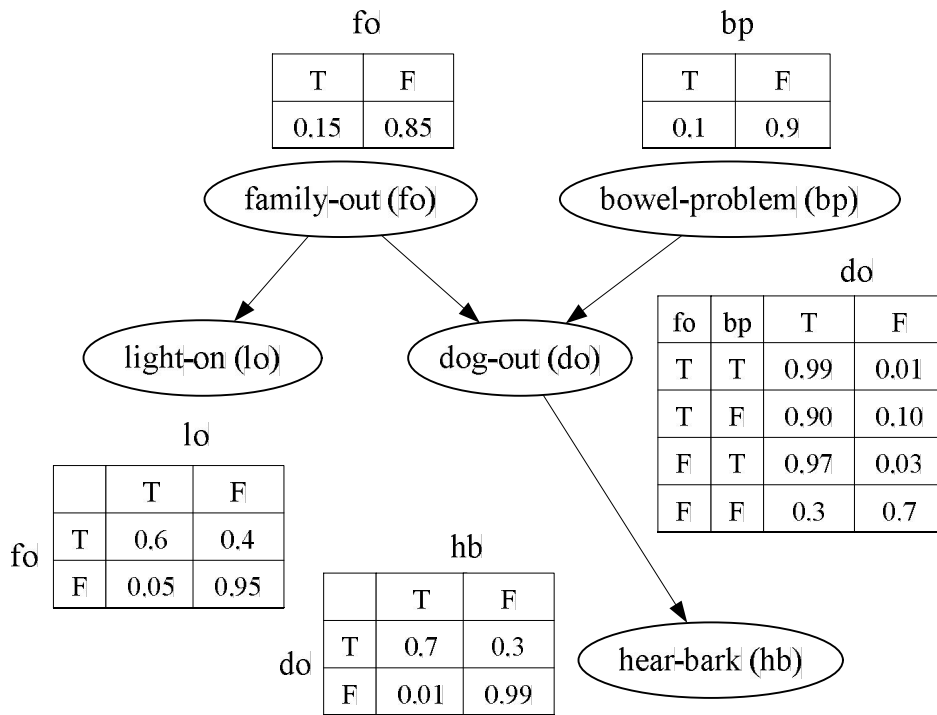


圖 二-2 貝氏網路範例：family-out 問題

(資料來源：Charniak, 1991)

圖 二-2 是一個包含機率表的範例 (Charniak 1991)。這是一個用以推測「家人是否在家」的問題，首先要決定問題的變數：family-out (家人不在家)、bowel-problem (狗的排便問題)、dog-out (狗在房子外面)、hear-bark (聽到狗叫聲)、light-on (門外燈亮著)。其中 family-out 代表家人是否出門；bowel-problem 表示狗是否想要排便；dog-out 表示狗是否在房子外面；hear-bark 表示是否聽到自家狗吠的聲音；light-on 表示門外燈是否亮著。每個隨機變數都擁有其事前機率表或條件機率表來表示其與相鄰節點之間的影响關係。

根據經驗，我們假設家人是否出門和狗的排便問題，與狗是否在房子外面是相關的，同時家人是否出門也會影响門外燈是否亮著，而狗是否在房子外面也影响著是否聽到自家狗吠的聲音之可能性。依據過去累積的記錄可以得到上述變數的條件機率，以建立圖 二-2 中的條件機率表和貝氏網路結構。

另一個以貝氏網路解決問題的例子是關於信用卡盜刷的問題。首先定義問

題的變數如下：盜刷 (F)、加油 (G)、買珠寶 (J)、年齡 (A)、以及性別 (S) 等。其中 F 表示該筆交易是否為盜刷；G 表示在之前 24 小時之內是否有去加油；J 表示在之前 24 小時之內是否有去買珠寶首飾；年齡 A 在此可以為一連續變數，表示該信用卡持有者之年齡；性別 S 表示該持有者為男性或女性。(Heckerman 1996)

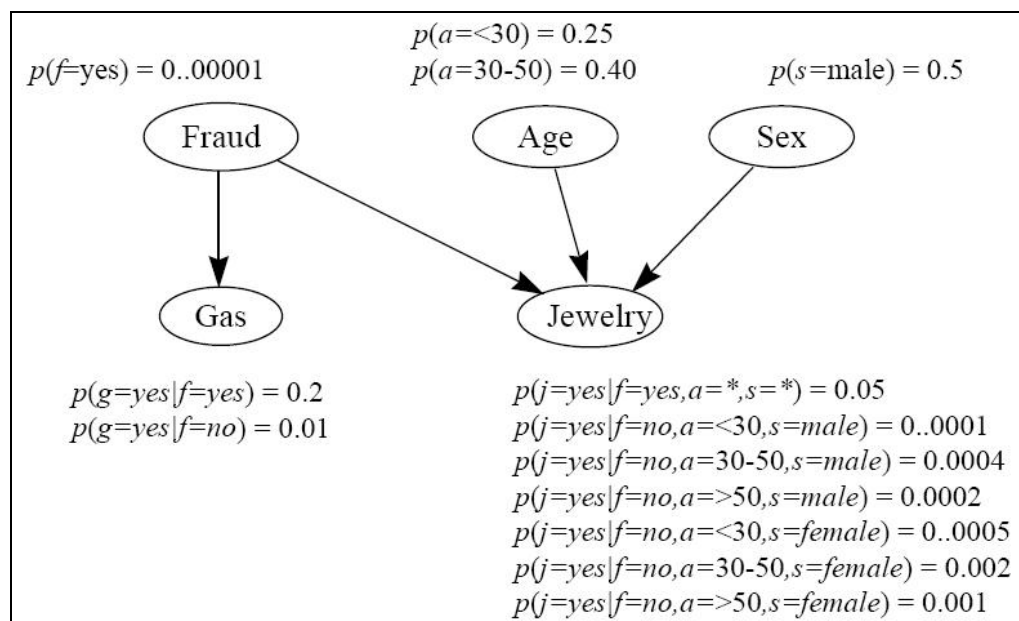


圖 二-3 貝氏網路範例：信用卡盜刷問題

(資料來源：Heckerman, 1996)

根據經驗，我們假設是否為盜刷、持卡者年齡與性別是與其購買珠寶之慾望和需求相關的，同時是否盜刷也影響著是否有加油的行為。根據這個關係，我們可以從以往信用卡消費的統計資料中得到相關的條件機率，藉此我們建立了圖 2-2 的貝氏網路。注意，在這個模型中，我們同時意涵著以下的條件獨立性：

$$\begin{aligned}
 p(a|f) &= p(a) \\
 p(s|f, a) &= p(s) \\
 p(g|f, a, s) &= p(g|f) \\
 p(j|f, a, s, g) &= p(j|f, a, s)
 \end{aligned}
 \tag{4}$$

由於貝氏網路給我們的是整個問題的聯合機率，因此我們可以藉由機率的定理來求得任何我們想知道的條件機率。舉例來說，我們想知道某一筆交易是否

為盜刷，我們可以這樣計算：

$$p(j|f, a, s, g) = \frac{p(f, a, s, g, j)}{p(a, s, g, j)} = \frac{p(f, a, s, g, j)}{\sum_{f'} p(f', a, s, g, j)} \quad (5)$$

然後根據方程式 3 的定義，方程式 5 可以再展開成為：

$$\begin{aligned} p(j|f, a, s, g) &= \frac{p(f)p(a)p(s)p(g|f)p(j|f, a, s)}{\sum_{f'} p(f')p(a)p(s)p(g|f')p(j|f', a, s)} \\ &= \frac{p(f)p(g|f)p(j|f, a, s)}{\sum_{f'} p(f')p(g|f')p(j|f', a, s)} \end{aligned} \quad (6)$$

### 2.2.3 貝氏網路之學習模式

貝氏網路的學習模式可分為 Structural learning 與 Parametric learning 兩種 (Enrique, Jose et al. 1997)。此兩種學習模式皆需要評量網路結構品質的函數 (quality measure) 以及搜尋演算法 (search algorithm)。傳統貝氏網路的建構，通常是由問題領域的專家依據經驗法則或問題領域特有的知識，建立出變數與變數之間的關聯。但若遇上較複雜之問題領域，結構關聯性之建立的複雜度則會大幅增加。為解決複雜度的問題，多數學者傾向利用搜尋演算法將資料隱含的結構建立出來，同時搭配一個評量網路結構品質的函數，以選出與資料最為相符合的結構。(范樹根，2005)

一般貝氏網路學習模式需要下列三個步驟：(Enrique, Jose et al. 1997)

1. 選擇 quality measure 與 search algorithm。
2. 利用 search algorithm 找出一些高品質的貝氏網路子集合。這步驟需要估計方法來估計參數，並以評估方法來評估所選之子集合中所有的貝氏網路。
3. 由前一步驟中的子集合找出品質最高的網路結構。

利用 search algorithm 建立貝氏網路的結構主要有兩大類的方法。一是限制基礎 (Constraint-Based) 的演算法，此種方法主要是以所有結點全部連結 (Fully-Connected) 的貝氏網路為初始結構，經由兩兩變數節點間條件獨立測試，消除沒有關係的連結，並給予一個衡量貝氏網路品質的方程式，最終得到與資料最為相符的貝氏網路結構。另一類的搜尋演算法則是 Search and Score，此種方法則是搜尋變數間所有可能形成的貝氏網路，並給予一個衡量貝氏網路品質的方程式，搜尋出品質最好的貝氏網路，常見的方法有 K2 Algorithm 與 Structural EM Algorithm 等。

quality measure 是當面對一群貝氏網路時，用以評估網路是否適合欲解之問題的準則，並可對這些網路依評估結果好壞作排序。quality measure 由三個元素所構成，其構成式子如下：

$$Q = f(\text{prior information}) + g(\text{available data}) + h(\text{network complexity})$$

一般常見的 quality measure 方法如下列三項：(Enrique, Jose et al. 1997)

- The Guiger and Hererman measure
- The Cooper-Herskovits measure
- The standard Bayesian measure

參數學習基本架構如下：



根據貝氏定理，每一步驟的事前機率 (prior) 與可能性 (likelihood) 加以計算得到事後機率 (posterior)。再以該事後機率作為下一步驟的事前機率，持續循環。

在參數的學習方面，最通用的方法則是運用 EM algorithm。EM algorithm 主

要分為兩步驟：

- 在 E 步驟中為運用現有的資料中的預估參數去補足不完整的資料；
- 在 M 步驟則是運用補足的資料去求得與資料最相符的參數，如此 E 步驟、M 步驟反覆計算，求得最符合資料的條件機率分配。

## 第三節 類神經網路

人類腦內複雜的神經網路組織，經由神經細胞的突觸構造 (synapses) 作彼此之間的訊息傳導，於是人類得以對許多變化作不同反應。從神經網路所得到的靈感，人工智慧學者嘗試著來模擬仿造人類大腦的組織及運作方式，以求增進電腦的能力，使之能夠具有思考、判斷、組織等能力。近年來由於倒傳遞類神經網路的出現及資訊科技的進步，使類神經網路無論在學術研究或其他應用領域皆有不錯的表現。

### 2.3.1 類神經網路理論及架構

類神經網路 (Artificial Neural Networks) 是一種平行分散式處理計算模式。其基本的運作原理是以大量、簡單的處理單元互相連接，藉由整體處理單元對外界輸入訊號的簡單運算來處理資訊，擁有類似於人腦的許多特性及優點。與傳統的專家系統需要事先定義好規則不同，類神經網路乃藉由自我學習來處理問題。對於類神經網路這個以神經網路為背景所發展出來的模式，經由不斷強化它的方法，盡可能達到人類所擁有的學習能力 (Learning Ability)、適應力 (Adaptivity)、前後相關資訊處理能力 (Inherent Contextual Information Processing)、容錯能力 (Fault Tolerance) 等。

然而經由許多研究證實，類神經網路在許多方面具有以往電腦系統所未能

擁有的特性，可應用於許多以往的電腦系統所未能達到的領域，例如：財務預測、樣式辨認等。除此之外，類神經網路更具有對於時間序列模型與預測以及無母數估計等等的潛在能力(Kuan and White 1994)。

類神經網路是由許多處理單元(類神經元)組成，各處理單元結構如圖 二-4 所示。它包括下面幾個重要概念(Berry and Linoff 1997)：

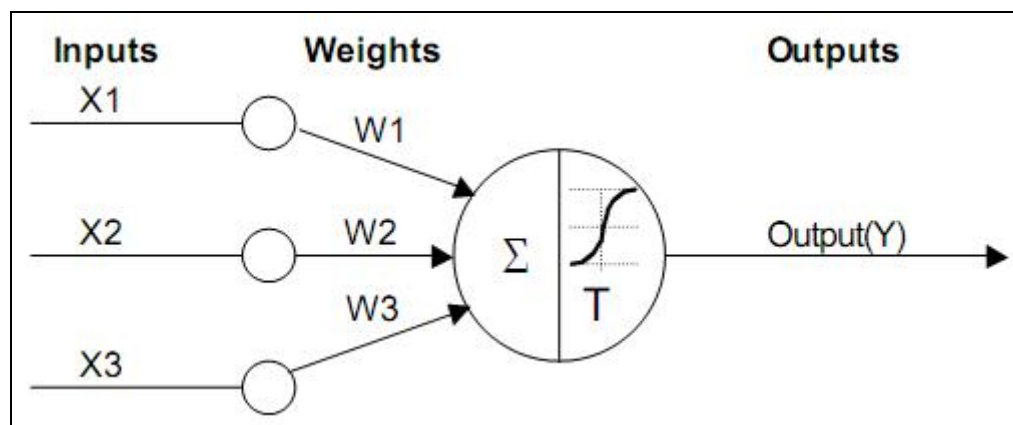


圖 二-4 類神經網路單元之模型

(資料來源：張振魁，2000)

- 輸入單元 (inputs unit)：來自外界環境或其他處理單元，每個輸入單元均對應於一個實際的屬性。輸入資料須利用前處理 (pre-processing) 將其正規化至 0 至 1 之間。
- 輸出單元 (outputs unit)：類神經網路結果輸出，成為某一特定問題之解。也可能輸出至其他處理單元，成為該處理單元輸入訊號的一部份。
- 權重 (weights)：代表前一層或外界輸入對該處理單元的相對強度。權重在類神經網路中相當重要，因為網路的學習基本上即是透過不斷地調整權重以減少誤差。
- 結合函數  $\Sigma$  (combination function)：將所有的輸入單元 ( $X_i$ ) 合併成一個單一的值 ( $Y_i$ )：
$$Y_i = \sum_{i=1}^n X_i W_i$$
。函數有一內定的門檻值， $Y_i$  越接近

此門檻值，則對輸出單元有較強烈的影響。

- 轉換函數  $T$  (transfer function)：將結合函數合併的單一值  $Y_i$  轉換為輸出訊號。最常見的轉換函數為 S 形的函數 (Sigmoid function)。

$$\text{Sigmoid} = f(x) = \frac{1}{1+e^{-x}}$$

一個類神經網路是由許多上述的處理單元所組成，並可以組成各種不同的網路模式 (張振魁，2000)。

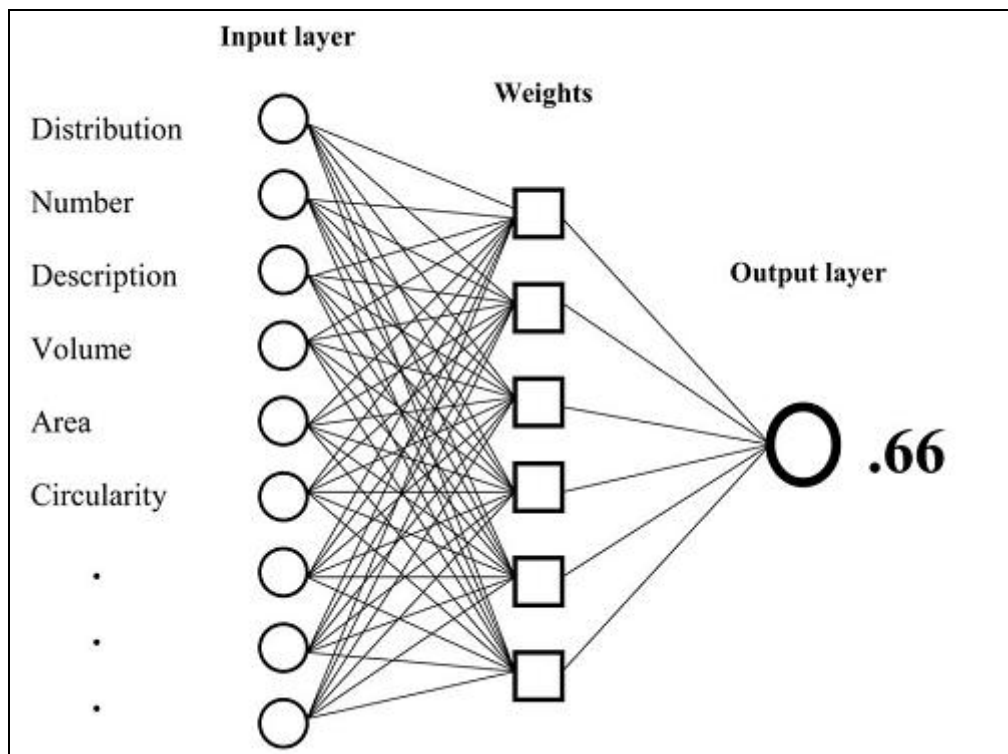


圖 二-5 類神經網路的範例

(資料來源：Burnside，2005)



### 2.3.2 倒傳遞類神經網路

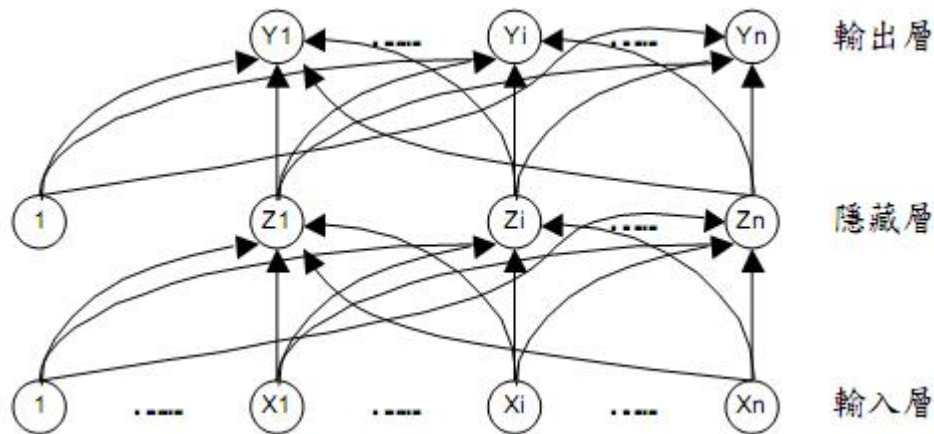


圖 二-6 標準倒傳遞類神經網路

(張振魁, 2000)

在眾多的類神經網路模型中，以倒傳遞演算法的監督式前饋網路應用最為廣泛，常用的非線性轉換函數為雙彎曲函數和雙曲線正切函數，適合診斷、預測等應用。倒傳遞類神經網路 (Back-Propagation Neural Network, BPN) 是 1986 年由 McClelland 和 Rumelhart 兩位學者共同發表，為層狀前饋式網路架構 (Layered Feedforward Network)，其學習演算法乃採用倒傳遞式學習演算法 (Backpropagation Learning Algorithm)，屬於監督式學習網路，轉換函數為 S 形的函數 (Sigmoid function)。

倒傳遞演算法是一種監督式學習的類神經網路，它會將錯誤的訊號反饋回來，以便及時修正權重。在利用訓練樣本「訓練」網路時，其實就是在調整網路權重，使得之後的樣本輸入至網路後，可以得到想要的結果輸出。

其基本原理是利用最陡坡降法 (the gradient steepest descent method) 的觀念，向後傳回輸出誤差，將誤差函數予以最小化，使輸出的預測值得以逐漸接近實際值。在學習階段裡會反覆執行兩個步驟：

1. 向前階段 (Forward phase)：網路系統會對於每一個供學習用的外界資料，執行前饋的動作而依次地計算出所有隱藏節點和輸出節點的對應激發狀態值；然後，對於每個輸出節點，其實際輸入值會與期望輸出值作比較，以得到該樣本誤差訊號。
2. 向後階段 (Backward phase)：第二個步驟則牽涉到網路系統的誤差訊號後向傳遞 (Backward Pass) 的運作現象。在輸出層所產生的誤差訊號會倒傳遞到每個隱藏節點，並且會適當地調整在輸出節點和隱藏節點之間的權重 (Weight) 以及輸出節點之間的偏倚值 (Bias)；然後，這個誤差訊號會更進一步地傳遞到每個輸入節點，同樣地，在隱藏節點和輸入節點之間的權重以及隱藏節點之間的偏倚值也會被調整，使網路向理想狀態收斂，縮小輸出的預測和實際值誤差。

以上稱為一個學習循環 (learning cycle)，可以設定訓練樣本反覆學習數個學習循環，直至收斂狀況為止 (林蔓蓁，1993)。但有時為求網路的推廣，訓練循環不宜設得太長久，以免權重過度符合訓練樣本的特性，新的樣本輸入時無法正確推導。

倒傳遞類神經網路的優點如下： (楊孟龍，2000) (葉怡成，1997)

1. 輸出值為連續，適用樣本辨識、分類、函數合成、雜訊過濾等，應用範圍最廣。
2. 學習精確度高，處理複雜的樣本識別與高度非線性的函數合成問題。
3. 回想速度快。
4. 可以建構非線性的模型。
5. 有良好的推廣性，對於未知的輸入亦可得到正確的輸出。
6. 可以接受不同種類的變數作為輸入，適應性強。

7. 可應用的領域相當廣泛。

倒傳遞類神經網路的缺點如下：(葉怡成，1997)

1. 學習速度慢且執行時間太長，訓練循環動輒上千次或萬次，需要等待時間。
2. 可能會遭遇到處理單元飽和的問題
3. 局部最小值，只找到局部最佳解，而非理想最佳解。
4. 網路架構和網路動態的決定尚缺乏非常系統化的方法。
5. 以迭代方式更新鍵結值與閾值，計算量大，相當耗費電腦資源。
6. 其解有無限多組，無法得知哪一組的解為最佳解。
7. 訓練過程中無法得知需要多少神經元個數，太多或太少的神經元均會影響系統的準確性，因此往往需以試誤的方式得到適當的神經元個數。
8. 輸入變數的和輸出結果需自行尋找，若二者因果關係過於薄弱，則更改多樣網路架構、網路動態亦難以收斂。

## 第四節 案例式推理

案例式推理 (Case-Based Reasoning, CBR) 為近年於人工智慧領域所發展的一種推理方法，發展起源於美國 1980 年代末期，繼之在歐洲興起，過去大家都是使用人工智慧領域中專家系統 (Rule-Based Expert System) 來解決問題，但在專家系統中存有明顯的限制，例如：規則難以獲取、專業知識難以清楚表達、程式設計者無法清楚了解問題等，為了避免發生上述問題重覆發生，而有新的研究方法出現，因此有了案例式推理方法的誕生，利用案例 (Case) 以取代規則 (Rule)，並隨著人工智慧領域的蓬勃發展，案例式推理愈來愈受重視，且已被

廣泛應用於各領域之中。

## 2.4.1 案例式推理概念

案例式推理的基本概念是在解決問題的過程中，當決策者遭遇到新的問題時，充分運用過去的經驗，運用相似性 (Similarity) 比對，尋求以過去的經驗模式，以作為解決問題的思考方向與解決方式，或針對過去的案例再做進一步的修正，以運用到目前決策的行為，稱之為案例式推理。

案例表達主要可分二大部分，如圖 二-7：問題描述 (Problem Description) 和結果儲存 (Stored Solution)，資料擷取是透過回答問題的方式，以產生新的解決方法(Kolodner 1993)。

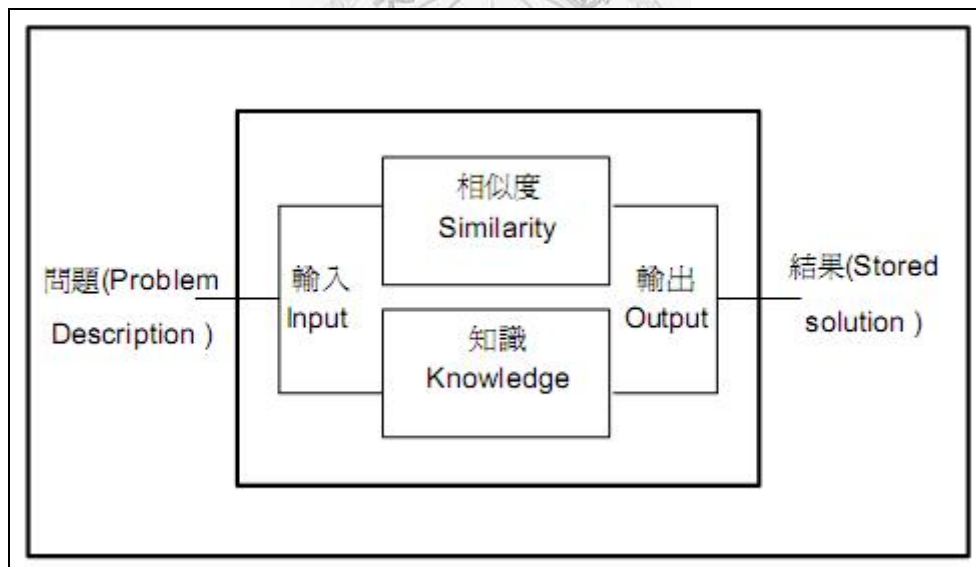


圖 二-7 案例表達方式

( 資料來源：Kolonder，1993 )

案例式推理完全以目前的問題與過去的案例的相似比較為基礎，這可以大幅免除一般人工智慧技術在知識抽取 (Knowledge Elicitation) 時所面臨的瓶頸 (Shi and Yeh 1999)。圖 二-8 是 CBR 系統的推論流程，圖 二-9 是 CBR 系統的一個範例。

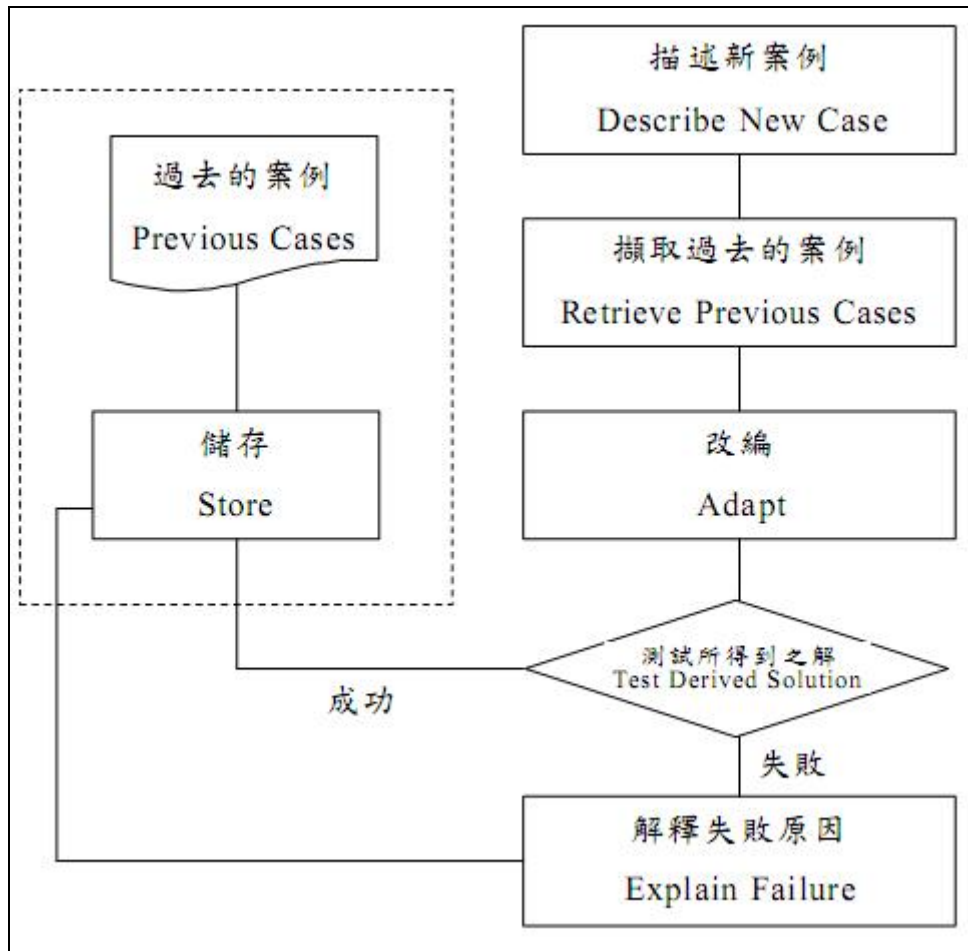


圖 二-8 CBR 系統推論流程

(資料來源：Montazemi & Gupta, 1996)

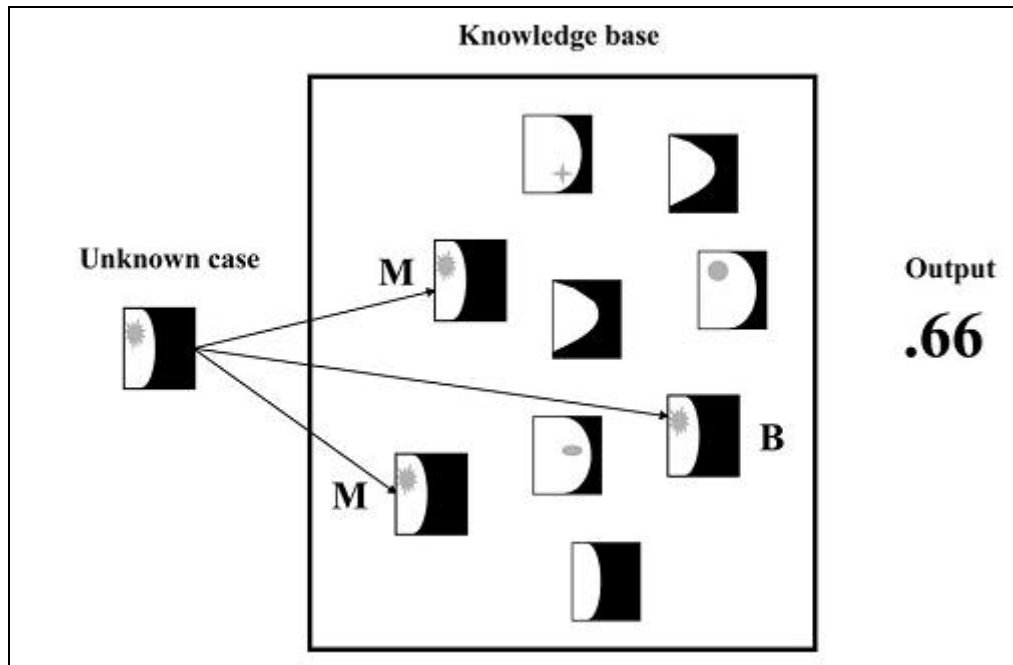


圖 二-9 案例式推理系統之範例

(資料來源：Burnside, 2005)

案例式推理和傳統的方法不同，使用案例式推理時不需要知道問題解決方法，只需要去組織與蒐集我們過去所做過相似的例子，並且不需再去引出專家所提出的理論與研究。

描述與使用案例式推理機制時，我們必然會使用到以下四個動作，資料擷取(retrieve)、再使用(reuse)、修正(revise)與保留(retain)。

1. 資料擷取 (retrieve)：運用相似性比對從資料庫中擷取和過去最相似的案例。
2. 再利用 (reuse)：運用相似案例解決方法來處理目前的問題。
3. 修正 (revise)：如果過去相似的案例不完全符合現今的問題，可以過去的案例為基礎或參考，修正解決方法。
4. 保留 (retain)：修正後的案例保存於資料庫中成為新的案例。

## 第三章、研究方法

如前所述，本研究重點在於發展一新的量化分析模式，藉以整合處理專家意見，進而達到將資訊進一步延展成知識，彰顯其中重要的知識脈絡。在此研究目的下，本章根據前一章文獻分析的結果以及相關研究流程的指引，規劃後續研究所需之研究方法內涵。第一節首先說明選取的個案實例及其內涵；第二節介紹本研究所提出的 IDTBN 模式之方法架構；第三節闡述本研究所應用的實證研究方法與研究設計內容；第四節則說明本研究在研究設計規劃下，整體的研究架構內涵；最後，第五節詳述本研究透過實證研究方法的應用，蒐集整理與過錄（coding）專家意見內容的過程。

### 第一節 應用實例—台電公司電力長期負載預測

在選取個案實例上，本研究考量的標準有二，其一是個案實例能夠符合本研究的目的所需，其二是個案實例本身亦具有實質應用上的價值；如此一來，不僅符合研究創新之學術價值需求，研究結果也同時可以作為提供實務應用參考。基於上述目的，本研究特選取「台電長期電力負載預測」作為應用實例。以下介紹個案既有長期預測作業實施情形。

就個案公司而言，在現有體制規範下，台電的業務經營範圍涵蓋發電、輸電、配電與售電，獨占整個電力市場，一個台電公司即等同於整個電力產業。目前，台電擁有逾 1,000 萬用戶，24 個區營業處（包括台灣本島及澎湖），並負擔所有用戶之供電義務與安全。從負有供電義務來看，個案公司一向以供給面資源進行長期電力負載預測，其誤差水準低於 3%（賴正文，2001），符合學理以及實務之需求。然而，自從 2000 年始，台灣出現第一次經濟成長率為負，當年度台灣三大經濟研究機構，包括：中華經濟研究院、台灣經濟研究院、台灣綜合研究院，以及中央研究院經濟研究所、政府主計處、經建會等，沒有任何一個經濟研究部

門精準預測到當年度的經濟成長率。自那時起，大家普遍驚覺到台灣社會發展正面臨結構轉換或劇變的來臨。個案公司進行長期電力負載預測之作業，由於預測期間長達 15 至 20 年，其受到的整體經濟社會結構轉變的影響，衝擊更大。

以個案公司來說，進行長期電力負載預測作業主要的目標在於提供未來電源開發、網路規劃、電業經營管理、電價結構及需求面管理之重要參考依據，期使電力供需得以整合。特別因為個案公司擁有的全台灣的電力系統，係獨立於其他國家電力網路之外，一旦發生電力短缺，無法自他國進口電力以供奧援；此外，過剩的電力供給則將嚴重浪費珍貴的能源資源，對於自產能源不足的台灣，以及現階段各種燃料價格不斷節節升高之際，個案公司終將難以支應。因此，在面臨台灣經濟社會結構轉變之際，如何在既有以供給面資源為預測基礎的作業模式下，納入各方影響因素之考量，特別是針對未來長期需求面可能的各種發展狀況，包括：經濟成長趨勢、環保議題、社會人口成長、產業發展、重大投資案等各式會影響未來電力負載的因素等，納入更精確的考慮，進行合宜之長期負載預測，作為電力供給資源規劃之基礎，可能是現階段在台灣有關電力事業發展一項重要的課題。

以長期電力負載預測作業模式之發展來看，鑑於電力為經濟發展之基礎，並與國計民生息息相關，因此，合宜的長期負載預測實與經濟、產業、人口、電價、氣溫及需求面管理等諸種變動條件，關係密切。在 1970 年代以前，世界各國之電力公司常使用簡單統計方法來進行負載預測，如：趨勢分析法、外插法等（許哲強，2002）（賴正文，2001）。其後各類方法逐漸發展，包括：計量經濟模式 (econometric model)、時間數列分析 (time series analysis)、人工類神經模型等（許哲強，2002）（賴正文，2001）。常用支援進行長期電力負載預測之計量方法比較如表 三-1 所示。一般而言，應用各種計量分析預測方法，只要預測誤差在 10% 以下（賴正文，2001），均可接受；而個案公司的長期電力負載預測誤差更達 4% 以下，詳見表 三-2 所示。個案公司面臨的問題不是計量分析預測方法選用的問



題，而是如何面對未來社會經濟結構劇變的問題，傳統的計量分析模式並無法涵蓋解決此一議題。

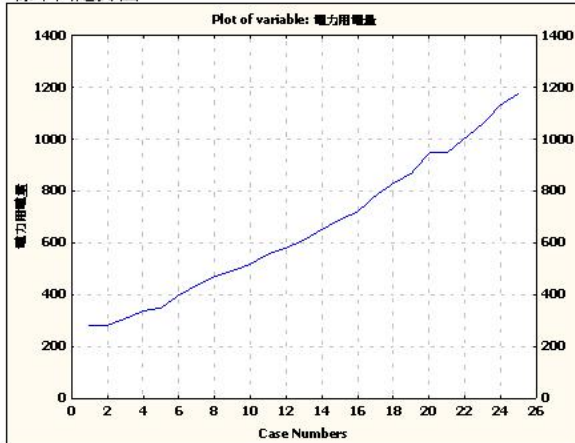
預測方法	特色	限制
趨勢分析法	<ul style="list-style-type: none"> <li>• 模型簡易。</li> </ul>	<ul style="list-style-type: none"> <li>• 難以反映複雜現象。</li> </ul>
時間序列法	<ul style="list-style-type: none"> <li>• 毋需加入外生變數，即可建構模型。</li> <li>• 可掌握週期性，季節性及循環性之變動趨勢。</li> </ul>	<ul style="list-style-type: none"> <li>• 模型選擇需高度技巧與經驗。</li> </ul>
傳統迴歸模式	<ul style="list-style-type: none"> <li>• 模型變數間具有因果關係，易於解釋。</li> </ul>	<ul style="list-style-type: none"> <li>• 較難處理自我相關問題。</li> </ul>
動態迴歸模式	<ul style="list-style-type: none"> <li>• 具傳統迴歸分析與時間序列分析整合之效果。</li> </ul>	<ul style="list-style-type: none"> <li>• 模型估計參數固定，降低模型對外來的衝擊反應能力。</li> </ul>
灰色預測	<ul style="list-style-type: none"> <li>• 模型操作簡單，原始數列操作較少。</li> </ul>	<ul style="list-style-type: none"> <li>• 不適合作長期預測。</li> </ul>
類神經網路	<ul style="list-style-type: none"> <li>• 適用於各種型態的輸出輸入對映關係。</li> </ul>	<ul style="list-style-type: none"> <li>• 容易產生樣本內資訊過度配適的問題。</li> </ul>
投入產出模型 (產業關聯分析)	<ul style="list-style-type: none"> <li>• 一般均衡分析可對所有相關產業部門作全面投入與產出的關聯分析。</li> </ul>	<ul style="list-style-type: none"> <li>• 模型所需資料複雜。</li> <li>• 預測時間較短。</li> </ul>
混合模型	<ul style="list-style-type: none"> <li>• 具擷長補短，增加模型應用上的彈性。</li> </ul>	<ul style="list-style-type: none"> <li>• 建模成本高，求解不易。</li> </ul>
終端需求模型	<ul style="list-style-type: none"> <li>• 模型結合工程法，可有效反映未來結構變遷。</li> </ul>	<ul style="list-style-type: none"> <li>• 建模成本高，模型更新不易。</li> </ul>
脊迴歸	<ul style="list-style-type: none"> <li>• 可解決變數之間共變的效應問題。</li> </ul>	<ul style="list-style-type: none"> <li>• 模型結構複雜</li> <li>• 缺乏明顯的經濟涵意。</li> </ul>

表 三-1 常用支援進行長期電力負載預測之計量方法比較

(資料來源：台電公司)

Input: 電力用電量 (負載預測歷史資料彙整表(94).sta)						
Transformations: D(1)						
Model:(1,1,0) MS Residual= 400.61						
Paramet.	Param.	Asympt. Std.Err.	Asympt. t( 22)	p	Lower 95% Conf	Upper 95% Conf
Constant	37.32678	4.091397	9.123236	0.000000	28.84174	45.81182
p(1)	-0.02513	0.219060	-0.114698	0.909725	-0.47943	0.42918

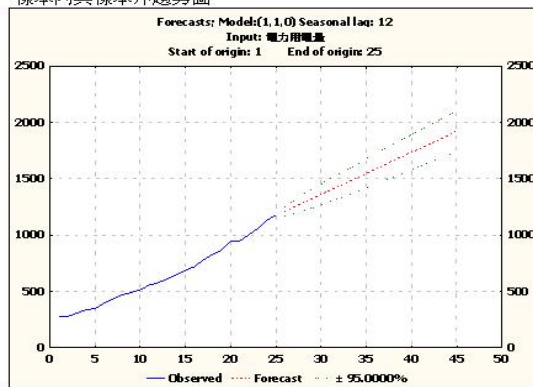
樣本內趨勢圖



樣本內預測

	Observed Value	Predicted Value	MAPE(%)
75	398.40	387.24	2.80
76	436.39	424.55	2.71
77	470.30	461.87	1.79
78	492.10	499.20	1.44
79	518.41	536.53	3.49
80	557.17	573.85	2.99
81	580.41	611.18	5.30
82	611.56	648.51	6.04
83	652.14	685.83	5.17
84	689.31	723.16	4.91
85	721.14	760.49	5.46
86	782.38	797.82	1.97
87	830.49	835.14	0.56
88	868.44	872.47	0.46
89	947.68	909.80	4.00
90	948.07	947.12	0.10
91	1004.35	984.45	1.98
92	1060.43	1021.78	3.65
93	1132.93	1059.10	6.52
94	1176.48	1096.43	6.80
Average			3.41

樣本內與樣本外趨勢圖



樣本外預測

	Forecast	Lower 95.0000%	Upper 95.0000%	Std.Err.
95	1213.65	1172.14	1255.16	20.02
96	1250.98	1193.01	1308.95	27.95
97	1288.31	1217.60	1359.02	34.10
98	1325.63	1244.15	1407.12	39.29
99	1362.96	1271.97	1453.95	43.87
100	1400.29	1300.69	1499.88	48.02
101	1437.61	1330.10	1545.12	51.84
102	1474.94	1360.06	1589.82	55.40
103	1512.27	1390.46	1634.08	58.74
104	1549.59	1421.23	1677.96	61.90
105	1586.92	1452.32	1721.52	64.90
106	1624.25	1483.69	1764.81	67.78
107	1661.58	1515.30	1807.85	70.53
108	1698.90	1547.13	1850.68	73.18
109	1736.23	1579.15	1893.31	75.74
110	1773.56	1611.34	1935.77	78.22
111	1810.88	1643.69	1978.08	80.62
112	1848.21	1676.18	2020.24	82.95
113	1885.54	1708.81	2062.27	85.22
114	1922.86	1741.55	2104.17	87.43

表 三-2 個案公司電力用電預測結果—時間序列分析模式

(資料來源：台電公司)

就個案公司的長期負載預測作業而言，所需進行的預測項目很多，包括：電燈用電預測、電力用電預測、尖峰負載預測、平均負載預測、區域別負載預測，以及產業別負載預測等；所需輸入與研析之各式相關資料源龐雜，包括：國內外經濟情勢、經濟成長預測、產業結構預測、能源情勢分析、大型開發案以及人口成長、電價和氣溫等相關情資分析；所需構建的預測模型，則須依據預測項目、輸入資料性質，以及搭配合適預測方法，進行各式長期負載預測模型交替方案模

擬，始能定案。

然而，定案後的長期電力負載預測結果，僅是提供從過去預測未來所產生的必要資訊而已，並無法從預測結果看出各項外生變數之間預測未來之後彼此的交互作用關係，也無法得知外生變數和所要預測的標的之間的變動關聯。本論文即在於協助個案公司，如何從既有資訊進一步延展成知識，從已被整合的單一觀點的資訊，透過良好的分析架構之應用，進一步延展成可包括多元觀點的知識，來呈現對於未來大型社會經濟結構可能變化趨勢之發展脈絡。而進一步以資訊管理學的領域來看，本論文所做的研究，除了提供資訊進一步延展成知識之功用外，以未來構建成資訊應用系統的觀點來看，其構建的核心應用程式則是從以模型為導向進展成以知識為導向，則本論文所能提供的參考價值，不可謂不大。

經由上述討論得知，本論文所要進行的研究，就是奠基在個案公司既有長期電力負載預測作業模式的基礎上，進一步透過本論文研究的結果，應用第二章文獻探討分析所發展出來的 IDTBN 模式，整合領域專家對於各項影響長期電力負載預測可能的因素之專家意見，輔助闡述既有計量分析預測模式的結果，讓原本單一預測的結果，因此而有了豐富的內涵和未來發展趨勢的知識內容。以下繼續說明本論文之整體研究設計和實證研究方法的應用。

## 第二節 IDTBN 模式

經由第二章的文獻探討，可得知四種常見之資料分析方法的基本概念及優缺點，決策樹及貝氏網路適合用以作為資料分析並建立出易於解釋分析的架構，並可用機率分布的模型來表示變數間相互影響的關係與強度，在專家意見模型的建構上，更能輕易歸納出專家意見的脈絡走向與所強調關鍵變數關係。

類神經網路雖然學習精確度高、可接受不同類型的變數作為輸入且可建構非線性模型，或找出人類無法覺察出的關係，但缺點是執行速度緩慢，計算量大

相當耗費電腦資源，網路架構和網路動態的決定尚缺乏非常系統化的方法，並且類神經網路之訓練與測試的細節無法得知，相對於貝氏網路以機率方式建構，類神經網路中複雜的數學計算，較難解釋其涵義(Burnside 2005)。

案例式推理系統的優點是只需要加入與系統特徵與輸出結果相符的新案例，便形同將新的知識涵括進來，不需要如同其它資料分析模型，例如：貝氏網路或類神經網路，只要新資料加入仍需要經過重新計算，更新到新的架構上，案例式推理也有其缺點，在使用其結果來解釋時需格外注意，因為它與類神經網路類似的是，它仰賴的是其知識庫裡的所有案例，除非知識庫裡的某案例完全符合新進入的未知案例，否則無法呈現出決策制定的機率(Burnside 2005)。

又貝氏網路相對於決策樹結構而言，較難直接萃取出推論法則，若將二者結合則可呈現出複雜的決策制定過程(Janssens, Wets et al. 2006)。本研究乃提出 IDTBN 模式 (Integrated Decision Tree and Bayesian Network)，建構程序如下：

1. 以專家意見過錄後的資料，經過初步的資料前置處理，例如：必要時的資料筆數擴充與數值資料離散化，再以 C4.5 決策樹演算法建構出決策樹架構，並依其結果可分析出主要的關鍵屬性，亦即影響分類結果最主要的變數。
2. 於決策樹中被分類錯誤的實例資料，表示與整體資料的一致性較不吻合，因此是可視為是離群值，於決策樹分析後將此種實例資料除去。
3. 由於決策樹依據關鍵屬性而呈現樹狀分支結構，在此階段將分類結果依據決策樹分支的相關性分為幾組資料，以進行接下來的貝氏網路分析，以期貝氏網路的結果能透過先行決策樹推論分類而更加收斂。
4. 依上一步驟資料分類各別進行貝氏網路的建構，經由參數學習 (parameter learning) 的過程，將不足的資訊以推算的方式補足，預期可得到準確度更高且更符合各組貝氏網路連結關係所隱含的意義。

5. 所建構出之個別貝氏網路可用作目標變數之分類預測，並可分析變數之間的相互關係以及機率分布值，藉此能瞭解變數間的影響強度。



### 第三節 實證研究方法與設計

本論文旨在於發展 IDTBN 模式，整合領域專家對於各項影響長期電力負載預測可能的因素變化之專家意見，因此，資料的蒐集係以專家意見為標的，而為讓專家意見能夠在最自然的情境下發表，讓研究者能在最自然的情勢下蒐集到專家意見，所以，在意見蒐集上係搭配個案公司的經常性作業模式進行。換言之，本論文係採經驗研究的方式進行相關資料的蒐集、研究觀察以及資料分析等過程，基此，本論文即以實證研究的方法進行相關研究過程，而所謂的實證研究的方法即是以經驗研究為基礎(Chava et al., 2003)。在整個實證研究的過程，本論文所進行的研究步驟如下：

- (一) 研究方法：實證研究式參與及觀察。
- (二) 資料蒐集方式：蒐集個案公司次級資料、專家諮詢會議現場錄音、錄音資料轉錄分析與處理。
- (三) 資料蒐集期間：參與及觀察 1 年、專家諮詢會議現場錄音自 2007 年 12 月至 2008 年 3 月。
- (四) 專家諮詢會議參與專家背景：專家諮詢會議共蒐集 3 場次，約 100 名人次專家到場，專家背景涵蓋電力、電機、經濟、產業、環保等，任職單位涵蓋電力公司、政府能源部門、政府產業經濟部門、政府環保單位、民間智庫、環保團體、學界、產業界等。
- (五) 資料過錄處理：分成兩組進行交叉比對，先將專家諮詢會議現場錄音資料轉錄成文字檔，再將文字檔資料轉錄成內容綱要檔，然後根據內容綱要檔發展資料格式，最後再根據資料格式將內容綱要檔進一步轉錄成數值檔資料，經由資料檢核與交叉比對過程確認資料無誤，始建檔提供後續資料分析使用。

(六) 資料分析處理：配合本論文研究設計的指引，進行資料分析之處理。

而在研究設計方面，依據本論文所要發展的 IDTBN 模式，主要表現在資料分析的方式上，以使用來驗證 IDTBN 模式的最優化。換言之，經由上述實證研究的過程，蒐集到本論文所需資料，在研究設計上，則將實證資料檔案先進行決策樹方法分析（第一組）、然後再將實證資料檔案進行貝氏網路模式分析（第二組）、最後再進行 IDTBN 模式分析（第三組），也就是將資料檔案依據策樹方法分析的結果重新進行歸類，然後將重新歸類處理後的新的資料檔案，再分別進行貝氏網路模式分析，所得的結果綜合分析就完成 IDTBN 模式之分析。在理論上，第三組的分析結果應當會優化於第一組和第二組，因為資料經由策樹方法分析指引，重新歸類後再個別構建所合適的貝氏網路模式，照道理應該要優於第一組和第二組的分析結果。

如果，本論文能夠如實驗證所提出的 IDTBN 模式具有整合專家預測意見的效能，那麼，經由本論文所發展的實證研究過程，未來可以考慮網頁化，也就是說搭配原先既有計量預測分析模式，架構成專家討論社群，讓專家根據既有計量預測模式分析結果，進行社群互動和討論，然後系統再發展 web mining 應用模組，將原先人工資料轉錄的作業經由 web mining 應用模組探勘分析的結果，自動轉錄成數值性資料檔案，然後再搭配本論文所提出的 IDTBN 模式分析結果，結合既有計量預測模式分析的結果，就可以進一步如同之前所述，協助個案公司，將長期電力負載預測作業，從既有資訊進一步延展成知識，從已被整合的單一觀點的資訊，透過良好的分析架構之應用，進一步延展成可包括多元觀點的知識，藉以呈現對於未來大型社會經濟結構可能變化趨勢之發展脈絡。而從資訊應用系統的觀點來看，洽足以支援核心應用程式從以模型為導向進展成以知識為導向。這是本論文展望未來希望能夠進一步發展的可能研究方向。

## 第四節 研究架構

從研究概念來看，本論文聚焦的重點在於發展 IDTBN 模式，所以，本論文整體的研究架構，實際上是建立在長期電力負載預測既有計量分析作業模式的基礎上，再透過個案公司的實務問題，以及相關文獻的指引，發展出 IDTBN 模式，然後，經由本論文研究設計，驗證出 IDTBN 分析模式的優化所在，最後進一步將既有計量預測模式和 IDTBN 分析模式結合在一起，提供個案公司將長期電力負載預測作業，從既有資訊進一步延展成包括多元觀點的知識，也可支援個案公司未來進一步發展相關資訊系統時，參考使用。本論文整體研究架構如圖 三-1 所示。

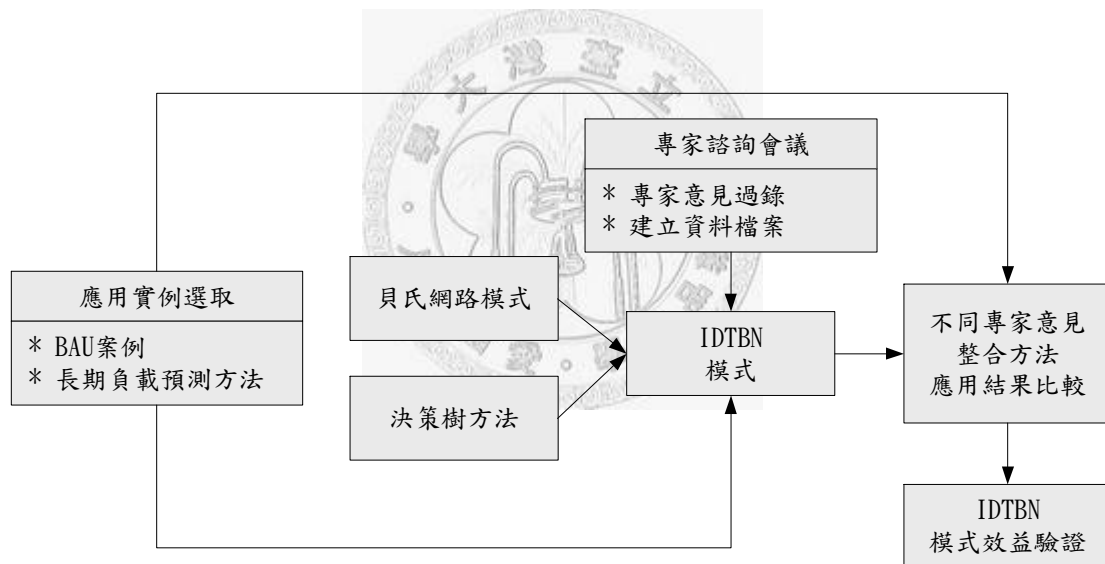


圖 三-1 本論文研究架構

(資料來源：本研究整理)

## 第五節 專家意見內容過錄

本節進一步詳加說明本論文有關專家意見內容過錄處理的情形，俾能掌握資料產生的過程、資料特性以及資料最終完成的格式，以進行下一階段之資料分析。



(一) 參與專家諮詢會議：

共 3 場。第一場會議時間 2007 年 12 月 27 日由個案公司主辦，召集產官學界相關電力、電機、能源、經濟、產業、環保等專家，針對主辦單位所提 2008 年至 2027 年之未來 20 年長期電力負載預測結果，提出廣泛性意見諮詢，這一場次與會人士約有 30 人。第二場會議時間 2008 年元月 3 日由個案公司主辦，召集個案公司內部相關部門，包括：電力調度、電力系統規劃、發電單位、企劃、業務、電源開發、會計、財務、環保、研究單位等，由主辦部門負責提報所提 2008 年至 2027 年之未來 20 年長期電力負載預測結果，請個案公司內部專家提出修正意見，與會人數約有 30 人。第三場會議時間 2008 年元月 22 日由個案公司主辦，召集個案公司內部相關部門以及公司外專家學者，由主辦部門負責提報所提 2008 年至 2027 年之未來 20 年長期電力負載預測結果，請個案公司內部專家和外部專家提出修正意見，與會人數約有 40 人。

(二) 會議記錄、現場錄音：除了參與觀察之外，事後並作成會議記錄和現場錄音，提供後續資料轉錄使用。

(三) 錄音檔逐字稿：第一輪先將現場錄音作成文字逐字稿。

(四) 過錄逐字稿、濃縮要義、交叉比對：此部分資料處理結果，詳附錄一：專家諮詢會議資料過錄表所示。

(五) 根據步驟四濃縮要義結果，萃取重要變數和內涵，並進而發展資料過錄格式表，提供第二輪資料過錄參考使用。此部分資料處理結果，詳附錄二：資料過錄格式表所示。

(六) 修定資料過錄表格、完成資料檔案建置，以及資料交叉檢核：此部分資料處理結果，詳附錄三：資料檔案所示。至此完成的資料檔案才送交下一研究程序進行資料分析使用。

## 第四章、研究過程與結果

本章節將分為資料前置處理、決策樹方法結果、貝氏網路方法結果、IDTBN 之方法應用結果及應用效益分析五大部分。首先第一節將說明專家意見過錄資料之前置處理過程；第二節是進行單一決策樹模型之建構；第三節是進行單一貝氏網路模型之建構；第四節為本研究所提出之 IDTBN 之方法應用，乃結合 Bayesian Network 與 Decision Tree 方法，對於專家意見的預測模型建立，能更加精確地過濾出可信度較高的資料，並期望以此專家意見模型能提升 BAU case 的完整度；第五節將以 IDTBN 的實驗結果，進行應用效益分析。

### 第一節資料前置處理

本研究於此階段將進行 IDTBN 的實驗，首先從包含電力、電機、能源、經濟、產業、環保等領域的專家意見中，將較為無關的資訊去除，經由資料過錄整理，挑選出專家不斷重覆提及之可能影響電力負載量的 17 個重要的外生變數，如表 四-1 所示。

變數名稱	變數意義
X1	國際經濟成長率
X2	CO <sub>2</sub> 減量效果
X3	環保效果
X4.1	國內平均經濟成長率
X4.2	GDP 所得彈性
X5.1	工業部門佔比預估
X5.2	服務業部門佔比預估
X6	重大投資案

X7	DSM 抑低量
X8	預估人口成長率
X9	氣溫效果，增加負載量預估
X10	預估電價調升
X11	能源技術提升抑低用電量效果
X12	物價上漲抑低電力負載效果
X13	預估用電量年平均成長率
X14	預估尖峰負載年平均成長率
X17	節能政策目標抑低電力負載效果

表 四-1 資料分析所選取之變數名稱與代表意義

### 資料轉換與增加資料變異性

專家的意見資料由於是以口頭陳述方式表達，專家背景不同，對於電力負載提及的意見著重點也不盡相同，因此必須先將專家意見轉換為以上述外生變數表示之分析性或分類性數值，此為資料過錄過程，以取得配合實驗所需之格式的資料。各變數的資料過錄時所參照的對照表如表 四-2 所示。

變數名稱	變數屬性	說明
X1	分析性	空白：代表未提及，為缺值。 餘者：填入實際數值。
X2	分析性	空白：代表未提及，為缺值。 強度：1-9。
X3	分類性	0：未提及。 1：認為環保具抑低電力負載量效果。
X4.1	分析性	空白：代表未提及，為缺值。 餘者：填入實際數值。
X4.2	分析性	空白：代表未提及，為缺值。 餘者：填入實際數值。
X5.1	分析性	空白：代表未提及，為缺值。 餘者：填入實際數值。

X5.2	分析性	空白：代表未提及，為缺值。 餘者：填入實際數值。
X6	分類性	0：代表無影響。 1：投資案增加，電力負載量增加。 2：產業外移，抑低電力負載量。
X7	分類性	0：未提及。 1：具抑低電力負載效果。
X8	分析性	空白：代表未提及，為缺值。 餘者：填入實際數值。
X9	分析性	空白：代表未提及，為缺值。 餘者：填入實際數值。
X10	分析性	空白：代表未提及，為缺值。 餘者：填入實際數值。
X11	分析性	空白：代表未提及，為缺值。 餘者：填入實際數值。
X12	分類性	0：未提及。 1：具抑低電力負載效果。
X13	分析性	空白：代表未提及，為缺值。 餘者：填入實際數值。
X14	分析性	空白：代表未提及，為缺值。 餘者：填入實際數值。
X17	分類性	0：未提及。 1：認為環保具抑低電力負載量效果。

表 四-2 專家意見過錄對照表

由於專家估計值過於保守，於是導致代表專家意見之數值資料中，各變數對應變數 X13 的影響不顯著，可能造成分類效果不彰。因此在此階段依照各變數對 X13 促使增加或抑低之正負面影響，對 X13 之數值進行加權方式的微調，目的在增加 X13 數值間之離散程度。其中，X11 因其數值太大且變化小，X2、X6、X9、X10 皆包含缺值，故不予以考量。X13 經調整之後為 X13'，上述變數之正負影響關係與加權的計算式如下：

$$X13' = X13 + (X1 - X3 + X4.1 + X4.2 + X5.1 - X5.2 - X7 + X8 - X12 - X17) / 50 - \text{Rand}() / 10$$

因為 Decision Tree 與貝氏網路分類法，應變數 X13' 需為分類性變數，於是

須將其作離散化分為四個類別，分別是>3.67、3.46-3.67、3.25-3.46 與<3.25，再接著進行實驗。

Weka (Waikato Environment for Knowledge Analysis) 是由紐西蘭的 Waikato 大學機器學習研究團隊所開發出來的軟體，是一套使電腦學習如何分析資料的演算工具，以 Java 撰寫可適用於多種平台，提供直接分析資料，或併入自行研發的 Java 應用軟體裡，Weka 包含了一些機器學習方面有用的工具集，例如 data pre-processing、classification、regression、clustering、association rules，以及 visualization 等等。以下的各研究流程將透過 WEKA 這套軟體所包含的決策樹及貝氏網路演算法進行實驗。

## 第二節 效益衡量指標

為驗證本研究效益，乃採以 Precision、Recall、Accuracy 與 F-measure 等作為衡量指標。表 四-3 為 confusion matrix 示意圖，用以表示實際與分類結果的所對應的筆數分布，是由 true positive、true negative、false positive 和 false negative 所組成的矩陣。相關之衡量指標分述如下：

實際 分類	E1	E2
E1	<b>tp</b> (true positive)	<b>fp</b> (false positive)
E2	<b>fn</b> (false negative)	<b>tn</b> (true negative)

表 四-3 confusion matrix 示意圖

- True positive rate : TP-rate =  $\frac{tp}{tp + fn}$ ，是指目標樣本分類正確的樣本數目比率，與 Recall 值同義。
- False positive rate : FP-rate =  $\frac{fp}{fp + tn}$ ，是指非目標樣本分類成目標樣本

的錯誤樣本數目比率。

➤ Precision (正確率):  $\text{Precision} = \frac{tp}{tp + fp}$ ，代表分類結果為 E1 的資料

中，實際為 E1 的資料筆數比率。

➤ Recall (召回率):  $\text{Recall} = \frac{tp}{tp + fn}$ ，代表實際為 E1 的資料中，被分

類為 E1 的資料筆數比率。

➤ Accuracy (準確率):  $\text{Accuracy} = \frac{tp}{tp + fp + fn + tn}$

➤ F-measure  $= \frac{2}{\frac{1}{\text{Recall}} + \frac{1}{\text{Precision}}} = \frac{2 \times \text{Recall} \times \text{Precision}}{\text{Recall} + \text{Precision}}$

在一般的情形之下 Precision 跟 Recall 存在反比關係，即提升 Precision 時，Recall 就會下降，反之亦然。因此為了結合這兩樣測量數值，就有了新的測量值 F-measure，是兩者的調和平均數。

### 第三節 決策樹方法結果

採用 WEKA 提供之 J48 演算法來建立決策樹的模型，主要先將各分析性變數離散化之後成為分類性變數，並找出重要之特徵變數。以 X13' 亦即「預估用電量年平均成長率」為應變數，變數 X1~X12 與 X17 為預測變數，此外為求得更精確的分類正確率，採以 Cross-validation (多次交叉確認模式) Folds = 10 來驗證以該資料建立決策樹之準確度，如圖 四-1。參數列設定為 weka.classifiers.trees.J48 -C 0.25 -M 2，參數意義及其設定如表 四-4 所示。

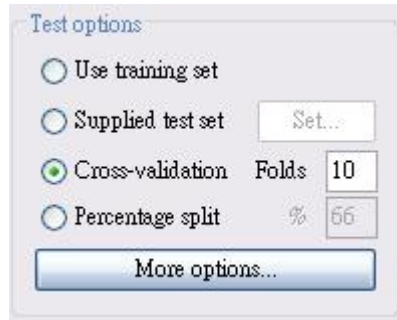


圖 四-1 WEKA 交叉確認模式設定

參數名稱	參數值
confidence Factor	0.25
進行決策樹修剪的參數，數值愈小表修剪程度愈高	
minNumObj	2
每個葉節點至少含有之 instance 數目	
numFolds	3
決定用以修剪決策樹的 fold 數，其餘則是用以建立決策樹	
subtreeRaising	True
在修剪決策樹時是否考慮 subtree raising operation	

表 四-4 WEKA J48 決策樹之參數設定

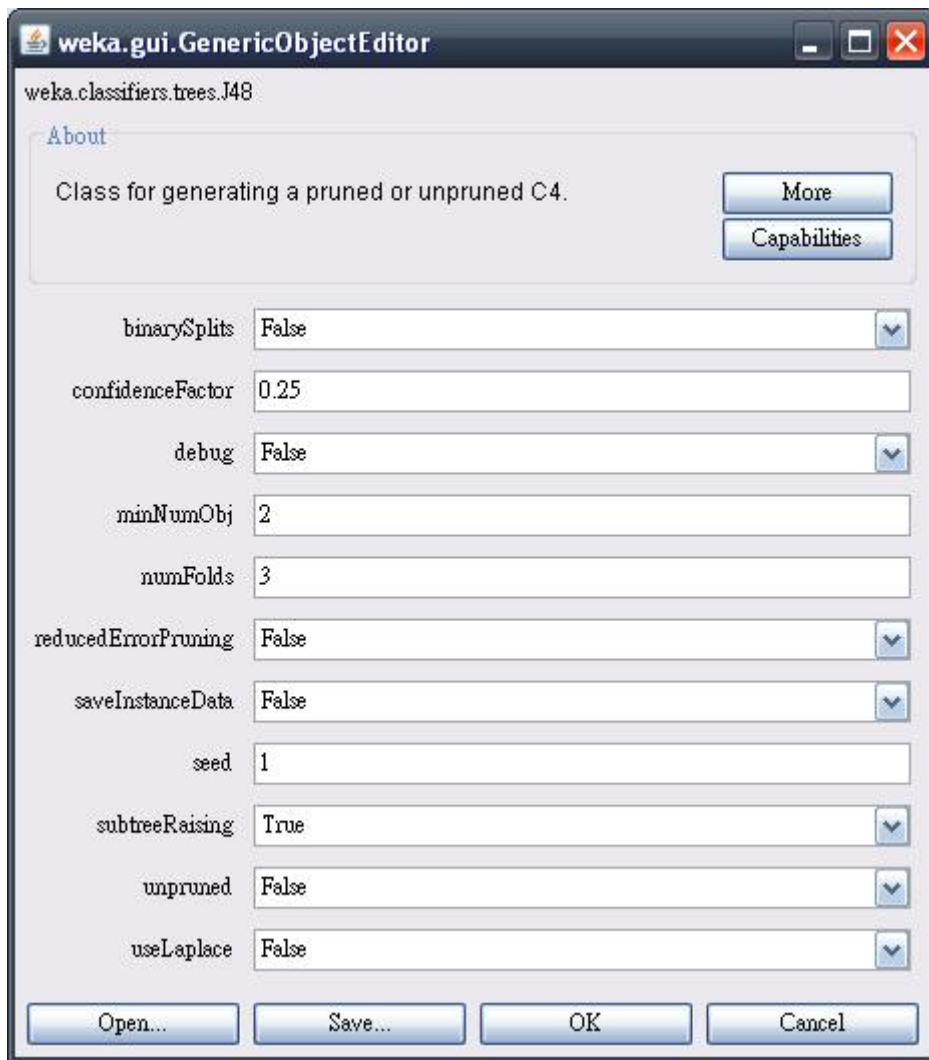


圖 四-2 WEKA J48 參數設定介面

Correctly Classified Instances 共 135 筆，分類準確率為 91.2162 %；Incorrectly Classified Instances 共 13 筆，分類錯誤率為 8.7838 %。使用 WEKA J48 演算法所建立之決策樹模型，如圖 四-3 所示。而分類結果衡量指標與 confusion matrix 詳見表 四-5 與表 四-6。



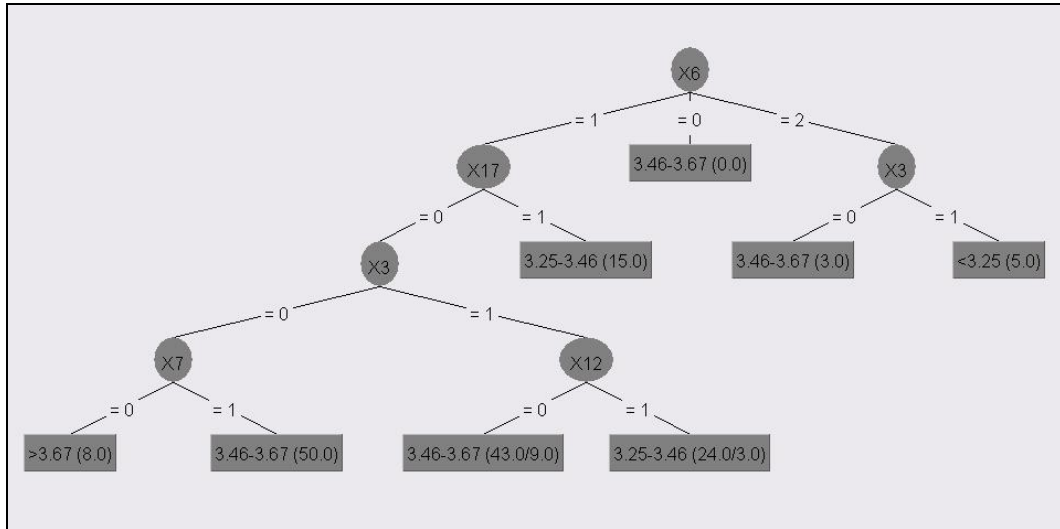


圖 四-3 WEKA J48 決策樹模型

TP Rate	FP Rate	Precision	Recall	F-Measure	Class
1	0.007	0.833	1	0.909	a = <3.25
0.9	0.028	0.923	0.9	0.911	b = 3.25-3.46
0.956	0.155	0.905	0.956	0.93	c = 3.46-3.67
0.615	0	1	0.615	0.762	d = >3.67

表 四-5 WEKA J48 分類結果衡量指標

	a	b	c	d
a	5	0	1	0
b	0	36	3	0
c	0	4	86	5
d	0	0	0	8

表 四-6 WEKA J48 之 confusion matrix

## 第四節貝氏網路方法結果

採用 WEKA 提供之 BayesNet 演算法以建立貝氏網路的模型。同樣以  $X_{13}$ ' 亦即「預估用電量年平均成長率」為應變數，變數  $X_1 \sim X_{12}$  與  $X_{17}$  為預測變數。參數列設定為 BayesNet -D -Q weka.classifiers.bayes.net.search.local.K2 -- -P 3 -S BAYES -E weka.classifiers.bayes.net.estimate.SimpleEstimator -- -A 0.5，參數意義與設定如表 四-7 所示。

參數名稱	參數值
estimator: SimpleEstimator	Alpha: 0.5
以 data 評價/評估所建立之貝氏網路結構的條件機率表 (conditional probability tables)。Alpha 值可視為是每個機率值之初始值。	
initAsNaiveBayes	True
用以進行 structure learning 的初始網路結構是 Naive Bayes Network。	
maxNrOfParents	3
設定貝氏網路中節點最多可擁有多少個父節點。當初始為 Naive Bayes 且 maxNrOfParents 設定為 1 時，會生成 Naive Bayes classifier。當 maxNrOfParents 設為 2，將建立 Tree Augmented Bayes Network (TAN)；當 maxNrOfParents 設定大於 2，則建立 Bayes Net Augmented	

Bayes Network (BAN)。	
scoreType	BAYES
決定要使用何種方式衡量網路結構的品質。有 Bayes、BDeu、Minimum Description Length (MDL)、Akaike Information Criterion (AIC)和 Entropy 可供選擇。	

表 四-7 WEKA BayesNet 貝氏網路之參數設定

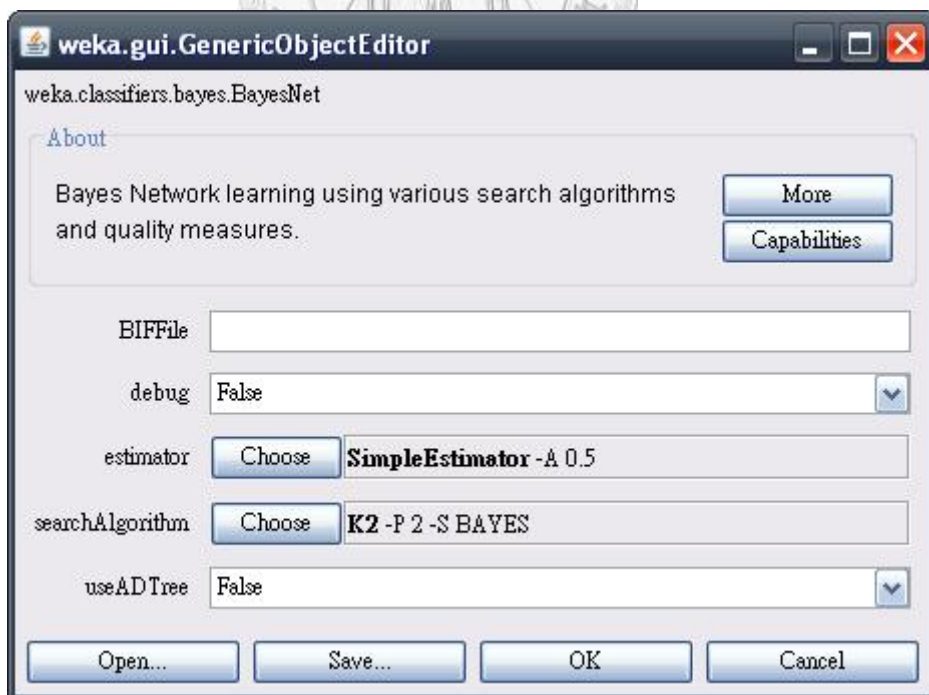


圖 四-4 WEKA BayesNet 參數設定介面

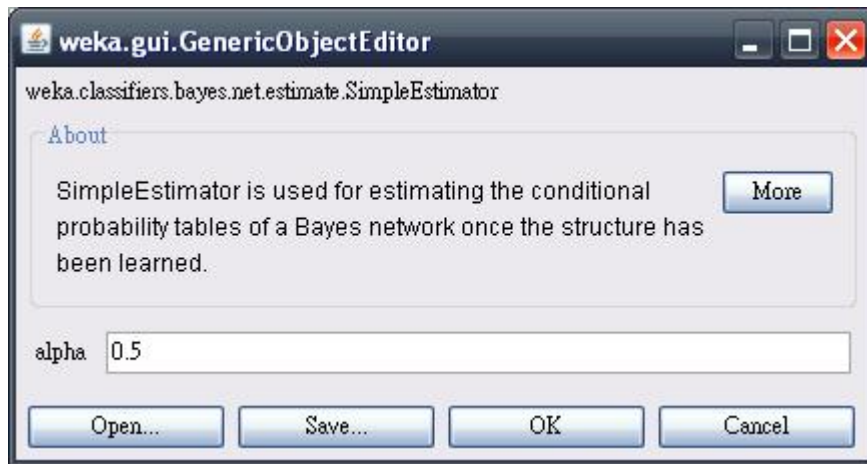


圖 四-5 WEKA BayesNet Simple Estimator 參數設定介面

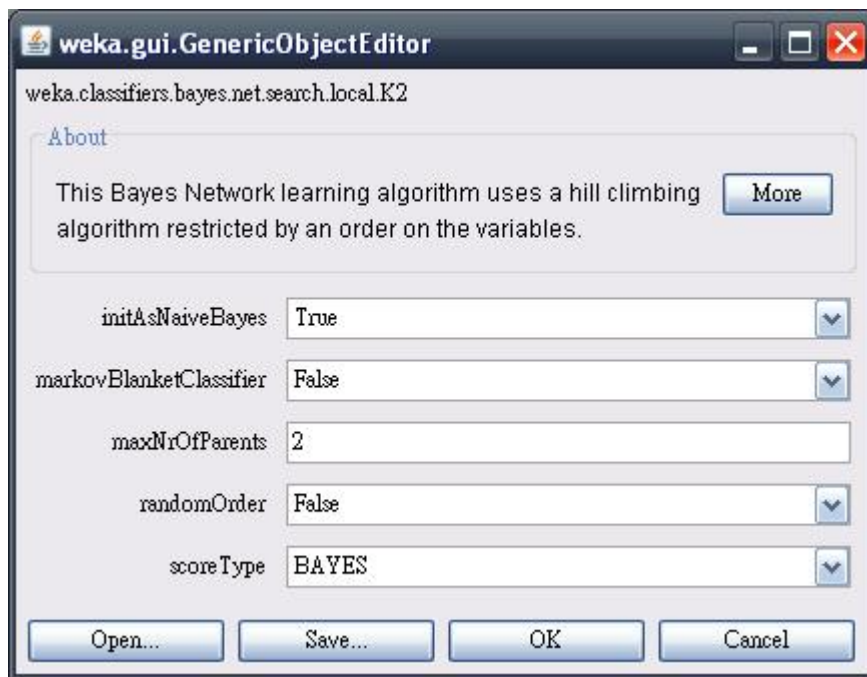


圖 四-6 WEKA BayesNet search algorithm : K2 參數設定介面

Correctly Classified Instances 共 136 筆，分類準確率為 91.8919 %; Incorrectly Classified Instances 共 12 筆，分類錯誤率為 8.1081 %。使用 WEKA BayesNet 演算法所建立之貝氏網路模型，如圖 四-7 所示。而分類結果衡量指標與 confusion matrix 詳見表 四-8 與表 四-9。

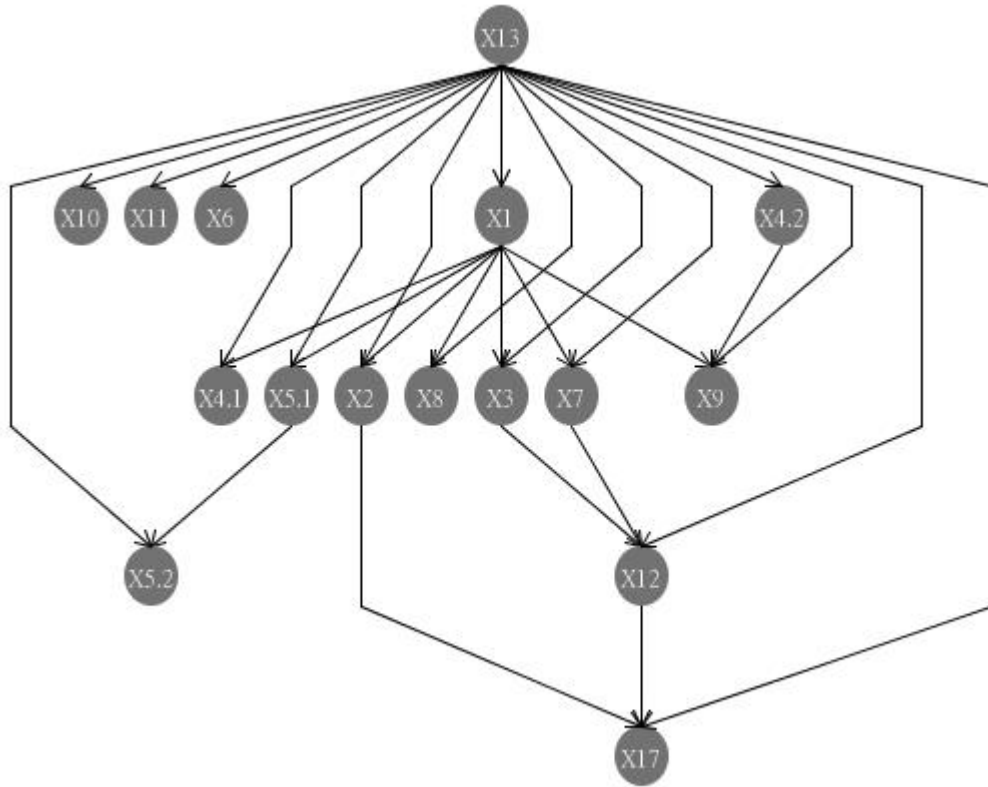


圖 四-7 WEKA BayesNet 貝氏網路模型

TP Rate	FP Rate	Precision	Recall	F-Measure	Class
1	0	1	1	1	<3.25
0.85	0.028	0.919	0.85	0.883	3.25-3.46
0.956	0.19	0.887	0.956	0.92	3.46-3.67
0.615	0.007	0.889	0.615	0.727	>3.67

表 四-8 WEKA BayesNet 分類結果衡量指標

	a	b	c	d
a	5	0	0	0
b	0	34	3	0
c	0	6	86	5
d	0	0	1	8

表 四-9 WEKA BayesNet 之 confusion matrix

但因為以所有資料直接進行貝氏網路建立，其結構過於複雜，其間之關係難以解釋，於是將依決策樹結果分類之兩大類資料如圖 四-8，於第四節將進行 IDTBN 方法，參照決策樹的分類結果，分別建立貝氏網路，期能使貝氏網路的結果更加收斂具解釋意義。

## 第五節 IDTBN 方法應用結果

由第二節決策樹之分類結果，依據它的分支走向與關連性分為兩大類，並考慮減少各類筆數差異之懸殊，可分為兩大類資料，分別以虛實線方框代表（虛線方框表示第一類，實線方框則表示第二類），如圖 四-8 所示。希望利用此分類改善第三節僅使用單一貝氏網路的分類效果，使原本複雜難以解釋的結構，能經由決策樹的事先分類後，轉變為較收斂之貝氏網路結構。

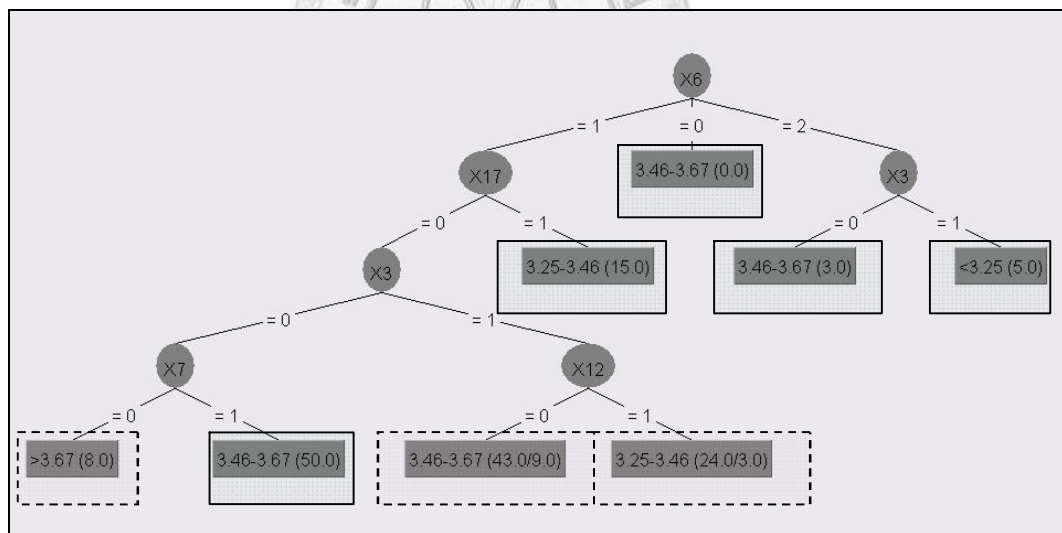


圖 四-8 參考決策樹模型之兩大類資料

第一類總資料筆數為 75 筆，分類結果：Correctly Classified Instances 共 63 筆，分類準確率為 84%；Incorrectly Classified Instances 共 12 筆，分類錯誤率為 16%。使用 WEKA BayesNet 演算法所建立之貝氏網路模型，如圖 四-9 所示。而分類結果衡量指標與 confusion matrix 詳見表 四-10 與表 四-11。

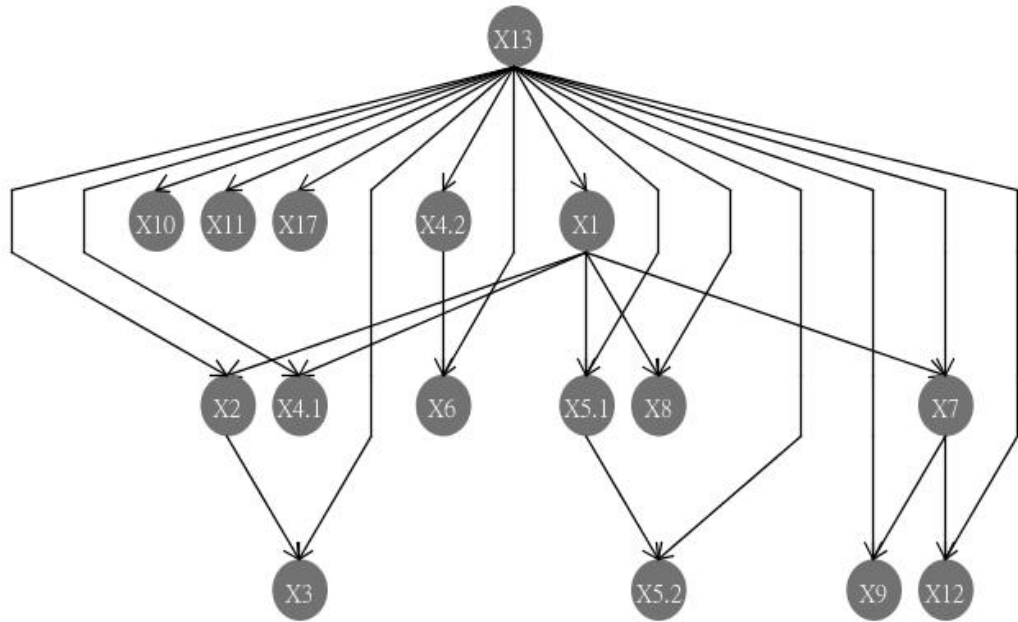


圖 四-9 WEKA BayesNet 貝氏網路模型：分類一

TP Rate	FP Rate	Precision	Recall	F-Measure	Class
0	0	0	0	0	<3.25
0.84	0.06	0.875	0.84	0.857	3.25-3.46
0.919	0.237	0.791	0.919	0.85	3.46-3.67
0.615	0	1	0.615	0.762	>3.67

表 四-10 WEKA BayesNet 分類結果衡量指標：分類一

	a	b	c	d
a	0	0	0	0
b	0	21	3	0
c	0	4	34	5
d	0	0	0	8

表 四-11 WEKA BayesNet 之 confusion matrix：分類一

由於正確率過低，將第一類資料剔除於決策樹模型中被分類錯誤的資料後，第一類總資料筆數為 64 筆，分類結果：Correctly Classified Instances 共 63

筆，分類準確率為 98.4375 %；Incorrectly Classified Instances 共 1 筆，分類錯誤率為 1.5625 %；可知此作法大幅提升準確率。使用 WEKA BayesNet 演算法所建立之貝氏網路模型，如圖 四-10 所示。而分類結果衡量指標與 confusion matrix 詳見表 四-12 與表 四-13。

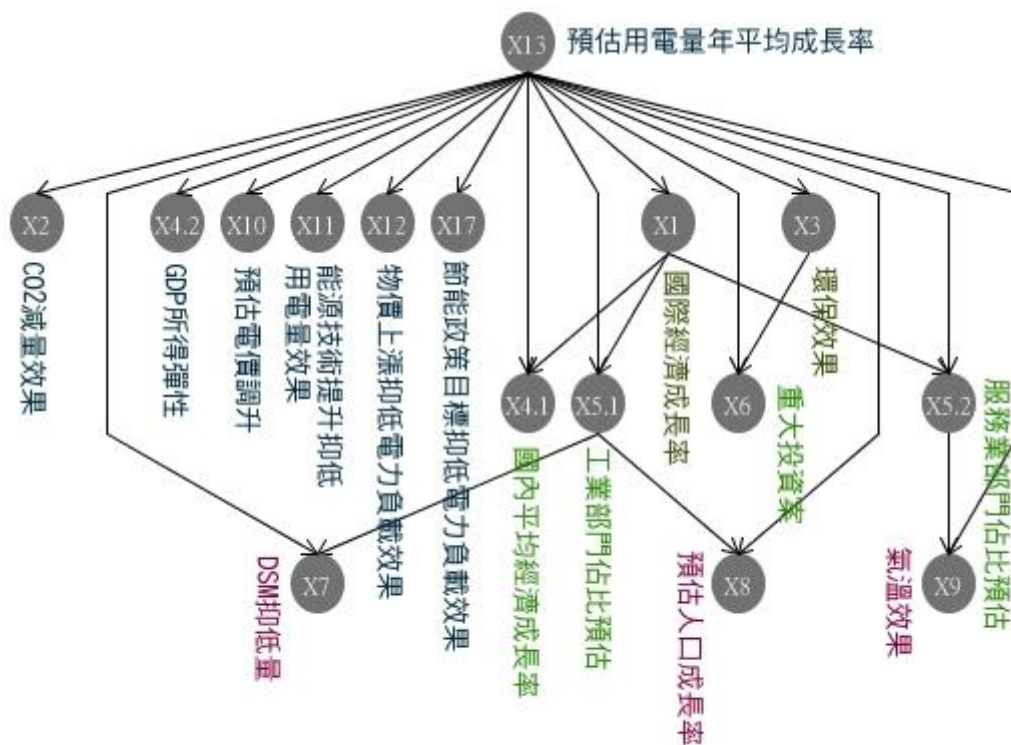


圖 四-10 WEKA BayesNet 貝氏網路模型：去除錯誤資料之分類一

TP Rate	FP Rate	Precision	Recall	F-Measure	Class
0	0	0	0	0	<3.25
1	0.023	0.955	1	0.977	3.25-3.46
0.971	0	1	0.971	0.986	3.46-3.67
1	0	1	1	1	>3.67

表 四-12 WEKA BayesNet 分類結果衡量指標：去除錯誤資料之分類一



	a	b	c	d
a	0	0	0	0
b	0	21	1	0
c	0	0	34	0
d	0	0	0	8

表 四-13 WEKA BayesNet 之 confusion matrix：去除錯誤資料之分類一

第二類總資料筆數為 75 筆，分類結果：Correctly Classified Instances 共 73 筆，分類準確率為 100 %。使用 WEKA BayesNet 演算法所建立之貝氏網路模型，如圖 四-11 所示。而分類結果衡量指標與 confusion matrix 詳見表 四-14 與表 四-15。

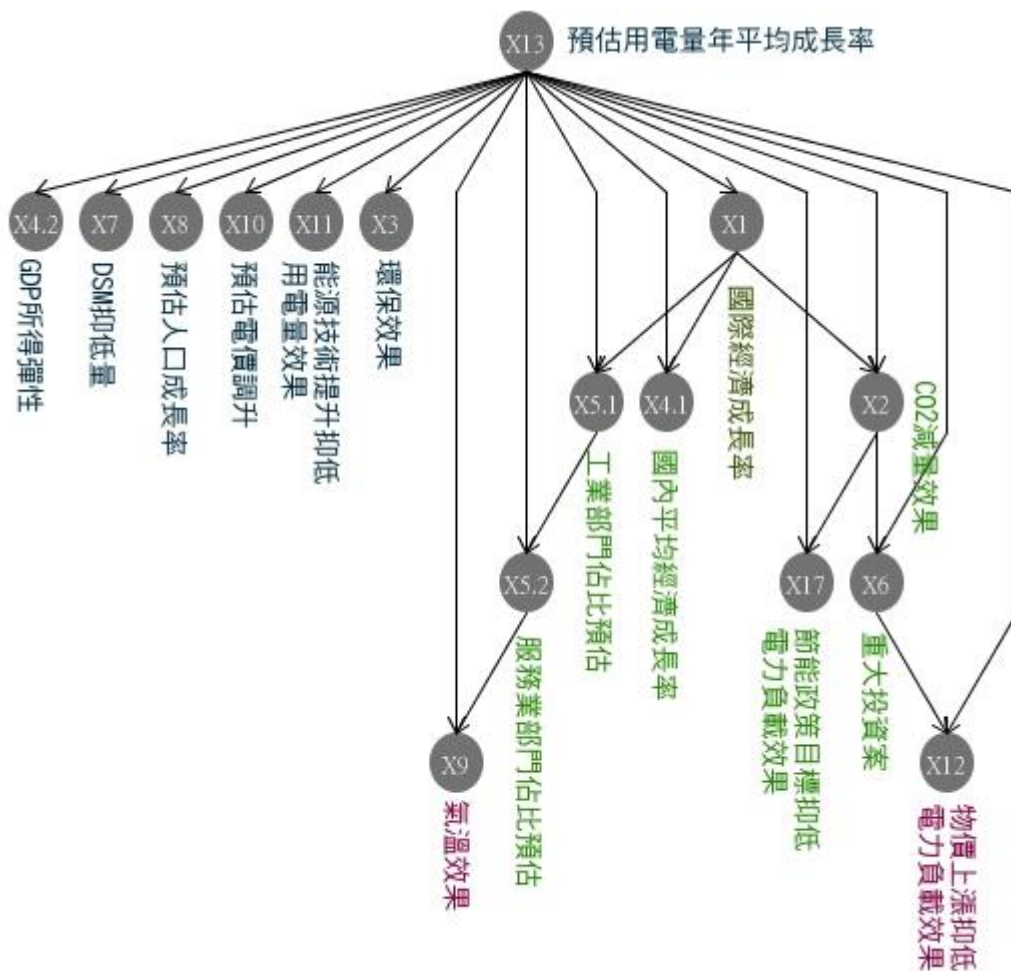


圖 四-11 WEKA BayesNet 貝氏網路模型：分類二

TP Rate	FP Rate	Precision	Recall	F-Measure	Class
1	0	1	1	1	<3.25
1	0	1	1	1	3.25-3.46
1	0	1	1	1	3.46-3.67
0	0	0	0	0	>3.67

表 四-14 WEKA BayesNet 分類結果衡量指標：分類二

	a	b	c	d
a	5	0	0	0
b	0	15	0	0
c	0	0	53	0
d	0	0	0	0

表 四-15 WEKA BayesNet 之 confusion matrix：分類二

由實驗結果可知，本研究所提出之 IDTBN 方法，能經由先建立決策樹分類，剔除錯誤資料，以提升貝氏網路實作的正確率；並經由決策樹模型找出重要關鍵變數，使貝氏網路建立的結果能更加收斂，表示之意義也更加明確合理。

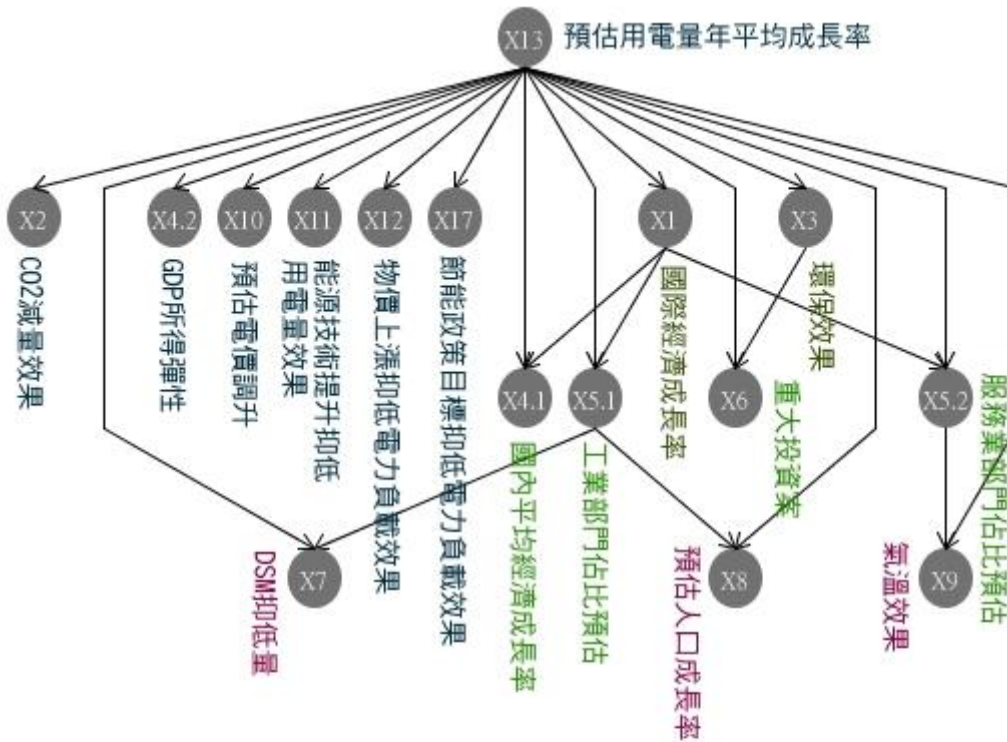
## 第六節 IDTBN 方法效益分析

本論文在資料分析過程，首先根據決策樹之樹狀分類結果，將資料分成組內同質性較高而組間差異性大的兩組，其在實質領域上的意涵實則包含一組具提高 X13 用電量年平均成長率效果之外生變數，該組以重大投資案為初始分隔節點，包括：預估投資案增加、電力負載量增加；未提及節能政策；未提及需求面管理之抑低用電量效果，從實質面來看，強調的是看好未來經濟成長和產業發展，而預估未來年平均用電量將會成長者，可以稱之為經濟成長組。而另一組之組成則是認為未來台灣會產業外移、主張環保、發展節能政策、發展電力需求面管理，從而預估未來年平均用電量之成長將會趨緩者，可以稱之為節能環保組。

接著，則將這兩組分別進行貝氏網路分析，企圖整合這兩組專家各自對於未來相關外生變數對於年平均用電量之可能產生的影響模式。經由本論文的研究過程與結果顯示，經濟成長組的貝氏網路模型從圖 四-10 可見，CO<sub>2</sub> 減量、GDP、電價、電力技術發展、物價水準以及節能政策發展等因素，均對未來年平均用電量之成長扮演直接的影響效果，而其餘因素則組成兩組環狀網絡產生對於 X13 用電量年平均成長率之影響效果，包括：一組以國際經濟成長情勢為首，影響國內經濟成長情勢、工業部門成長、服務業部門成長等而對 X13 用電量產生影響；另一組則以環保為首，藉由影響重大投資案發展，進而帶動對於 X13 用電量之影響。然而，從圖 四-10 亦得知，第三階所掛的影響因素，如：電力需求面管理、人口和氣溫等，由於串接在工業部門與服務部門成長之後，在實質應用面上難以詮釋，這可能是構建貝氏網路模型之限制。

而節能環保組所構建出來的貝氏網路模型詳見圖 四-11，其中可見 GDP、電力需求面管理、人口、電價、電力技術發展以及環保等因素，均對未來年平均用電量之成長扮演直接的影響效果，其餘因素則以國際經濟之發展情勢為重要節點，組成環狀網絡產生對於 X13 用電量年平均成長率之影響效果，包括：透過影響工業部門成長、服務部門成長、國內經濟成長情勢、CO<sub>2</sub> 減量、節能政策發展和重大投資案發展等，進而帶動對於 X13 用電量之影響。同樣地，節能環保組之貝氏網路模型其中第三階所掛的影響因素，如：氣溫和物價，亦如經濟成長組的貝氏網路模型一般在實質應用面上難以詮釋，再一次驗證了貝氏網路模型之限制。

不過，儘管本研究在構建貝氏網路模型所遭遇之限制，研究上仍可以從貝氏網路模型上各節點所提供的機率值分佈，進一步精確地瞭解各外生變數對於 X13 用電量之影響關係，例如：如圖 四-12 所示。未來更可透過所構建的經濟成長組的貝氏網路模型（詳見圖 四-12）和節能環保組所構建出來的貝氏網路模型（詳見圖 四-13），用來預測不同組別的專家其所提的對於未來各種外生變數之預估意見變化情形，對於 x13 用電量所產生的影響和預估 x13 用電量之變化數值。



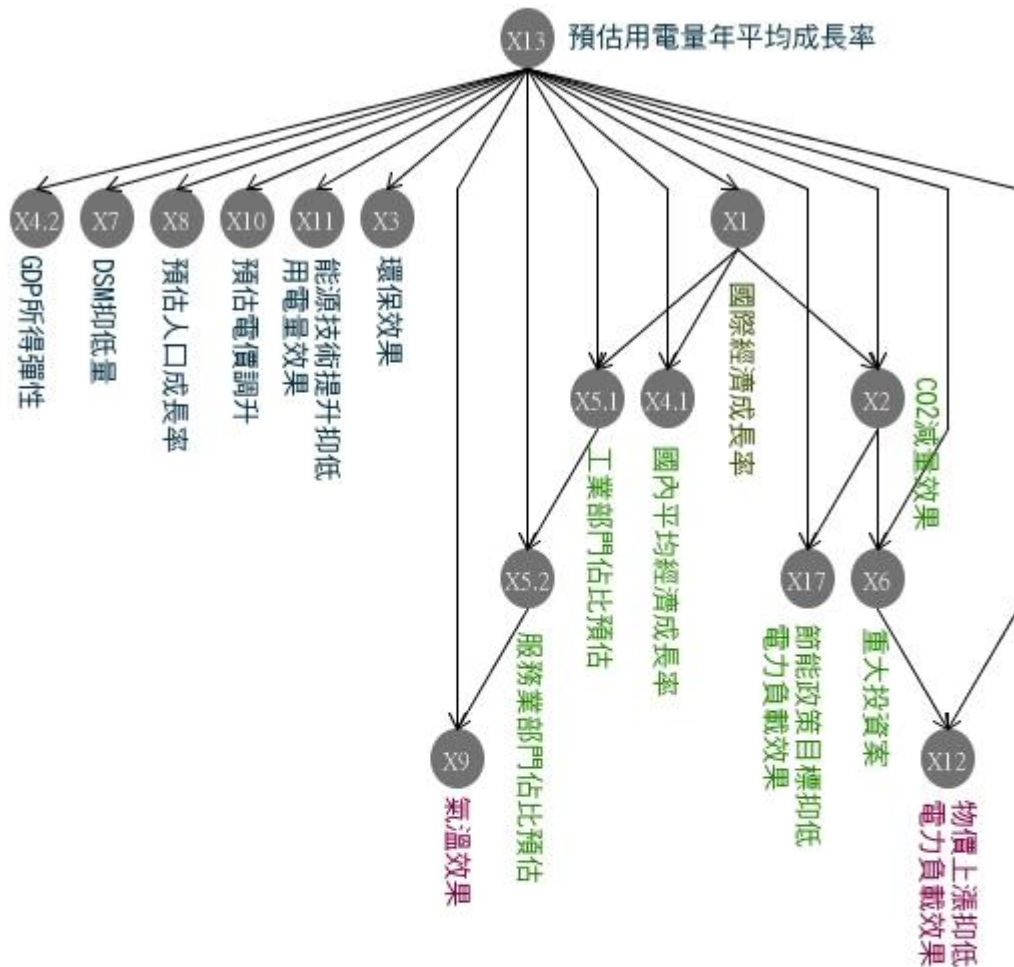
Probability Distribution Table For X13

<3.25	3.25-3.46	3.46-3.67	>3.67
0.008	0.326	0.538	0.129

Probability Distribution Table For X4.2

X13	0.65	0.84
<3.25	0.5	0.5
3.25-3.46	0.023	0.977
3.46-3.67	0.014	0.986
>3.67	0.056	0.944

圖 四-12 WEKA BayesNet 貝氏網路模型：去除錯誤資料之分類一 與其部分機率分布表



Probability Distribution Table For X13

<3.25	3.25-3.46	3.46-3.67	>3.67
0.073	0.207	0.713	0.007

Probability Distribution Table For X7

X13	0	1
<3.25	0.083	0.917
3.25-3.46	0.969	0.031
3.46-3.67	0.009	0.991
>3.67	0.5	0.5

圖 四-13 WEKA BayesNet 貝氏網路模型：分類二 與其部分機率分布表

歸納上述分析來看，本研究應用 IDTBN 模型用來整合分析專家意見的結果，其獲得的實質效益包括：

1. 有效提升模型的準確度，各階段所構建的模型準確度及其衡量指標 F-measure 結果如圖 四-14 和表 四-16 所示。圖 四-14 中本研究所提出

之 IDTBN 方法能使準確度提升至 98.4375 % (BN 分類一\*，表去除錯誤資料之分類一)與 100%(BN 分類二);表 四-16 中，本研究之 IDTBN 方法亦能使 F-measure 值達到 0.977 以上，因此可知 IDTBN 可用以處理專家意見整合問題。

2. 精確的資料分組與建模結果，讓實質應用層面的詮釋更容易進行。
3. 精確的資料分組與建模結果，讓實質應用層面的預估作業更能有效率地進行。

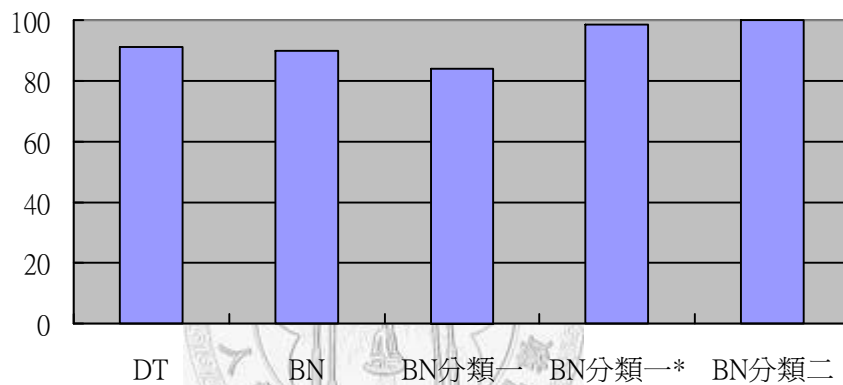


圖 四-14 各階段方法之 Accuracy 比較圖

	DT	BN	BN 分類一	BN 分類一 (去除錯誤資料)	BN 分類二
a	0.909	1	0	0	1
b	0.911	0.883	0.857	0.977	1
c	0.93	0.92	0.85	0.986	1
d	0.762	0.727	0.762	1	0

表 四-16 各階段方法之 F-measure 指標比較表

## 第五章、結論與建議

本章將對此研究之分析流程與實驗結果進行彙整，指出其中之研究價值與重要研究成果，並進一步探討本研究之研究限制及未來可能之研究方向，以供後續相關研究之參考。

### 第一節 結論

企業在進行生產研發等活動皆須仰賴大量資訊，如何將資訊轉換成為可輔助企業決策的知識，乃是資訊爆炸時代之下所面對的首要問題之一。在獲取知識的途徑中，通常針對某一特定專業領域，萃取並整合專家意見的作法是為常見的模式，因此若能有效收斂專家意見並將其以結構化模型呈現，提供日後分析與應用，進一步能支援各相關領域之決策制定。

目前資料探勘技術廣泛應用於資料模型之建立，或建構輔助決策的知識模型，例如：貝氏網路以條件機率為理論基礎，在給定各屬性之發生機率後，根據各屬性間相互關係計算出條件機率關係，建立貝氏網路，用以預測事件發生之機率；決策樹是以樹狀圖為基礎，主要用於分類與問題預測，與類神經網路不同的是，它是人類可理解的推論模式，可用文字來表達易於瞭解的規則，並能找出各決策分支點的關鍵變數，提供決策者作參考；類神經網路和決策樹類似的是，同樣應用於分類與預測，具強大學習能力及容錯能力，卻由於其較複雜的數學模式結構，與決策樹相較之下，推論過程較難以理解；案例式推理透過改編那些既有案例以嘗試解決新問題，是一種利用尋找相似案例的推理方法，不過須建立龐大的案例知識庫才能使預測更為準確。由於資料探勘的方法各有其優缺點，因此僅使用任何單一方法可能無法充分發展出專家意見整合模型，使專家意見收斂並表達出整合後之意見脈絡。

為能使專家意見有效整合，找出關鍵的目標值影響變數，並過濾可能不具代表性的意見，使所建立之結構簡化收斂，於是本研究整合決策樹與貝氏網路兩種資料分析方法，以決策樹作初步分類，整合專家意見建立出推論規則，並建構結果收斂之貝氏網路模型，不但能以計量方法達到分析目的，又能保留其中重要的知識脈絡，更可作為未來預測之用。

本研究選定台電作為個案，由於目前電力負載預測值之調整乃由專家諮詢會議討論後彙整開會記錄，再依未來環境之外生變數可能之情況，微調其負載預測值，卻沒有把專家意見的知識脈絡保存下來。本研究期能改善目前傳統以專家直接對預測值作微調，而沒有將輔助決策過程的知識架構完整呈現保存的不足之處，並針對部分專家分歧意見作處理，以機率為基礎的數值化貝氏網路模型呈現在專家知識中各變數相互影響的關係與程度。

經整體分析流程之構建及其內各步驟之執行與驗證，及研究結果中以決策樹與貝氏網路所建立的專家意見整合模型，可知本研究所提出之 IDTBN 模型能有效整合並收斂專家意見為兩大類，分別是經濟成長組與節能環保組，並在除去極少數分歧意見後，有效提升專家意見資料所建構出之貝氏網路的準確度達 98.43%。由以上結果可知，此整合性分析技術，可應用於專家意見整合上。

最後，經由本研究結果顯示，本研究所提之 IDTBN 模型可適用於台電之電力負載預測個案上，所提出之整合專家意見架構作為一輔助決策支援系統的角色，可輔助於電力長期負載預測或未來相關研究參考之用，亦即在現有台電之決策支援模型與其計量方法預測負載值結果之外，讓專家可參考此意見整合模型，瞭解關鍵環境變數之間的影响程度，視情況微調電力負載預測值。同時也可為其它專業領域之專家意見整合所參考並加以推廣發展。



## 第二節 建議

### 5.2.1 研究限制

本研究之研究範圍以 2007 年底至 2008 年初共三場專家諮詢會議之會議記錄及錄音檔，並過錄為符合本研究模型可用之資料格式作為實驗資料來源，主要分析對象為參與會議之電力、經濟、環保等研究與業務單位所提出之專家意見。然而本研究基於資料取得、應用實例之產業特性以及資料探勘技術等觀點之考量下，有其研究限制詳敘如下：

#### 一、資料觀察期間不足

本研究所蒐集到的專家意見資料來自於三個場次專家諮詢會議，雖說在實務上專家會議召開曠日費時，而且所費不貲，可是對於本研究來說，由於無法對每個專家進行細部數據的訪談，僅根據專家諮詢會議中的參與觀察、錄音、會議記錄、簡報資料等，進行資料萃取的工作，造成資料取得上的限制較大。這是後續研究可以補強的部分。

#### 二、應用實例之產業特性

受到電力負載所牽涉的層面與變數限制，本研究所著重之切入點可能也與其它產業不同，故研究結果的解義和評量標準，也需要依照實際應用的領域作調整。

#### 三、資料探勘技術之應用限制

資料探勘易受到資料品質的影響，因此以此類技術所建立的模型，能否真實呈現專家的知識推論脈絡，則牽涉到資料本身的準確度與完整程度，所以如果資料有偏差就會影響資料分析的結果，可能造成難以解釋分析結果的情況。每種分類分析方法都有其優缺點與限制，因此

視資料特性和分析目的選擇特定的方法，也就必須承受該方法在某些方面的不足之處，並要透過反覆實驗測試找到最適參數設定，並依實際情況進行調整，才能使資料在該分析方法的架構下，能準確呈現資料蘊含的意義。

## 5.2.2 未來研究方向

依據本研究各步驟之研究流程與分析結果，建議未來可以本研究為基礎，對下列各研究方向進行延伸與探討：

- 一、本研究的實驗中所使用的原始資料，乃是從專家諮詢會議錄音檔與會議記錄所過錄而成，這其中需要人工將資料轉換為實驗所需格式，若能結合語意辨識分析技術，從專家知識社群的討論中，自動找出關鍵詞彙與變數之間相互影響的關係程度，或提供友善的人機操作介面，讓專家得以直接以制式化的格式表達意見，則能使資料處理這部分減少耗費額外的人力與時間。
- 二、除了將所有專家意見整合建立為模型之外，能再依專業領域分類，各別將歸屬不同領域別的專家意見分別作整合模型之構建，如此則能比較在不同領域的專家，所重視的變數和提出的影響程度也許不盡相同，進一步分析比較不同背景對模型所造成的差異。
- 三、目前本研究是針對三次專家會議中的意見匯整進行模型分析，主要是建立一個以目前環境情況為主的靜態資料模型，但考慮到外在環境發生變化，而外生變數對於電力負載預測值的影響程度改變時，模型中之相關機率值應就實際情況進行調整，建議可以本研究架構為基礎，加以延伸並建立能自動修正微調機率分布的電力負載預測模型。

附錄一：專家諮詢會議資料過錄表（前 5 筆資料）

Raw data 來源 #1 第一次專家諮詢會議簡報內容		
專家背景	欄位名稱	重要敘述
簡報單位 a. 政府部門 b. 電力	1. 國際經濟發展情勢	<ul style="list-style-type: none"> <li>➢ 全球經濟穩定成長率可達 3.7%。</li> <li>➢ 短期略降為 3.3%。</li> <li>➢ 全球經濟成長可望帶動台灣經濟成長，進而影響電力負載攀升。</li> </ul>
	2. CO <sub>2</sub> 減量限制	<ul style="list-style-type: none"> <li>➢ 產生直接抑低電力負載效果。</li> <li>➢ 發揮間接抑低電力負載效果。</li> </ul>
	3. 環保主張	<ul style="list-style-type: none"> <li>➢ 帶動節約能源，抑低電力負載。</li> </ul>
	4. 國內經濟成長	<ul style="list-style-type: none"> <li>➢ 微幅攀升，年平均成長 4%（中案）。</li> <li>➢ 可望帶動電力負載攀升。</li> <li>➢ GDP 每增加 8.14%，則用電量將增加 9.32%。</li> </ul>
	5. 產業結構變化	<ul style="list-style-type: none"> <li>➢ 工業部門佔比預估從 29.2%降為 26.69%。</li> <li>➢ 服務業部門佔比預估從 69.59%提升為 73.27%。</li> <li>➢ 服務業部門佔比愈高、工業部門佔比愈低，可望抑低電力負載量。</li> </ul>
	6. 重大投資案	<ul style="list-style-type: none"> <li>➢ 可望帶動電力負載急劇攀升。</li> <li>➢ 未來台灣景氣好轉下，重大投資案也會增加。</li> </ul>
	7. 電力需求面管理措施	<ul style="list-style-type: none"> <li>➢ 既有措施並無新的成長。</li> <li>➢ 需求面管理成效，可望抑低電力負載量。</li> </ul>
	8. 人口成長預測	<ul style="list-style-type: none"> <li>➢ 人口成長會帶動電力負載自然成長。</li> <li>➢ 至 115 年台灣人口平均成長率為 0.05%。</li> </ul>
	9. 氣溫效果	<ul style="list-style-type: none"> <li>➢ 換算成夏季平均冷氣度，未來每年預估增加 1.5 度。</li> <li>➢ 對於尖峰負載影響甚鉅。</li> <li>➢ 冷氣度每上升 20.41%，則用電量將增加 0.7%。</li> </ul>
	10. 電價效果	<ul style="list-style-type: none"> <li>➢ 電價調昇，可望抑低電力負載量。</li> </ul>
	11. 能源技術提升	<ul style="list-style-type: none"> <li>➢ 例如全面使用 LED 省電照明、提升能源效率管理、提升冷凍空調效率等，將可使每年用電量減少 282.8 億度。</li> </ul>
	12. 物價上漲	<ul style="list-style-type: none"> <li>➢ 未提及。</li> </ul>
	13. 預估用電量呈成長趨勢	<ul style="list-style-type: none"> <li>➢ 預估用電量年平均成長率為 3.3%。</li> </ul>

	14. 預估尖峰負載呈成長趨勢	➤ 預估尖峰負載年平均成長率為 3.3%。
	15. 兩岸經貿關係	➤ 未提及。
	16. 經建目標	➤ 未提及。
	17. 節能目標	➤ 未提及。

Raw data 來源 #2 第二、三次專家諮詢會議簡報內容

專家背景	欄位名稱	重要敘述
簡報單位 a. 政府部門 b. 電力經濟	1. 國際經濟發展情勢	<ul style="list-style-type: none"> <li>➤ 國際經濟情勢溫和擴張。</li> <li>➤ 全球經濟成長可望帶動台灣經濟成長，進而影響電力負載攀升。</li> </ul>
	2. CO <sub>2</sub> 減量限制	<ul style="list-style-type: none"> <li>➤ 將造成產業重大影響，進而抑低電力負載。</li> <li>➤ 將造成經濟發展重大影響，進而抑低電力負載。</li> </ul>
	3. 環保主張	➤ 未提及。
	4. 國內經濟成長	<ul style="list-style-type: none"> <li>➤ 微幅攀升，年平均成長 4.1% (中案)。</li> <li>➤ 可望帶動電力負載攀升。</li> <li>➤ GDP 每增加 1%，則用電量將增加 0.84%。</li> </ul>
	5. 產業結構變化	<ul style="list-style-type: none"> <li>➤ 工業部門佔比預估從 30.2% 降為 28.29%。</li> <li>➤ 服務業部門佔比預估從 68.4% 提升為 72.5%。</li> <li>➤ 服務業部門佔比愈高、工業部門佔比愈低，可望抑低電力負載量。</li> </ul>
	6. 重大投資案	<ul style="list-style-type: none"> <li>➤ 可望如期完成，促使經濟發展。</li> <li>➤ 帶動電力負載成長。</li> <li>➤ 未來因有環評因素，仍須審慎樂觀。</li> </ul>
	7. 電力需求面管理措施	<ul style="list-style-type: none"> <li>➤ 預估需求面管理措施績效從 67.8 萬千瓦增加為 84.7 萬千瓦。</li> <li>➤ 需求面管理成效，可望抑低電力負載量。</li> </ul>
	8. 人口成長預測	<ul style="list-style-type: none"> <li>➤ 人口成長會帶動電力負載自然成長。</li> <li>➤ 至 115 年台灣人口平均成長率為 0.1%。</li> </ul>
	9. 氣溫效果	<ul style="list-style-type: none"> <li>➤ 對於尖峰負載影響甚鉅。</li> <li>➤ 預估冷氣度每年上升 1 度。</li> </ul>
	10. 電價效果	<ul style="list-style-type: none"> <li>➤ 電價調昇，可望抑低電力負載量。</li> <li>➤ 預估實質電價平均成長率為 2%。</li> </ul>
	11. 能源技術提升	➤ 例如全面使用 LED 省電照明、提升能源效率管理、提升冷凍空調效率等，將可使每年用

		電量減少 282.8 億度。
	12. 物價上漲	➤ 導致民間消費減少，抑低經濟成長。
	13. 預估用電量呈成長趨勢	➤ 預估用電量年平均成長率為 3.5%。
	14. 預估尖峰負載呈成長趨勢	➤ 預估尖峰負載年平均成長率為 3.5%。
	15. 兩岸經貿關係	➤ 未提及。
	16. 經建目標	➤ 未提及。
	17. 節能目標	➤ 未提及。

Raw data 來源 #3 第一次會議：能源局召集能源經濟專家		
專家背景	欄位名稱	重要敘述
1. 曾文清 a. 政府部門 b. 產業經濟	1. 國際經濟發展情勢	<ul style="list-style-type: none"> <li>➤ 全球經濟穩定成長率可達 3.7%。</li> <li>➤ 短期略降為 3.3%。</li> <li>➤ 全球經濟成長可望帶動台灣經濟成長，進而影響電力負載攀昇。</li> </ul>
	2. CO <sub>2</sub> 減量限制	<ul style="list-style-type: none"> <li>➤ 產生直接抑低電力負載效果。</li> <li>➤ 發揮間接抑低電力負載效果。</li> </ul>
	3. 環保主張	<ul style="list-style-type: none"> <li>➤ 帶動節約能源，抑低電力負載。</li> </ul>
	4. 國內經濟成長	<ul style="list-style-type: none"> <li>➤ 微幅攀昇，年平均成長 4% (中案)。</li> <li>➤ 可望帶動電力負載攀昇。</li> <li>➤ GDP 每增加 8.14%，則用電量將增加 9.32%。</li> </ul>
	5. 產業結構變化	<ul style="list-style-type: none"> <li>➤ 工業部門佔比預估從 29.2% 降為 26.69%。</li> <li>➤ 服務業部門佔比預估從 69.59% 提升為 73.27%。</li> <li>➤ 服務業部門佔比愈高、工業部門佔比愈低，可望抑低電力負載量。</li> </ul>
	6. 重大投資案	<ul style="list-style-type: none"> <li>➤ 可望帶動電力負載急劇攀昇。</li> <li>➤ 未來台灣景氣好轉下，重大投資案也會增加。</li> </ul>
	7. 電力需求面管理措施	<ul style="list-style-type: none"> <li>➤ 既有措施並無新的成長。</li> <li>➤ 需求面管理成效，可望抑低電力負載量。</li> </ul>
	8. 人口成長預測	<ul style="list-style-type: none"> <li>➤ 人口成長會帶動電力負載自然成長。</li> <li>➤ 至 115 年台灣人口平均成長率為 0.05%。</li> </ul>
	9. 氣溫效果	<ul style="list-style-type: none"> <li>➤ 換算成夏季平均冷氣度，未來每年預估增加 1.5 度。</li> </ul>

		<ul style="list-style-type: none"> <li>➤ 對於尖峰負載影響甚鉅。</li> <li>➤ 冷氣度每上升 20.41%，則用電量將增加 0.7%。</li> </ul>
	10. 電價效果	<ul style="list-style-type: none"> <li>➤ 電價調昇，可望抑低電力負載量。</li> </ul>
	11. 能源技術提升	<ul style="list-style-type: none"> <li>➤ 例如全面使用 LED 省電照明、提升能源效率管理、提升冷凍空調效率等，將可使每年用電量減少 282.8 億度。</li> </ul>
	12. 物價上漲	<ul style="list-style-type: none"> <li>➤ 未提及。</li> </ul>
	13. 預估用電量呈成長趨勢	<ul style="list-style-type: none"> <li>➤ 考量環評阻礙電廠設置，負載預測不宜低估。</li> <li>➤ 預估用電量年平均成長率為 3.65%。</li> </ul>
	14. 預估尖峰負載呈成長趨勢	<ul style="list-style-type: none"> <li>➤ 考量環評阻礙電廠設置，負載預測不宜低估。</li> <li>➤ 預估尖峰負載年平均成長率為 3.6%。</li> </ul>
	15. 兩岸經貿關係	<ul style="list-style-type: none"> <li>➤ 未提及。</li> </ul>
	16. 經建目標	<ul style="list-style-type: none"> <li>➤ 未提及。</li> </ul>
	17. 節能目標	<ul style="list-style-type: none"> <li>➤ 未提及。</li> </ul>

Raw data 來源 #3 第一次會議：能源局召集能源經濟專家

專家背景	欄位名稱	重要敘述
2. 何金巡 a. 政府部門 b. 經濟	1. 國際經濟發展情勢	<ul style="list-style-type: none"> <li>➤ 全球經濟穩定成長率可達 3.7%。</li> <li>➤ 短期略降為 3.3%。</li> <li>➤ 全球經濟成長可望帶動台灣經濟成長，進而影響電力負載攀昇。</li> </ul>
	2. CO <sub>2</sub> 減量限制	<ul style="list-style-type: none"> <li>➤ 限制經濟發展，從而抑低電力負載。</li> <li>➤ 須兼顧環境永續，從而抑低電力負載。</li> </ul>
	3. 環保主張	<ul style="list-style-type: none"> <li>➤ 帶動節約能源，抑低電力負載。</li> </ul>
	4. 國內經濟成長	<ul style="list-style-type: none"> <li>➤ 微幅攀昇，年平均成長 4% (中案)。</li> <li>➤ 可望帶動電力負載攀昇。</li> <li>➤ GDP 每增加 8.14%，則用電量將增加 9.32%。</li> </ul>
	5. 產業結構變化	<ul style="list-style-type: none"> <li>➤ 工業部門佔比預估從 29.2% 降為 26.69%。</li> <li>➤ 服務業部門佔比預估從 69.59% 提升為 73.27%。</li> <li>➤ 服務業部門佔比愈高、工業部門佔比愈低，可望抑低電力負載量。</li> </ul>
	6. 重大投資案	<ul style="list-style-type: none"> <li>➤ 可望帶動電力負載急劇攀昇。</li> </ul>

		<ul style="list-style-type: none"> <li>➤ 未來台灣景氣好轉下，重大投資案也會增加。</li> </ul>
	7. 電力需求面管理措施	<ul style="list-style-type: none"> <li>➤ 既有措施並無新的成長。</li> <li>➤ 需求面管理成效，可望抑低電力負載量。</li> </ul>
	8. 人口成長預測	<ul style="list-style-type: none"> <li>➤ 人口成長會帶動電力負載自然成長。</li> <li>➤ 至 115 年台灣人口平均成長率為 0.05%。</li> </ul>
	9. 氣溫效果	<ul style="list-style-type: none"> <li>➤ 換算成夏季平均冷氣度，未來每年預估增加 1.5 度。</li> <li>➤ 對於尖峰負載影響甚鉅。</li> <li>➤ 冷氣度每上升 20.41%，則用電量將增加 0.7%。</li> </ul>
	10. 電價效果	<ul style="list-style-type: none"> <li>➤ 電價調昇，可望抑低電力負載量。</li> </ul>
	11. 能源技術提升	<ul style="list-style-type: none"> <li>➤ 例如全面使用 LED 省電照明、提升能源效率管理、提升冷凍空調效率等，將可使每年用電量減少 282.8 億度。</li> </ul>
	12. 物價上漲	<ul style="list-style-type: none"> <li>➤ 未提及。</li> </ul>
	13. 預估用電量呈成長趨勢	<ul style="list-style-type: none"> <li>➤ 預估用電量年平均成長率為 3.3%。</li> </ul>
	14. 預估尖峰負載呈成長趨勢	<ul style="list-style-type: none"> <li>➤ 預估尖峰負載年平均成長率為 3.3%。</li> </ul>
	15. 兩岸經貿關係	<ul style="list-style-type: none"> <li>➤ 未提及。</li> </ul>
	16. 經建目標	<ul style="list-style-type: none"> <li>➤ 未提及。</li> </ul>
	17. 節能目標	<ul style="list-style-type: none"> <li>➤ 未提及。</li> </ul>

Raw data 來源 #3 第一次會議：能源局召集能源經濟專家

專家背景	欄位名稱	重要敘述
3. 余勝雄 a. 政府部門 b. 電力電機	1. 國際經濟發展情勢	<ul style="list-style-type: none"> <li>➤ 全球經濟穩定成長率可達 3.7%。</li> <li>➤ 短期略降為 3.3%。</li> <li>➤ 全球經濟成長可望帶動台灣經濟成長，進而影響電力負載攀昇。</li> </ul>
	2. CO <sub>2</sub> 減量限制	<ul style="list-style-type: none"> <li>➤ 對於電力負載會有衝擊性的影響。</li> </ul>
	3. 環保主張	<ul style="list-style-type: none"> <li>➤ 帶動節約能源，抑低電力負載。</li> </ul>
	4. 國內經濟成長	<ul style="list-style-type: none"> <li>➤ 微幅攀昇，年平均成長 4% (中案)。</li> <li>➤ 可望帶動電力負載攀昇。</li> <li>➤ GDP 每增加 8.14%，則用電量將增加 9.32%。</li> </ul>
	5. 產業結構變化	<ul style="list-style-type: none"> <li>➤ 工業部門佔比預估從 29.2% 降為 26.69%。</li> </ul>

		<ul style="list-style-type: none"> <li>➤ 服務業部門佔比預估從 69.59%提升為 73.27%。</li> <li>➤ 服務業部門佔比愈高、工業部門佔比愈低，可望抑低電力負載量。</li> </ul>
	6. 重大投資案	<ul style="list-style-type: none"> <li>➤ 可望帶動電力負載急劇攀升。</li> <li>➤ 未來台灣景氣好轉下，重大投資案也會增加。</li> </ul>
	7. 電力需求面管理措施	<ul style="list-style-type: none"> <li>➤ 既有措施並無新的成長。</li> <li>➤ 需求面管理成效，可望抑低電力負載量。</li> </ul>
	8. 人口成長預測	<ul style="list-style-type: none"> <li>➤ 人口成長會帶動電力負載自然成長。</li> <li>➤ 至 115 年台灣人口平均成長率為 0.05%。</li> </ul>
	9. 氣溫效果	<ul style="list-style-type: none"> <li>➤ 換算成夏季平均冷氣度，未來每年預估增加 1.5 度。</li> <li>➤ 對於尖峰負載影響甚鉅。</li> <li>➤ 冷氣度每上升 20.41%，則用電量將增加 0.7%。</li> </ul>
	10. 電價效果	<ul style="list-style-type: none"> <li>➤ 電價調昇，對於未來電力負載量發展會有衝擊。</li> <li>➤ 電價調整幅度應足以反應目前台電之虧損。</li> </ul>
	11. 能源技術提升	<ul style="list-style-type: none"> <li>➤ 例如全面使用 LED 省電照明、提升能源效率管理、提升冷凍空調效率等，將可使每年用電量減少 282.8 億度。</li> </ul>
	12. 物價上漲	<ul style="list-style-type: none"> <li>➤ 未提及。</li> </ul>
	13. 預估用電量呈成長趨勢	<ul style="list-style-type: none"> <li>➤ 預估用電量年平均成長率為 3.3%。</li> </ul>
	14. 預估尖峰負載呈成長趨勢	<ul style="list-style-type: none"> <li>➤ 尖峰負載預估值因基期氣溫及降雨較多，以致偏低。</li> <li>➤ 預估尖峰負載年平均成長率為 3.6%。</li> </ul>
	15. 兩岸經貿關係	<ul style="list-style-type: none"> <li>➤ 對於電力負載會有衝擊性的影響。</li> </ul>
	16. 經建目標	<ul style="list-style-type: none"> <li>➤ 未提及。</li> </ul>
	17. 節能目標	<ul style="list-style-type: none"> <li>➤ 未提及。</li> </ul>



附錄二：資料過錄格式表

變數內容	變數名稱	數值範圍	數值代碼
1. 流水號	No	001-163	依序給於流水號
2. 部門	Sec	1-4	1. 政府部門 2. 民間智庫 3. 財團法人 4. 學術界
3. 專長	Exp	1-3	1. 電力經濟 2. 工業、經濟、產業經濟、會計、環保 3. 電力電機
4. 國際經濟成長率 (%)	X1	空白、依實際數值所示。	空白：代表未提及，為缺值。 餘者：填入實際數值。
5. CO <sub>2</sub> 減量效果 (強度)	X2	空白、1-9	空白：代表未提及，為缺值。 強度：1-9。
6. 環保效果 (有或無)	X3	0、1	0：未提及。 1：認為環保具抑低電力負載量效果。
7. 國內平均經濟成長率 (%)	X4.1	空白、依實際數值所示。	空白：代表未提及，為缺值。 餘者：填入實際數值。
8. GDP 所得彈性 (彈性)	X4.2	空白、依實際數值所示。	空白：代表未提及，為缺值。 餘者：填入實際數值。
9. 工業部門佔比預估 (%)	X5.1	空白、依實際數值所示。	空白：代表未提及，為缺值。 餘者：填入實際數值。
10. 服務業部門佔比預估 (%)	X5.2	空白、依實際數值所示。	空白：代表未提及，為缺值。 餘者：填入實際數值。
11. 重大投資案 (共兩類)	X6	空白、1-2	0：未提及。 1：投資案增加，電力負載量增加。 2：產業外移，抑低電力負載量。
12. DSM 抑低效果 (電力需求面管理措施)	X7	0、1	0：未提及。 1：具抑低電力負載效果。
13. 預估人口成長率 (%)	X8	空白、依實際數值所示。	空白：代表未提及，為缺值。 餘者：填入實際數值。
14. 氣溫效果，增	X9	空白、依實際	空白：代表未提及，為缺值。

加負載量預估 (度)		數值所示。	餘者：填入實際數值。
15. 預估電價調 升(%)	X10	空白、依實際 數值所示。	空白：代表未提及，為缺值。 餘者：填入實際數值。
16. 能源技術提 升抑低用電量效 果(億度)	X11	空白、依實際 數值所示。	空白：代表未提及，為缺值。 餘者：填入實際數值。
17. 物價上漲抑 低電力負載效果 (有或無)	X12	0、1	0：未提及。 1：具抑低電力負載效果。
18. 預估用電量 年平均成長率 (%)	X13	空白、依實際 數值所示。	空白：代表未提及，為缺值。 餘者：填入實際數值。
19. 預估尖峰負 載年平均成長率 (%)	X14	空白、依實際 數值所示。	空白：代表未提及，為缺值。 餘者：填入實際數值。
20. 節能政策目 標抑低電力負載 效果(有或無)	X17	0、1	0：未提及。 1：認為環保具抑低電力負載量效 果。



No.	X1		X2		X3		X4.1		X4.2		X5.1		X5.2		X6		X7		X8		X9		X10		X11		X12		X13		X17		
	Nominal		Nominal		Nominal		Nominal		Nominal		Nominal		Nominal		Nominal		Nominal		Nominal		Nominal		Nominal		Nominal		Nominal		Nominal		Nominal		
31	3.3	8	1	4	0.84	27.62	71.38	1	0	0.05	1.5	2	282.8	0	3.25-3...	1																	
32	3.3	8	1	4	0.84	27.62	71.38	1	0	0.05	1.5	2	282.8	0	3.25-3...	1																	
33	3.3	8	1	4	0.84	27.62	71.38	1	0	0.05	1.5	2	282.8	0	3.25-3...	1																	
34	3.3	8	1	4	0.84	27.62	71.38	1	0	0.05	1.5	2	282.8	0	3.25-3...	1																	
35	3.3	8	1	4	0.84	27.62	71.38	1	0	0.05	1.5	2	282.8	0	3.25-3...	1																	
36	3.3	8	1	4	0.84	27.62	71.38	1	0	0.05	1.5	2	282.8	0	3.46-3...	0																	
37	3.3	8	1	4	0.84	27.62	71.38	1	0	0.05	1.5	2	282.8	0	3.46-3...	0																	
38	3.3	8	1	4	0.84	27.62	71.38	1	0	0.05	1.5	2	282.8	0	3.46-3...	0																	
39	3.3	8	1	4	0.84	27.62	71.38	1	0	0.05	1.5	2	282.8	0	3.46-3...	0																	
40	3.3	8	1	4	0.84	27.62	71.38	1	0	0.05	1.5	2	282.8	0	3.25-3...	0																	
41	3.3	8	1	2.7	0.65	25.9	73.5	1	1	0.05	1.5	2	282.8	0																			
42	3.3	8	1	2.7	0.65	25.9	73.5	1	1	0.05	1.5	2	282.8	0																			
43	3.3	8	1	2.7	0.65	25.9	73.5	1	1	0.05	1.5	2	282.8	0																			
44	3.3	8	1	2.7	0.65	25.9	73.5	1	1	0.05	1.5	2	282.8	0																			
45	3.3	8	1	2.7	0.65	25.9	73.5	1	1	0.05	1.5	2	282.8	0																			
46	3.3		1	4	0.84	27.62	71.38	1	0	0.05	1.5	2	282.8	0	3.46-3...	0																	
47	3.3		1	4	0.84	27.62	71.38	1	0	0.05	1.5	2	282.8	0	3.46-3...	0																	
48	3.3		1	4	0.84	27.62	71.38	1	0	0.05	1.5	2	282.8	0	3.46-3...	0																	
49	3.3		1	4	0.84	27.62	71.38	1	0	0.05	1.5	2	282.8	0	3.46-3...	0																	
50	3.3		1	4	0.84	27.62	71.38	1	0	0.05	1.5	2	282.8	0	3.46-3...	0																	
51	3.3		1	4	0.84	27.62	71.38	1	0	0.05	1.5	2	282.8	0	3.46-3...	0																	
52	3.3		1	4	0.84	27.62	71.38	1	0	0.05	1.5	2	282.8	0	3.46-3...	0																	
53	3.3		1	4	0.84	27.62	71.38	1	0	0.05	1.5	2	282.8	0	3.46-3...	0																	
54	3.3		1	4	0.84	27.62	71.38	1	0	0.05	1.5	2	282.8	0	3.46-3...	0																	
55	3.3		1	4	0.84	27.62	71.38	1	0	0.05	1.5	2	282.8	0	3.46-3...	0																	
56	3.3	8	1	4	0.84	27.62	71.38	1	0	0.05	1.5	2	282.8	0	3.25-3...	1																	
57	3.3	8	1	4	0.84	27.62	71.38	1	0	0.05	1.5	2	282.8	0	3.25-3...	1																	
58	3.3	8	1	4	0.84	27.62	71.38	1	0	0.05	1.5	2	282.8	0	3.25-3...	1																	
59	3.3	8	1	4	0.84	27.62	71.38	1	0	0.05	1.5	2	282.8	0	3.25-3...	1																	
60	3.3	8	1	4	0.84	27.62	71.38	1	0	0.05	1.5	2	282.8	0	3.25-3...	1																	

## 參考文獻

### 一、中文部分

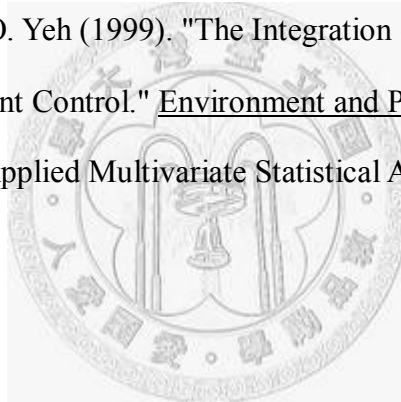
- 【1】 許士軍(1982),「管理學」,二版,東華書局。
- 【2】 廖嘉雄(1999),「以貝氏網路為基礎的知識處理機制之研究」,東海大學工業工程研究所碩士論文。
- 【3】 潘俊帆(2000),「以貝氏網路為基礎的個人差異化知識處理機制」,東海大學工業工程研究所碩士論文。
- 【4】 范樹根(2005),「結合模糊集合與貝氏分類 應用於無線設備測試良率之研究」,中原大學資訊管理研究所碩士論文。
- 【5】 張振魁(2000),「以類神經網路提高股票單日交易策略之獲利」,國立中央大學資訊管理研究所碩士論文。
- 【6】 周政宏(1995),「神經網路-理論與實務」,松崗圖書。
- 【7】 林蔓蓁(1994),「銀行授信客戶之風險評估」,中央大學資訊管理研究所碩士論文。
- 【8】 楊孟龍(2000),「類神經網路於股價波段預測及選股之應用」,中央大學資訊管理研究所碩士論文。
- 【9】 葉怡成(1997),「應用類神經網路」,儒林圖書公司。
- 【10】 許家成(2001),「案例式推理於地理資訊系統的應用-以颱風路徑預測為例」,國立臺灣大學地理環境資源研究所碩士論文。
- 【11】 戴志洋(2001),「文章信賴分配模型與綜合加權評分機制:運用貝氏網路於知識討論群組系統」,台灣大學資訊管理研究所碩士論文。
- 【12】 許哲強(2002),「台灣區域電力負載預測分析系統之建立與應用研究」,國立成功大學資訊工程研究所博士論文。
- 【13】 賴正文(2001),「人工神經網路在長期負載預測之應用研究」,台電公司委託研究計畫。
- 【14】 陳家榮(1993),「負載預測發展及制度之規劃研究」,台灣電力公司。

- 【15】 葉怡成(1995),「類神經網路模型應用與實作」,儒林圖書。
- 【16】 台灣經濟研究院,歷年「國內總體經濟預測暨景氣動向調查月報」。
- 【17】 台灣電力公司企劃處,歷年「長期負載預測報告」。
- 【18】 台灣電力公司企劃處,歷年「統計年報」。
- 【19】 吳仁和、林信惠(2001),「系統分析與設計:理論與實務應用」,智勝文化出版。
- 【20】 徐世勳、李篤華、蘇漢邦、林幸君(2005),「台灣產業結構變動之動態一般均衡預測—TAIGEM-E模型的應用」。
- 【21】 許哲強(2002),「台灣區域電力負載預測分析系統之建立與應用研究」,成大資工所博士論文。
- 【22】 陳順宇(2005),「多變量分析」,華泰文化出版,台北。
- 【23】 梁定澎(2006),「決策支援系統與企業智慧」,智勝文化出版,台北。
- 【24】 萬哲鈺(2003),「時間序列講義」。
- 【25】 賴正文等(2001),「人工神經網路在長期負載預測之應用研究」,台電公司委託研究計畫。
- 【26】 謝邦昌、易丹輝編著(2003),「統計資料分析:以 STATISTICA6.0 為例」,中華資料採礦協會出版。
- 【27】 Chava Frankfort-Nachmias & David Nachmias 著,潘明宏、陳志瑋譯(2003),「最新社會科學研究方法」,韋伯文化。
- 【28】 哈嘉琪(2005),「2030 年數位生活情境分析-數位家庭應用之研究」,國立交通大學科技管理研究所碩士論文。
- 【29】 鄭味亭(2007),「支援情境變動下的潛在目標用戶探勘系統之構建—以可停電力潛在用戶探勘為例」,國立臺灣大學資訊管理研究所碩士論文。

## 二、英文部分

- 【1】 Baupin, N. and K. Zreik (2000). "Remote decision support system: a distributed information management system." Knowledge-Based Systems **13**(1): 37-46.
- 【2】 Berry, M. J. and G. Linoff (1997). Data Mining Techniques: For Marketing, Sales, and Customer Support, John Wiley & Sons, Inc. New York, NY, USA.
- 【3】 Burnside, E. S. (2005). "Bayesian networks Computer-assisted diagnosis support in radiology1." Academic Radiology **12**(4): 422-430.
- 【4】 Charniak, E. (1991). "Bayesian Networks without Tears." AI Magazine **12**(4): 50-63.
- 【5】 Enrique, C., Jose, et al. (1997). Expert Systems and Probabilistic Network Models. New York, Springer-Verlag.
- 【6】 Heckerman, D. (1996). "A tutorial on learning with Bayesian networks." Microsoft Technical Report MSR-TR-05-06.
- 【7】 Heckerman, D. and J. S. Breese (1995). "Causal independence for probability assessment and inference using Bayesian networks." Systems, Man and Cybernetics, IEEE Transactions **26**(6): 826-831.
- 【8】 Heckerman, D., A. Mamdani, et al. (1995). "Real-world applications of Bayesian networks." Communications of the ACM **38**(3): 24 - 26.
- 【9】 Janssens, D., G. Wets, et al. (2006). "Integrating Bayesian networks and decision trees in a sequential rule-based transportation model." European Journal of Operational Research **175**(1): 16-34.
- 【10】 Kolodner, J. (1993). Case-based reasoning, Morgan Kaufmann Publishers Inc. San Francisco, CA, USA.
- 【11】 Kuan, C. M. and H. White (1994). "Artificial neural networks: an econometric perspective." Econometric Reviews **13**(1): 1-91.

- 【12】 Mingers, J. (1989). "An Empirical Comparison of Pruning Methods for Decision Tree Induction." Machine Learning **4**(2): 227-243.
- 【13】 Pearl, J. (1988). Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference, Morgan Kaufmann.
- 【14】 Quinlan, J. R. (1986). "Induction of decision trees." Machine Learning **1**(1): 81-106.
- 【15】 Quinlan, J. R. (1993). C4. 5: Programs for Machine Learning, Morgan Kaufmann.
- 【16】 Shachter, R. D. (1988). "Probabilistic Inference and Influence Diagrams." Operations Research **36**(4): 589-604.
- 【17】 Shi, X. and A. G. O. Yeh (1999). "The Integration of Case-based Systems and GIS in Development Control." Environment and Planning B **26**(3): 345-364.
- 【18】 Johnson (2002). Applied Multivariate Statistical Analysis. Prentice Hall.





### 三、參考網頁

- 【1】 Weka 3 - Data Mining with Open Source Machine Learning Software in Java.  
Available at: <http://www.cs.waikato.ac.nz/ml/weka/>. Accessed 25 Feb 2008
- 【2】 Classification via Decision Trees in WEKA. Available at:  
<http://maya.cs.depaul.edu/~classes/ect584/WEKA/classify.html>. Accessed 25 Feb 2008
- 【3】 關於類神經網路，網址：  
<http://www.gct.ntou.edu.tw/Lab/aiwww/neural.html>，閱覽日期：2008/4/16
- 【4】 台電企業網站，網址：<http://www.taipower.com.tw>，閱覽日期：2008/3/22
- 【5】 行政院主計處，中華民國總體經濟統計資訊網，閱覽日期：2008/3/22

