

# 行政院國家科學委員會專題研究計畫 成果報告

## Data Warehousing 及 Data Mining 在營建財務及成本管理 之決策支援之研究(I)

計畫類別：個別型計畫

計畫編號：NSC91-2211-E-002-090-

執行期間：91年08月01日至92年07月31日

執行單位：國立臺灣大學土木工程學系暨研究所

計畫主持人：荷世平

計畫參與人員：林芳瑩

報告類型：精簡報告

處理方式：本計畫可公開查詢

中 華 民 國 93 年 2 月 16 日

# 行政院國家科學委員會補助專題研究計畫期中進度報告

## Data Warehousing 及 Data Mining 在營建財務及成本管理 之決策支援系統( I )

計畫類別：個別型計畫

計畫編號：NSC 91-2211-E-002-090-

執行期間：91年08月01日至92年07月31日

計畫主持人：荷世平

共同主持人：

計畫參與人員：林芳瑩

成果報告類型(依經費核定清單規定繳交)：精簡報告

本成果報告包括以下應繳交之附件：

- 赴國外出差或研習心得報告一份
- 赴大陸地區出差或研習心得報告一份
- 出席國際學術會議心得報告及發表之論文各一份
- 國際合作研究計畫國外研究報告書一份

處理方式：除產學合作研究計畫、提升產業技術及人才培育研究計畫、  
列管計畫及下列情形者外，得立即公開查詢

涉及專利或其他智慧財產權， 一年  二年後可公開查詢

執行單位：國立台灣大學土木工程學系暨研究所

中華民國九十二年十月三十一日

# Data Warehousing 及 Data Mining 在營建財務及成本管理之決策支援系統( I )

荷世平\*

## 摘要

資料採擷(Data Mining)於 1990 年開始蓬勃發展之際,1992 年 M.Y. 便首度以 ID3 之演算法應用於財務危機預測上,其以事實導向方式選取變數,除可更全方位的考量影響變數外,更可避免專家經驗選取變數之人為偏差,本研究首先將就財務之資料,驗證資料採擷(Data Mining)進行之分析結果,優於過去以專家經驗選擇較佳變數後再進行分析之結果。未來一年,本研究期以 John Argenti (1976)所提出之『企業生存因素』(Factors Affecting Survival of Business)作為研究背景理論,期能在上百個與營建業相關之變數中以序列分析(Sequence Pattern Analysis)提出財務危機發生之階段序列型樣(Sequence Pattern),再採用集群(Cluster)與 ID3 決策樹(Decision Tree),選取各階段財務危機預測之相關變數與其關聯性,以作為營建業者自我省察財務狀況之指標,更可作為資金提供者之信用評估指標,進而降低財務危機之發生,鞏固營建業與金融體系財務之健全。而成本管理之相關研究亦於本研究成果之後續發展中說明研究進度。

關鍵字：財務危機(Financial Distress)、成本管理(Cost Management)、營建業(Construction Industry)、資料採擷(Data Mining)、決策支援系統(Decision Support System)。

# **Financial and Cost Management in Data Warehousing and Data Mining : Building a Decision Support System in Construction**

## **Industry ( I )**

S. Ping Ho\*

### **Abstract**

The purpose of this study is to develop a dynamic prediction model for financial distress in construction industry. Three main issues are discussed in this study. First, we study the basic process of financial distress in construction and identify the dependencies used for sequence pattern technique. Second, this study compares two methodologies in financial distress prediction for their accuracy and stability. Lastly, the performances of various distress prediction approaches will be evaluated and compared. The impacts of economic, industrial, and business characteristics on prediction performance will be assessed. Real world financial distress data and statistics from the construction companies in Taiwan are used to formulate and validate the distress prediction models derived in this study. This research expects to provide a quantitative framework in evaluating the financial standing of firms in construction industry. The financial standing evaluation information is crucial to both firm's managers for understanding the clients and suppliers, and fund suppliers for assessing the credit worthiness of the firm. The early awareness of a potential financial distress can also help to reduce the overall numbers of actual distress in construction industry.

**KEY WORD** : Financial Distress 、 Cost Management 、 Construction Industry 、 Data Mining 、 Decision Support System ◦

## 壹、前言

資料倉儲(Data Warehousing)與資料採擷(Data Mining)之發展，為知識發掘技術 1990 年代晚期以回顧性質的主動傳輸方式解決企業問題的一種決策支援工具，其可針對企業所欲解決的問題，藉由過去儲存的資料日積月累，經整理後形成資料庫的形式，經資料前置整理(Data Prepare)，以資料採擷(Data Mining)的方法取代專家經年累月的經驗，以事實中擷取之智慧取代過專家經驗的解決方案。

民國 84 年，營建產業發生財務危機之公司接踵而來，銀行也因不良資產的累積導致逾放比過高，但若營建業公司可預知其財務狀況，則可是先加以預防，而信用供應商或公司債權人，或商業銀行等其他提供信用的金融機構，若可預先得知此債務人或授信對象之財務狀況，藉此降低借款之錯誤決策，讓不良公司無所遁形，進一步改善金融機構之體質，更間接鞏固金融市場。因此，本研究將以資料採擷(Data Mining)進行『營建業財務危機預測模型』之研究，希望能發掘適合營建業之財務危機預測模型，確保營建業與金融市場財務之健全。

本文之章節結構將遵循之研究流程，包括壹、緒論，貳、文獻回顧，參、研究方法與流程，肆、資料蒐集與整理，伍、分析結果與陸、結論與後續研究說明，茲如後所述。

## 貳、文獻回顧

### 2.1 財務危機之定義

過去許多研究學者對財務危機之定義，不外乎無法償還當期利息、企業重整、宣告破產等，茲整理如附錄表 1『財務危機定義之彙整表』，而本研究財務危機之定義將採用『台灣經濟新經濟報(TAIWAN ECONOMIC JOURNAL DATA BANK，簡稱 TEJ)信用風險評估之 9 種事件』，包括公司宣佈倒閉破產、公司聲明重整、公司跳票、向外界紓困求援，公司被外界接管、CPA 意見，對其繼續經營有疑慮、公司淨值為負、公司轉列全額交割股或下市與公司財務吃緊停工等九種事件將列為本研究財務危機之操作型定義。

### 2.2 財務危機之理論

財務危機理論可為實證時邏輯上之引導，並確保預測模型可應用在不同資料的研究上(Scott, 1981)，因此包括 Wilcox(1971, 1973)繼 Feller(1968)之後將賭博者損失模型之機率理論(Gambler's Ruin Model of Probability Theory)應用於財務危機預測模型之建立，但因大部份樣本與假設不符而於 1976 年屏棄理論方程式；Black-Scholes(1973)以選擇權評價理論(Option Pricing Theory)提出營利與現金流量變數。

John Argenti (1976)提出『企業生存因素』(Factors Affecting Survival of Business)包括內在因素(Internal Factor)之企業特性 Business Characteristics)與外在因素(External Factor)產業特性(Industrial Characteristics)、經濟環境(Economics Characteristics)等，文中指出若依企業生存理論，當財務危機預測模型指出企業保持最低維生水準之財務與管理資產是困難重重，則將有高度財務危機甚至破產之機會，建議債權人盡速回收對其財務上之支應。Scapens, Ryan and Fletcher(1981)提出由 Thom(1972)

發展之巨災理論(Catastrophe Theory),能從會計資訊中解釋企業債信(Credit Worthiness)突然改變之情況,並說明一體系達平衡狀態之因素,隨時間逐漸變化而造成體系有昭一日突然改變之行為,文章最後藉由會計資訊與授信者反應行為間之關係建立極端巨災模型(Cusp Catastrophe Model);Asquith et al (1994)以財務解剖之研究提出利息費用太高、營業績效比同業差與產業衰退等三項因素;Bhattacharya(1995)以財務比率分析之資金原理(Funds Theory),解釋企業生週期不同階段之財務狀況,其所採用之財務比率為資金週轉率(Finds Turnover Ratio)與資金流動率(Funds Flow Ratio),並將這些資金比率綜合起來導出單一企業健全比率(Health Ratio),用來預測企業之財務危機。

根據 Roberta Penny Marquette(1980)提出企業的失敗(Failure)為一動態過程(Dynamic Process)。可知財務危機乃是採階段性發生;George Foster(1978)將財務危機依據理論之推導將其階段區分為:主力產品銷售頓挫(Sales Decline in Major Products)、短期債務拖延償付(Deferment of Payments to Short-term Creditors)、優先股權未能發放(Omission of Preferred Dividends)、債券到期違約未付(Default of Bond)與宣告破產(Filing of Bankruptcy)等五階段;陳鞏榮(1983)根據 John Argenti (1976)所提出之『企業生存因素』(Factors Affecting Survival of Business)理論,將財務危機定義為包括財務危機、失調與破產倒閉等三階段;而 Lau(1987)以專家經驗之方式將財務危機定義為包括財務穩定、停止支付或大幅減少支出(40%以上)、貸款違約、受破產管束與破產清算等五階段。

由上述驗證 Bhattacharya(1995)提出『企業生命不同階段將有不同之財務狀況』更加合理,亦為本研究進行財務危機階段之理論根基,而與過去相關之研究,差異點在於過去為監督式之研究法,本研究期以非監督式,且類別間之距離相同之叢集法(Cluster)進行研究,以補足過去區別分析法屬值不屬量之限制。2.3 財務危機之實證模型

Foulke(1937)所倡導的驗證主義(Empiricism),指出唯有經得起實證考驗的知識,才是有用的知識。Scott(1981)亦認為實證模型與理論模型並不十分吻合,實證應証既有理論,而既有理論除提供時正進行上之支持外,同時也驗證模型之適切性。Jones(1987)更進一步提出對財務危機預測之研究者而言,相較於正確模型之建立,缺乏完整理論,並非嚴重之缺陷。因此以下首先就 Beaver(1966)單變量與 Altman(1968)多變量之創始者說明財務危機預測之實證模型,並簡單介紹人工智慧方法於財務危機預測應用之現況。Beaver(1966)以公司財務比率的資料,首度應用單變量迴歸預測財務危機,並首先採用二分類選擇法(Binary Choice),與隨機抽樣後以 A-B 樣本配對法以使公司規模與產業一致之樣本設計,更為後來的研究者樹立典範。

直至 Mark E. Zmijewski(1984)指出此非隨機抽樣法(即 A-B 配對法)有兩種錯誤,第一種為選擇基礎偏差(Choice-Based Biases),因為現實中財務危機企業僅佔極少數,以 A-B 之 1:1 或 1:2 之配對方式,將造成危機企業過度抽樣(Over Sampling);第二種為樣本選擇偏差(Sample Selection Biases),亦即以完整樣本作為抽樣之選擇標準,而實際上,財務危機企業有很多無法具有完整資料,但其強調此誤差所造成之結果並不一定會在整體歸類效果(Overall Classification)和預測上有顯著之改變,只有個別群體會有顯著影響,且若考量成本效益,非隨機抽樣法有其存在之必要性。

而財務危機的發生乃伴隨多項因素,以使單變量迴歸之預測能力受限,Altman(1968)隨即以逐步多元區別分析(Stepwise Multiple Discriminant Analysis;簡稱 MDA)雖考量

更多因素，並 Z-Scores 之迴歸式博得大眾初步之認同，但後來研究發現此一研究方法將遺漏重要變數(losing important variable)，進而嚴重影響預估能力而遭批評。

1990 年人工智慧可季逐漸發展成形，Odom、Sharda (1990)、Cadden, D. T.、Rahimian(1991)、Singh hote、Vitmani(1992)則紛紛引入於財務危機預測模型之應用領域，其分別以區別分析與導傳遞類神經網路進行預測；Salchenberger、Cinar、Lash(1992)以 Logit 機率模型、導傳遞類神經網路進行預測；Tam, K. y. and Kiang。其中，類神經網路最為人詬病乃因其運作過程如一黑箱子(Black Box)，使研究者無從得知研究結果之成因，而受批評。區別分析則是因函數具正負符號易被窗飾(Window Dressing)，且區分點為一絕對水準，使其預測能力受質疑。而 Logit 機率模型則是計算較為繁瑣。

M. Y. (1992)則以區別分析、Logit 機率模型、ID3 決策樹、導傳遞類神經進行預測，其中 ID3 決策樹即為資料採擷(Data Mining)之方法。Armonk, Summer (1999)更以資料採擷(Data Mining)進行動態的模型預測財務危機；而 Dean P. Foster and Robert A. Stine 亦以資料採擷(Data Mining)進行公司破產預測模型變數之選擇，將可考量變數擴充至上百個，使財務危機預測模型之考量因子更具全面性，進而提高預測之準確性。

#### 2.4 資料倉儲(Data Warehousing)與資料採擷(Data Mining)

資料倉儲(Data Warehousing)與資料採擷(Data Mining)等詳細之文獻回顧，情參閱本研究 92 年 1 月所提出之期中報告，在此不再贅述。

### 參、研究方法與流程

Altman(1993)提出各產業有因其所持有之產業特性各不相同，因此所產生之財務特性亦有所差異，因此研究首先將研究之產業索訂於營建業，而為考量資料之公信力，研究中財務危機預測模型之建立資料，將以台灣新經濟報資料庫(TAIWAN ECONOMIC JOURNAL DATA BANK)經審核之資料為研究資料。

本研究將以 IBM DB2 建立資料採擷之資料倉儲(Data Warehousing)，在輔以 Intelligent Miner 建立採擷模型，以分類(Classification)與類神經(N Network)之方式建立財務預測模型。其步驟如下所述：

1. 資料蒐集、整理。
2. 全部欄位進行財務正常與危機公司之分類。
3. 類神經進行敏感度分析。
4. 取敏感度大於 1 之欄位，再進行樹狀分類。
5. 比較兩模型之區別能力。

### 肆、資料蒐集與整理

本財務危機模型建立將以民國 82~民國 91 年營建業上市上櫃公司，資料期間包括營建業之景氣完整循環週期，資料內容共包括 40 家上市上櫃之營建公司，如附錄表 2 所示。其中，財務正常之公司 31 家，資料共 1301 筆；財務危機公司 9 家，資料共 162 筆；全部資料比數為 1463 筆(每筆定義：每家公司，每季之財務資料)。而以公司危機發生與否為分類

之類別標籤，進行分類與預測，因此須外加之資料欄位『財務危機公司』，其仍依台灣新經濟報資料庫(TAIWAN ECONOMIC JOURNAL DATA BANK)中財務危機公司之資料(如附錄表 3 所示)所標記，其標記原則如下：

1. 『財務危機公司』資料欄位型態：絕對數值。
2. 資料內容：0、1。
3. 資料內容定義---0：財務正常公司，1：財務危機公司。

本研究所使用之財務資料欄位如附錄表 3 所示，其為台灣新經濟報資料庫(TAIWAN ECONOMIC JOURNAL DATA BANK)中之財務資料，而為確保資分結果之準確性，將以危機公司與正常公司兩者之資料，皆齊全者之欄位，為首要原則依本研究資料處理之原則，將 77 個欄位刪減成 66 個欄位，刪減欄位包括 1. 用人費用率、2. 呆帳費用率、3. 現金流量允當、4. 財務危機公司、5. 員工人數(人)、6. 財務槓桿度、7. 或有負債比率、8. TCRI 信用評等、9. 營運槓桿度、10. 現金再投資等十個欄位資料，整理後之 66 個資料欄位如附錄表 4 所示。

## 伍、分析結果

### 5.1 樹狀分類法

本研究首先依據 82~91 年之營建財務資料，進行樹狀分類法，其錯誤率為 2.597%，如附錄圖 1 與經刪剪(Pruning)的樹如附錄圖二所示。而根據上述決策樹所選取之變數為 1. 有息負債利率、2. 存貨及應收帳款/淨值，3. 內部保留比率，4. 總資產週轉率(次)，5. 每人配備率，6. 營業費用率、7. 每人營收(千元)與 8. 固定資產週轉率(次)等八項變數，分別定義如附錄表 5，而所產生之規則如下頁所列：

規則 1：有息負債利率 $<3.43$  and 存貨應收帳款/淨值 $<5.86$   
Than 類別=1(純度 85.7%)

規則 2：有息負債利率 $<3.43$  and 存貨及應收帳款/淨值 $\geq 5.86$   
and 總資產週轉率(次) $<0.64$  and 每人配備率 $<31346.75$   
Than 類別=0(純度 98.2%)

規則 3：有息負債利率 $<3.43$  and 存貨及應收帳款/淨值 $\geq 5.86$   
and 總資產週轉率(次) $<0.64$  and 每人配備率 $\geq 31346.75$   
and 固定資產週轉率(次) $=0.275$  and 營業費用率 $<9.87$   
Than 類別=0(純度 100%)

規則 4：有息負債利率 $<3.43$  and 存貨及應收帳款/淨值 $\geq 5.86$   
and 總資產週轉率(次) $<0.64$  and 每人配備率 $\geq 31346.75$   
and 固定資產週轉率(次) $\geq 0.275$  and 營業費用率 $\geq 9.87$   
Than 類別=1(純度 92.3%)

規則 5：有息負債利率 $<3.43$  and 存貨及應收帳款/淨值 $\geq 5.86$



and 總資產週轉率(次) $\geq 0.64$

Than 類別=1(純度 100%)

規則 6：有息負債利率 $\geq 3.43$

and 每人營收(千元) $< 11209.835$

Than 類別=0(純度 100%)

and 內部保留比率 $< 96.25$

規則 7：有息負債利率 $\geq 3.43$

and 每人營收(千元) $\geq 11209.835$

Than 類別=0(純度 84.6%)

and 內部保留比率 $< 96.25$

規則 8：有息負債利率 $\geq 3.43$

and 內部保留比率 $\geq 104.520004$

Than 類別=1(純度 94.7%)

and 內部保留比率 $< 96.25$

規則 9：有息負債利率 $\geq 3.43$

and 內部保留比率 $\geq 104.520004$

Than 類別=1(純度 94.7%)

and 內部保留比率 $> 96.25$

## 5.2 敏感度分析

本研究以類神經決策分類法所進行之敏感度分析，為避免產生無法解釋之研究偏誤，去除 1. 編號、2. 公司、3. 年月與 4. 財務危機與否等非財務資料欄位，其結果如附錄表 5 所示；其分類之精確度為 89.13%，10.73%不正確，0.14%不明，如附錄圖 3 與附錄圖 4 所示。後續將以敏感度指數大於 1 資料欄位(具有底色之欄位)，再進行樹狀分類，期能驗證以原始資料之進行分析，將可獲得較佳之區別能力。

## 5.3 決策樹分類法(篩選變數)

為瞭解專家或經篩選後變數之解釋能力，本研究在此依 5.2 敏感度分析出之指標大於 1 之變數(如附錄表 6 所示)，再進行一次樹狀分類，其錯誤率為 3.759%，如附圖 5 所示，而其樹狀之分類結果如附錄圖 6 所示，挑選之變數整理如附錄表 7。

## 陸、結論與後續研究說明

### 6.1 結論

根據研究顯示，以原始欄位資料所分類之錯誤率為 2.597%，而若以類神經敏感度分析所得之高敏感度(資料敏感度 $> 1$ )欄位資料(自變數)進行分類，其錯誤率反而升高為 3.759%，此研究結果可應證過去研究不論以質性(專家經驗)或定性(因素分析、迴歸分析等統計方法)方式選取資料欄位(自變數)，其僅以已被認定之資料欄位(自變數)進行分類，模型將可能遺漏重要變數(Miss Important Variable)，如本研究即遺失 1. 內部保留比率，2. 營業費用率、3. 每人營收(千元)與 4. 固定資產週轉率(次)等重要變數，而對分類結果之產生偏差(使分類能力降低)。

此外，研究中所篩選出之變數有息負債利率、存貨及應收帳款/淨值、內部保留比率、總資產週轉率(次)、每人配備率、營業費用率、每人營收(千元)與固定資產週轉率(次)等八項變數，可與 Asquith et al (1994)以財務解剖之研究提出利息費用太高、營業績效比同業差與產業衰退等三項因素相對應，詳如附錄表 8 所示。

## 6.2 後續研究說明

### 6.2.1 營建業財務管理

由本研究初步成果可知，營建業財務危機預測模型之建立，應將所有變數都納入考量，再以使用軟體去除資料共線性問題，以避免發生遺漏重要變數(Miss Important Variable)邏輯上之錯誤；因此，未來本研究將循研究目的，將 John Argenti (1976)所提出之『企業生存因素』(Factors Affecting Survival of Business)作為研究背景理論，包括內在因素(Internal Factor)之企業特性 Business Characteristics)與外在因素(External Factor)產業特性(Industrial Characteristics)、經濟環境(Economics Characteristics)等變數，將 TAIWAN ECONOMIC JOURNAL DATA BANK 中所有之相關變數(依營造、建設相關之全部資料欄位)納入後續之研究。而在資料選去上，則不限定任何時間點，以確保能以全部資料詮釋現實狀況，進而發展準確性高之營建業財務危機預測模型。

此外，根據 Stephen A. DeLurgio (1995)對預測的定義：『預測是一個點、一個範圍及機率的估計(forecast should be a point, range, and probability estimate)，預測是對未來值或狀況機率上的估計或敘述，其機率強調了預測不該只是單一值，而是一個區域，而好的規劃與控制要求預測要包括平均值、範圍，和對該範圍的機率估算等項目。』可知過去部份學者未將資料依時間先後，將訓練資料(Training Data)、測試資料(Test Data)與應用資料(Application Data)分開，亦會產生邏輯上錯誤，此已納入後續研究之範疇。

### 6.2.2 營建業成本管理

營建業成本控制(Cost Management)，由 ASCE 之相關研究可知，營建業的成本超支(Cost Overrun)，對產業界有致使工程因現金周轉不靈而導致失敗的隱憂，在官方更有預算執行的績效問題，學術界亦致力於研究成本超支之關鍵因素，此重要性為各方所關切；但營建業過去從低價搶標，雖演變至近已發展出多種的競標方式，以杜絕不當廠商賺取暴利，影響工程品質，但成本不論是多為業主方的政府公家機關，亦或乙方民間營建公司，都是一個『不欲人知』的秘密，而標得後的工程變更案更是許多弊案的管道。

本研究於 91 年 10 月曾發公文向「台灣交通部國道興建工程局(Ministry of Transportation and Communication in Taiwan Area National Expressway Engineering Bureau)」，申請台北第二高速公路工程之相關資料，其雖提供光碟將過去相關資料掃描，但其變更項目不明，並將多項變更金額以總額表示，使資料無法應用於調查個別研究工程變更與追加預算之緣由，且因是施工處凌散各地、工程期間長與人員調動頻繁等因素，使責任無法歸屬，資料更是不可考，使成本控制之研究難以推展，由上述可知『工程資料庫之建立』實為營建業學術研究之重要資料根基。

本研究之調查除「台灣交通部國道興建工程局(Ministry of Transportation and Communication in Taiwan Area National Expressway Engineering Bureau)」外，亦包括『公共工程委員會』，調查發現『公共工程委員會』自 xx 年已逐步進行工程資料庫之建立，但所累積之資料量可能不足以作為資料採擷的樣本數，因此，本研究未來將再深入調查，以研擬相關資料不足時之應對辦法，期能尋求營建成本超支之原因，作為本研究之研究貢獻。

[参考文献]

1. John Argenti, Corporate Collapse : The Causes and Symptoms, McGraw-Hill, (1976).
2. Roy A. Foulke, " Financial ratios Become of Age" , Journal of Accountancy, pp. 209~210, (1937).
3. Mark E. Zmijewski , "Methodological Issue Related to the Issue to the Estimation of Financial Distress Prediction Model" , Journal of Accounting Research, (1984).
4. L. C. Gupta , " Financial Ratio For Monitoring Corporate Sickness - Toward a More Systematic Approach" , Oxford University Press, (1983).
5. Friedman, J. "Data mining and statistics: What is the connection?" , The 29th Symposium on the Interface, Houston, TX, (1997).
6. Hand, D. J., Blunt, G. , Kelly, M. G. and Adams, N. M. , " Data mining for fun and profit." , Statistical Science, Vol.15, NO.2, pp.111-131. (1999)
7. Berry, M. J. and Linoff, G. Data Mining Techniques for Marketing, Sales and Customer Support. Wiley, 1997.
8. Beaver , Willian ,H., Financial Ratios as Predictor of Failure ,Accounting-Selected Studies, (1966).
9. ,Altman,,E.,I. , " Financial Ratios Discriminate ,Analysis and the Prediction of Corporate Bankruptcy" , Journal of Finance, (1968).
10. Cadden,,D.,T., "Neurnal Networks and the Mathematics of Chaos-An Methodologies as Accurate Predictors of Corporate Bankruptcy" , First International Conference on Artificial Intelligence Applications in Wall Street , IEEE Computer Society, (1991).
11. ,Tam K. y. and Kiang M. Y., " Applications of Neural Networks: The Case of Bank Failure Prediction" , Journal of Science Management (1992).
12. Coats P. k. and F. L.Fantlin, "Distress Patterns Using a Natural Network Tool" ,Financial Management Recognizing (1993).
13. Altman E. I. G. Marco and F. Varetto , " Bankruptcy and Finance Corporate Distress Diagnosis: Comparisions Using Liner Discriminate Analysis and Neural Network." Journal of Finance(1994).
14. Wilson R. L. and R. Sharda , "Bnakruptcy Prediction Using Neural Network " , Journal of Decission Support System , (1994).
15. Wu Xiaodan and Andrew Flitman " Evaluating Corporate Failure Risk with a New Intelligent Intellingent Process Approach" IEEE International Conference on Intelligence Process System vol.2 (1997).
16. Alam P. D. Booth. K. Lee T. "The Use of Fuzzy Clustering Algorithm and Self-Organizing Neural Networks for Identifying Potentially Failing Banks: an Experimental Study" Thordarson Expert Systems with Applications, (2000). Stephen A. DeLurgio, Forecasting Principles and applications,

McGraw-Hill, (1998). Jiawei Han, Micheline Kamber, Data Mining: Concept and Techniques, Wiley, (2000). Michael J. A. Berry, Gordon S. Linoff, Data Mining: The Art and Science of Customer Relationship Management, Wiley, (2000) Boris Kovalerchuk, Evgnii Vityaev, Data Mining in Finance: Advance in Relational and Hybrid Methods., Kluwer Academic, (2000). 陳肇榮, 『財務比率預測企業財務危機之實證研究博士論文』, 國立政治大學企業管理研究所, 台北(1983)。陳明賢, 『財務危機預測之計量分析模型碩士論文』, 國立台灣大學商業管理研究所, 台北(1996)。

## 計畫成果自評

本計畫依照原研究計畫內容執行，過程中遭遇困難包括：資料蒐集不易，業界之資料電腦化程度不佳致使許多資料難以分析，及需從多次錯誤中學習較佳之 mining 方式等。然本研究仍克服上述困難並達成此二年期計畫之首年目標。Data Mining 於營建管理之應用上極其有限，尤其於財務之應用更由於資料蒐集及研究設計問題，使得此類研究更為不易。故本研究之學術價值及實務應用價值應皆屬極高。本研究已栽培一碩士班學生進行其碩士論文研究並已有初步成果，此成果已整理成研討會報告並投稿中。預期經過第二年之研究，其成果可於國際學術期刊中發表。

# 附 錄

表 1. 財務危機定義之彙整表

年份	作者	財務危機之定義								
2003	台灣經濟新報資料庫(TAIWAN ECONOMIC JOURNAL DATA BANK) 信用風險評估之 9 種事件	公司宣佈 倒閉破產	公司聲 明重整	公司 跳票	向外界紓 困求援	公司被外 界接管	CPA 意見，對其繼續 經營有疑慮	公司淨 值為負	公司轉列全額交 割股或下市	公司財務 吃緊停工
1966	Beaver	✓		✓				✓		
1968	Altman			✓				✓		
1972	Deakin	✓		✓				✓		
1978	Dun	✓	✓	✓	✓	✓		✓		✓
1986	陳明賢		✓	✓					✓	
1990	張正忠	✓	✓	✓				✓	✓	✓
1990	Wruck			✓						
1991	Gertner and Scharfftein			✓				✓		
1993	John			✓						

表 2 正常營建公司

2501 國建	2526 大陸	2528 皇普	2548 華固	9933 中鼎	2540 林三號
2504 國產	2527 宏璟	2516 新建	2514 龍邦	2524 京城	2542 宏巨
2505 國揚	2520 冠德	2530 華建	2511 太子	2523 德寶	2543 皇昌
2506 太設	2521 宏總	2533 昱成	2509 全坤	2515 中工	2544 昌益光電
2547 日勝生	2536 宏普	2534 宏盛	2545 皇翔	2546 根基	2535 達欣工程
9945 潤泰新					



表 3 營建財務危機公司(9 家)

編號	公司簡稱	公司類別	發生危機日	危機類型(危機日期)
1	2503 復木	建設	1973/4/24	重整(1984/04/24)
2	2512 寶建	建設	1990/5/7	跳票擠兌(2001/05/07)
3	2517 長谷	建設	1988/8/7	重整(1999/08/07)
4	2522 啟阜	營造	1987/12/2	跳票擠兌(1998/12/02)
5	2525 寶祥	建設	1991/9/6	全額下市(2002/09/06)
6	2529 仁翔	建設	1987/12/29	跳票擠兌(1998/12/29)
7	2537 春池	建設	1989/12/15	紓困求援(2000/12/15)
8	2538 基泰	建設	1991/2/11	全額下市(2002/02/11)
9	2539 櫻建	建設	1988/1/26	紓困求援(1999/01/26)

表 4 資料欄位名稱(共 66 個欄位)

編號	內部保留比率	資產報酬率(稅前息前)
資產報酬率(稅後息前)	總負債/總淨值	營業毛利率
營業利益率	利息保障倍數	常續性利益率(稅後)
營業費用率	應收帳款週轉率(次)	每股盈餘(元)
每股淨值(B)(元)	淨值週轉率(次)	每股現金流量(元)
每股營業額(元)	每人配備率	營業毛利成長率
營業利益成長率	存貨週轉率(次)	總資產成長率
淨值成長率	總資產週轉率(次)	常續性利益成長稅後
流動比率	固定資產週轉率(次)	負債比率
淨值/總資產	每人營業利益(千元)	營業利益/實收資本
稅前純益/實收資本	應付帳款收現天數(天)	應收帳款收現天數(天)
公司	年月	財務危機公司
資產報酬率(息前折舊前)	淨值報酬率(稅後)	淨值報酬率(常續利益)
稅前淨利率	稅後淨利率	業外收支率
現金流量比率	有息負債利率	有效稅率
每股淨值(A)(元)	每股淨值(C)(元)	常續性 EPS
每股營業利益(元)	每股稅前淨利(元)	營收成長率
稅前淨利成長率	稅後淨利成長率	經常淨利成長率
固定資產成長率	總資產報酬成長率	存貨應收帳款/淨值
速動比率	利息支出率	平均銷售天數(天)
長期資金適合率	借款依存度適合率(A)	每人營收(千元)
淨營業週期(天)	當季季底 PBR	當季季底 PER

表 5 自變數(5.1 原始變數)之定義

自變數	定義
1. 有息負債利率	(利息支出 + 資本化利息) / 平均長短期借款 *100
2. 存貨及應收帳款/淨值	(存貨 + 應收帳款及票據) / 淨值 *100
3. 內部保留比率	分配後盈餘 / 稅後淨利 *100
4. 總資產週轉率(次)	營收 / 平均總資產
5. 每人配備率	固定資產 / 員工人數
6. 營業費用率	營業費用 / 營業收入淨額 *100
7. 每人營收(千元)	營收 / 員工人數
8. 固定資產週轉率(次)	營業收入淨額 / 平均固定資產

備註：『有息負債利率』之平均長短期借款為去年同期(不含去年同期)至本期止各季平均長短期借款。若該比率為 9999.00 表示此計算式分母過小致有息負債利率過高(超過 9999.99), 則以 9999.00 代入。

表 6 類神經欄位敏感度分析

編號	資料欄位	敏感度指標	編號	資料欄位	敏感度指標	編號	資料欄位	敏感度指標
1	存貨週轉率(次)	7.8	22	總資產成長率	1.4	43	利息保障倍數	0.5
2	淨值/總資產	6.2	23	稅前淨利成長率	1.4	44	營業費用率	0.5
3	有息負債利率	6	24	每股營業額(元)	1.3	45	每人營收(千元)	0.4
4	應收帳款週轉率(次)	5.6	25	利息支出率	1.3	46	每股淨值(B)(元)	0.4
5	負債比率	5.2	26	每人營業利益(千元)	1.2	47	淨值報酬率(常續利益)	0.4
6	總資產週轉率(次)	4.2	27	稅後淨利成長率	1	48	平均銷售天數(天)	0.4
7	速動比率	4	28	經常淨利成長率	1	49	淨營業週期(天)	0.3
8	常續性 EPS	3.4	29	現金流量比率	0.9	50	稅前純益/實收資本	0.3
9	流動比率	3.4	30	淨值成長率	0.9	51	內部保留比率	0.3
10	每人配備率	3.3	31	當季季底 PBR	0.8	52	固定資產週轉率(次)	0.2
11	總負債/總淨值	3.3	32	營業毛利率	0.8	53	稅前淨利率	0.2
12	應收帳款收現天數(天)	3.3	33	總資產報酬成長率	0.7	54	稅後淨利率	0.2
13	存貨應收帳款/淨值	2.5	34	營業利益成長率	0.7	55	業外收支率	0.2
14	長期資金適合率	2.5	35	營業利益率	0.7	56	資產報酬率(稅前息前)	0.2
15	借款依存度適合率(A)	2.1	36	常續性利益成長稅後	0.7	57	資產報酬率(稅後息前)	0.2
16	固定資產成長率	2	37	淨值報酬率(稅後)	0.6	58	營業利益/實收資本	0.2
17	每股現金流量(元)	2	38	每股淨值(A)(元)	0.6	59	每股盈餘(元)	0.2
18	淨值週轉率(次)	2	39	當季季底 PER	0.6	60	每股淨值(C)(元)	0.2
19	應付帳款收現天數(天)	1.9	40	營業毛利成長率	0.6	61	每股稅前淨利(元)	0.2
20	有效稅率	1.7	41	資產報酬率(息前折舊前)	0.6	62	每股營業利益(元)	0.2
21	常續性利益率(稅後)	1.5	42	營收成長率	0.5			

表 7 自變數(5.3 挑選後變數)之定義

自變數	說明
有息負債利率	$(\text{利息支出} + \text{資本化利息}) / \text{平均長短期借款} * 100$
常續性利益率(稅後)	$\text{常續性利益(稅後)} * 100\% / \text{營收淨額}$
應收帳款週轉率(次)	$\text{營業收入淨額} / \text{平均(應收帳款及票據} + \text{應收票據貼現)}$
存貨及應收帳款/淨值	$(\text{存貨} + \text{應收帳款及票據}) / \text{淨值} * 100$
總資產成長率	$\text{總資產增減額} / \text{ABS(去年同期總資產)}$
每人營業利益(千元)	$\text{營業利益(千元)} / \text{員工人數}$

備註：『有息負債利率』之平均長短期借款為去年同期(不含去年同期)至本期止各季平均長短期借款。若該比率為 9999.00 表示此計算式分母過小致有息負債利率過高(超過 9999.99), 則以 9999.00 代入。

表 8 Asquith et al (1994)與本研究原始變數之對應表

Asquith et al (1994)變數	對應本研究原始變數	說明
利息費用太高	有息負債利率	$(\text{利息支出} + \text{資本化利息}) / \text{平均長短期借款} * 100$
營業績效比同業差	內部保留比率	$\text{分配後盈餘} / \text{稅後淨利} * 100$
	固定資產週轉率(次)	$\text{營業收入淨額} / \text{平均固定資產}$
	營業費用率	$\text{營業費用} / \text{營業收入淨額} * 100$
	總資產週轉率(次)	$\text{營收} / \text{平均總資產}$
	存貨及應收帳款/淨值	$(\text{存貨} + \text{應收帳款及票據}) / \text{淨值} * 100$
產業衰退	每人配備率	$\text{固定資產} / \text{員工人數}$
	每人營收(千元)	$\text{營收} / \text{員工人數}$

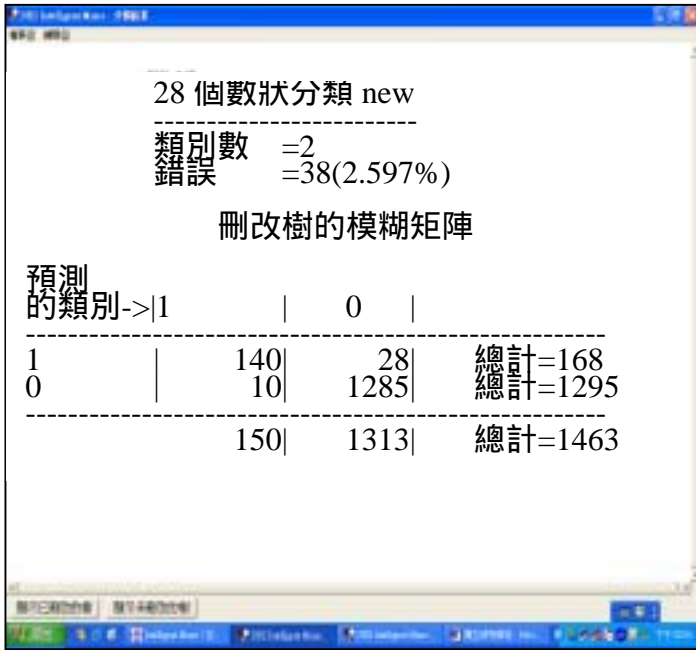


圖 1. 分析結果圖

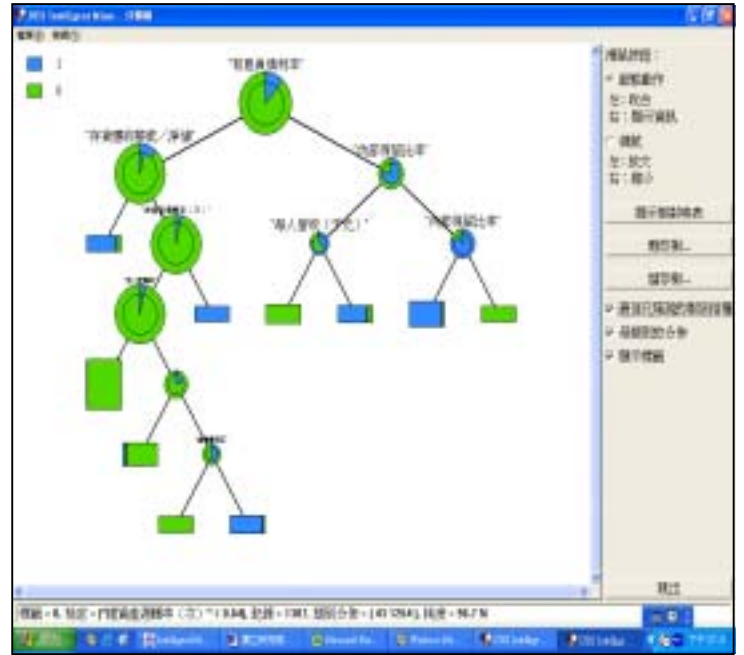


圖 2 已刪剪(Pruning)的樹(樹的深度=7)

備註：a. 類別 1：財務危機公司；類別 0：正常公司。

b. 未經刪減之樹狀分類(深度為 13)因可能會過度解釋(Over Fitting)，因此多以已刪剪(Pruning)的樹(樹的深度=7)進行研究解釋。

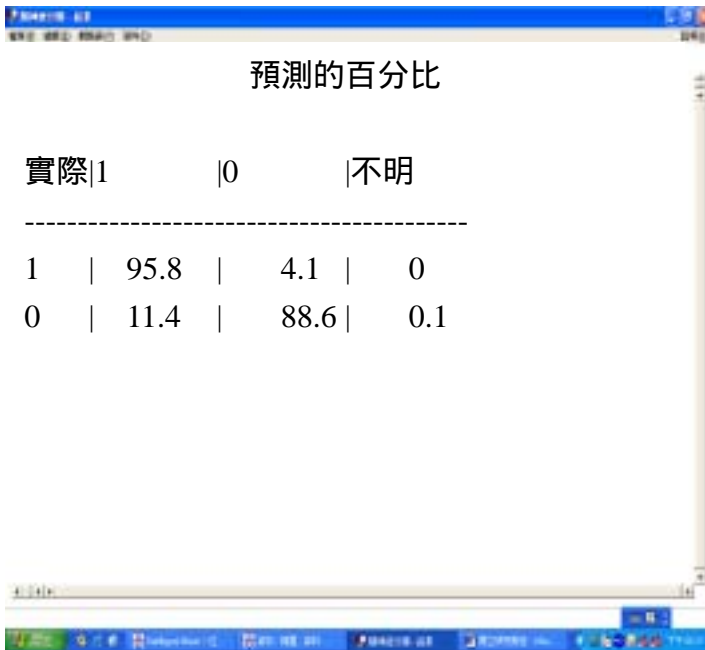


圖 3 預測的百分比—文字表示

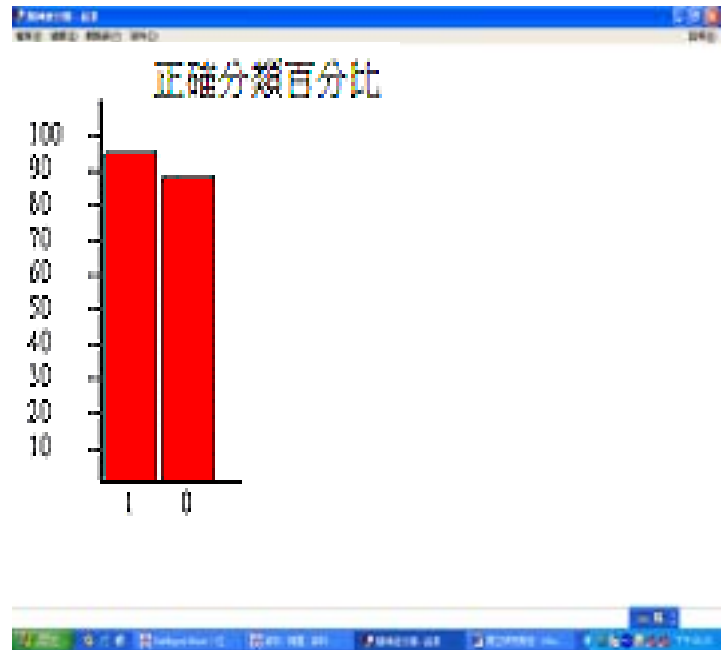


圖 4 預測的百分比—長條圖表示

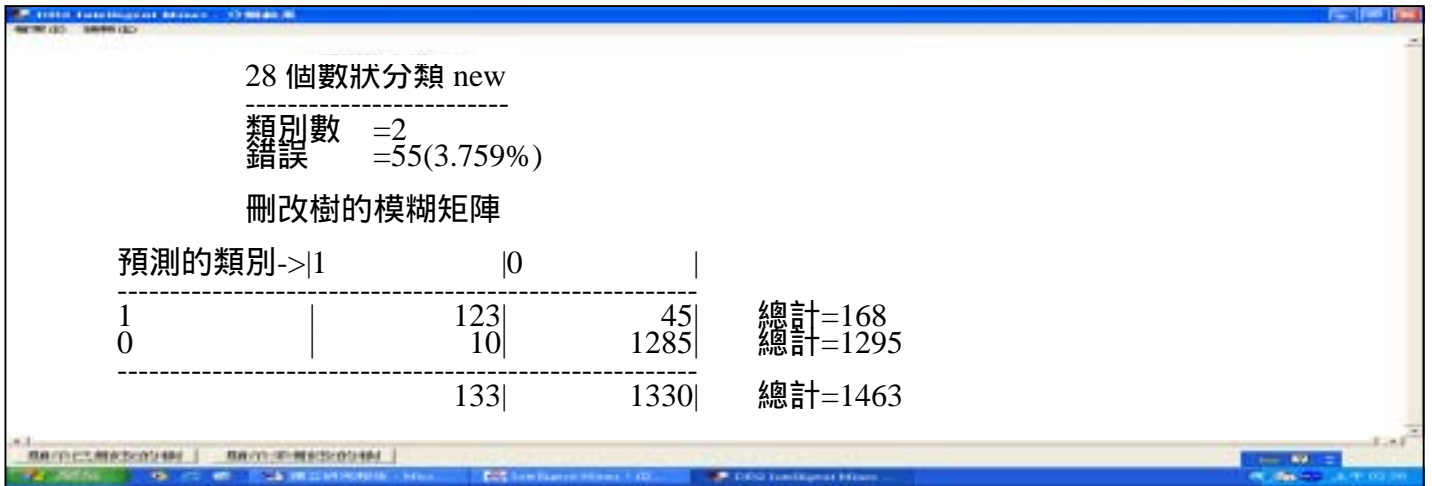


圖 5 預測的原始值—文字表示

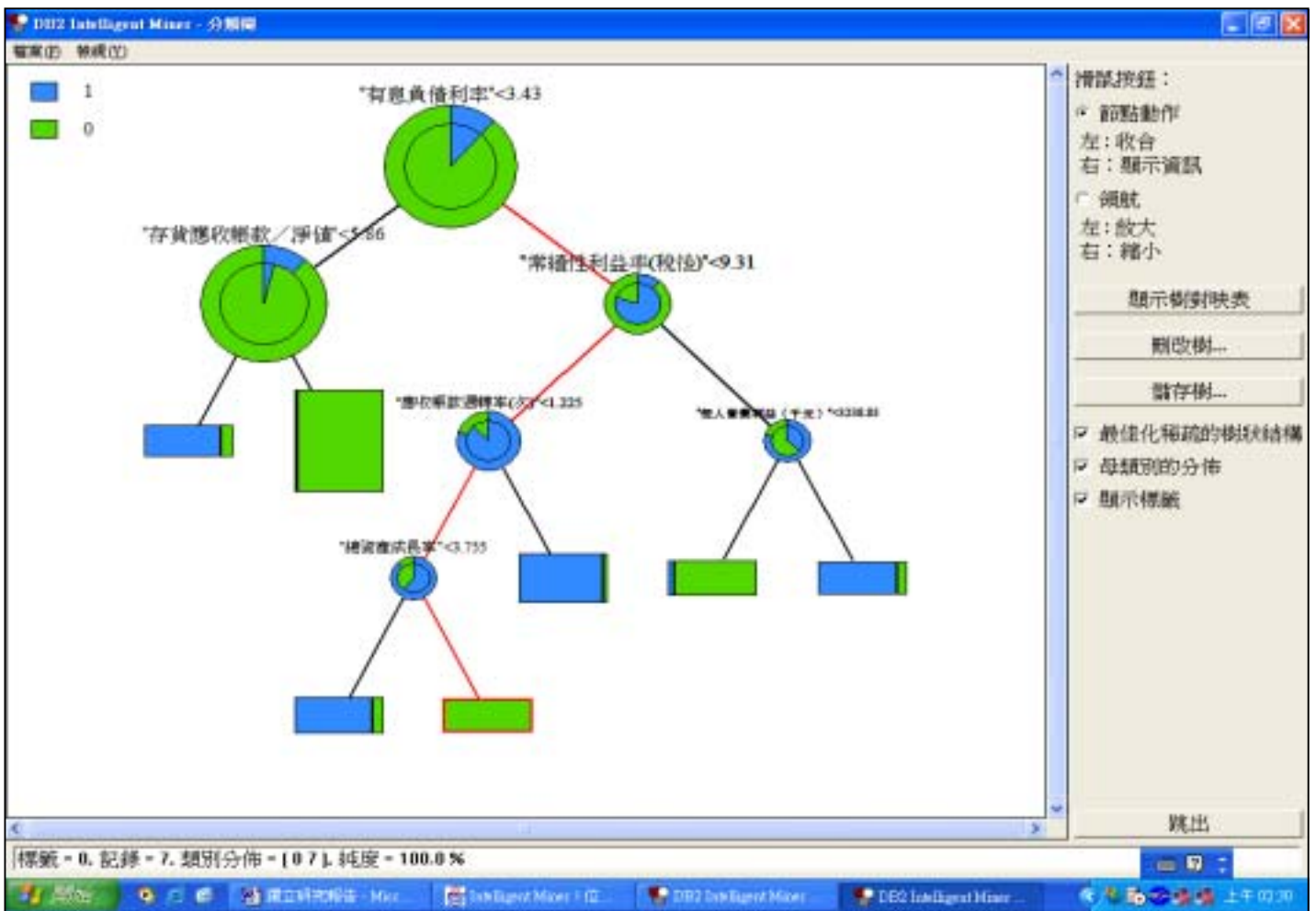


圖 6 樹狀之分類結果(篩選變數)