



嶺東科技大學
LING TUNG UNIVERSITY

資訊管理系

以決策樹和線性判 別分析探索晨星評 鑑基金之等級

指導教授： 萬絢 張士勳 教授

組員名單： 陳韋勝 A18C158
溫智硯 A18C164
謝均陽 A18C095
周筱惠 A18C140

中華民國 105 年 6 月

嶺東科技大學 資訊管理系

以決策樹和線性判別分析探索晨
星評鑑基金之等級

中華民國一零五年六月



嶺東科技大學
LING TUNG UNIVERSITY

資訊管理系專題口試委員審定書

以決策樹和線性判別分析探索晨星
評鑑基金之等級

指導教授： 萬絢 張士勳 教授

組員名單： 陳韋勝 A18C158

溫智硯 A18C164

謝均陽 A18C095

周筱惠 A18C140

指導教授： _____

口試委員： _____

中華民國 105 年 5 月 11 日

誌 謝

本論文得以順利完成，有許多人需要感謝。但首先要誠摯的感謝我們的指導教授萬絢教授以及張士勳教授，在研究專題的過程中，兩位教授總是在忙碌的課程之餘還花費了許多心思於我們的論文研究上，並悉心的教導使我們得以一窺基金的深奧，不時的討論並指點我們正確的方向，使我們在這其中獲益匪淺，也學習到團隊合作的精神。另外感謝嶺東科技大學資訊管理系陳明華教授、陳建興教授擔任口試委員，因為教授們的寶貴建議，使得專題論文的內容能更加完善，在此對於教授們的諄諄教誨，致上最深的謝意。

陳韋勝、溫智硯、謝均陽、周筱惠

謹誌

中華民國一〇五年五月嶺東科技大學

摘要

共同基金是現代投資人投資組合的重要工具之一，對於基金投資人而言如何在眾多基金當中，選出適合的投資標的，即成為重要的議題。基金的評級資料即為許多投資人參考的資訊，晨星的星等評級具有相當的知名度，且基金公司亦多提供其獲得的星等評級資料供投資人參考。一般投資大眾信賴晨星評級，往往以其結果進行投資決策。然而晨星評級是否合理有效？是否被人為控制？因此需以客觀資料探勘的方式進行評估。本研究以晨星金磚四國(俄羅斯、巴西、印度、中國)分組與亞洲(不包括日本)分組為例。另外，晨星評級只對基金本身進行評級，若基金經理人異動可能對基金之表現產生影響，所以晨星評級並不能完全揭露基金目前的狀態，所以我們利用 2012 年 5 月 15 號至 2015 年 5 月 15 號之基金明細資料，並運用決策樹方法使用 WEKA 軟體及線性判別分析(Linear Discriminant Analysis, LDA)進行預測分類，(亞洲(不包括日本)為訓練組，金磚四國為測試組)，並使用大盤法加以比較，以此檢驗晨星評級是否合理有效。

關鍵字：共同基金、基金評級、線性判別分析、決策樹

目 錄

摘 要.....	I
目 錄.....	II
圖目錄.....	IV
表目錄.....	V
第一章 緒論.....	1
1.1 研究背景與動機.....	1
1.2 研究目的.....	2
1.3 研究設計.....	3
1.4 研究流程.....	4
第二章 文獻回顧與探討.....	5
2.1 相關資料.....	5
2.2 專有名詞.....	7
第三章 研究方法.....	11
3.1 研究架構.....	11
3.2 決策樹(Decision Tree).....	12
3.4 線性判別分析(Linear Discriminant Analysis, LDA)..	15

第四章 實證研究	17
4.1 基金篩選.....	17
4.2 訓練資料及測試資料.....	18
4.3 晨星評級.....	32
4.3.1 決策樹分析結果(案例 a).....	32
4.3.2 LDA 分析結果(案例 b).....	39
4.3.3 兩方法誤差矩陣表(案例 c).....	42
4.4 大盤法.....	43
4.4.1 決策樹分析結果(案例 a).....	43
4.4.2 LDA 分析結果(案例 b).....	50
4.4.3 兩方法誤差矩陣表(案例 c).....	53
第五章 結論	54
參考文獻	55
附錄	57
附錄 A WEKA.....	57
附錄 B 誤差矩陣說明.....	64

圖目錄

圖 1.1 本研究之研究流程圖	4
圖 3.1 研究架構圖	11
圖 3.2 決策樹範例圖	14
圖 4.1 使用訓練資料產生決策樹圖	32
圖 4.2 導入決策樹之訓練資料交集聯集圖	35
圖 4.3 測試資料導入訓練資料之決策樹交集聯集圖	38
圖 4.4 測試資料導入訓練資料使用 LDA 產生之分類交集聯集圖	41
圖 4.5 使用訓練資料產生決策樹圖	43
圖 4.6 導入決策樹之訓練資料交集聯集圖	46
圖 4.7 測試資料導入訓練資料之決策樹交集聯集圖	49
圖 4.8 測試資料導入訓練資料使用 LDA 產生之分類交集聯集圖	52
圖 A.1 進入 WEKA 之介面	58
圖 A.2 Explorer 介面	58
圖 A.3 Knowledge Flow 介面	60
圖 A.4 KF	61
圖 A.5 探勘結果	62
圖 A.6 BayesNet 分類法	63
圖 A.7 決策樹成品圖	63

表目錄

表 2.1 晨星 Morningstar 星號評級的意義	7
表 4.1 訓練資料明細	19
表 4.2 測試資料明細	26
表 4.3 導入決策樹之訓練資料結果	33
表 4.4 導入決策樹之訓練資料分類結果表	35
表 4.5 測試資料導入訓練資料之決策樹結果	36
表 4.6 測試資料導入訓練資料之決策樹結果表	38
表 4.7 測試資料導入訓練資料使用 LDA 產生之分類結果	39
表 4.8 測試資料導入訓練資料使用 LDA 產生之分類結果表	41
表 4.9 決策樹誤差矩陣表	42
表 4.10 線性判別分析誤差矩陣表	42
表 4.11 導入決策樹之訓練資料結果	44
表 4.12 導入決策樹之訓練資料分類結果表	46
表 4.13 測試資料導入訓練資料之決策樹結果	47
表 4.14 測試資料導入訓練資料之決策樹結果表	49
表 4.15 測試資料導入訓練資料使用 LDA 產生之分類器結果	50
表 4.16 測試資料導入訓練資料使用 LDA 產生之分類結果表	52
表 4.17 決策樹誤差矩陣表	53

表 4.18 線性判別分析誤差矩陣表53

表 4.19 綜合結果表53

第一章 緒論

1.1 研究背景與動機

現代人有很多的投資工具除了股票、基金、期貨、選擇權和定存等等，投資工具多得令人眼花撩亂，面對這麼多樣的投資工具，到底哪種投資工具比較好，賺錢比較容易呢？[2]

所謂共同基金即在經理與保管分開原則下，由投信公司公開募集小額投資的資金，將資金交由保管機構管理，並經由專家管理操作，比個人投資更能掌握全球經濟與市場脈動，以及投資獲利的機會，由於基金具備專業知識的經理人操盤、分散風險、資金流動性高、成本較低等特色，對不諳操作策略的投資人而言，提供了一種較直接投資更為有效的投資的工具。

基金評級以科學的方法判斷基金的特點和績效表現，國際知名的評效機構包括晨星機構、理柏公司的 leader 評級系統、標準普爾星號評等、Value Line 評級、在國內部份則有台灣經濟新報基金評比系統，這些評級系統普遍的受到基金理財網站的使用，而決定基金績效因素例如大盤未來走勢、投資組合風險、基金經理團隊之優劣，而研究對象，探討如何篩選基金及探討其績效表現。

一般而言，多數基金投資者信任信評機構所提供的基金好壞指標，如晨星(morning star)所提供的五星評等，鮮少有人探討過評等的好壞，因此，本研究以線性判別分析和決策樹分別檢驗亞洲(不包括日本)和金磚四國之投資評等與績效的關聯性，並利用大盤法與五星評等加以比較。

從國內基金選擇到國外基金的引入，使得基金琳琅滿目，面對眾多的投資標的，投資者無法容易決定投資哪一基金，國內基金雖然高報酬具有相當吸引力，但伴隨之高風險性同時也會讓投資者卻步，但其境外基金，相對於國內基金投資者能獲取資訊更有所限制，投資者僅能於報章雜誌與網路獲取國外市場之資訊，其資訊獲取較為困難，而實務上最常使用基金評級也就是績效評估，基金之績效反應基金過去的歷史表現，雖然基金績效無法對未來保證，但以投資人的角度來看，若過去基金表現良好，則其未來表現良好的機率相當高，因此對於投資者來說基金是否具有績效持續性為其重要參考指標，而在基金績效持續性的研究領域中，國外與國內學者並無統一之定論，而且國內與國外學者甚少研究國內與境外基金的績效持續性議題，因此引發了研究基金之績效持續性的動機。

1.2 研究目的

隨著股市規模持續擴大，投資人理財觀念也逐漸成熟，共同基金快速成長，成為國內僅次於股票的投資工具，共同基金標榜著小額投資、風險分散及專家代為經營管理的優點，投資人希望所選擇的基金在經理人專業領域及操作下，擊敗市場擊敗同業獲得更高的報酬。

基金績效評比是針對基金過去的績效表現進行考評，以呈現基金在過去某段期間的表現情形，若基金被賦予最高等級評價，僅能說明過去績效良好，但未來仍是未知數，然而基金績效評比在某些報章雜誌、媒體以及網站部落格的推崇與宣傳下，以被渲染成基金市場中的北極星，產生「只要選擇高評價基金，即可獲得良好報酬」的觀念，但這卻是十分危險的觀念，畢竟基金績效評比並未對基金未來績效予以保證；而大盤法是指基金在該組別中的表現，以淨值高於或低於該組別平均為分界點。

基金是否具有績效持續性對於投資者來說是一重要資訊，加上區域型基金為基金成長型基金，投資者所預期之高報酬同時伴隨高風險，因此本研究目的在於欲提供投資區域型基金之投資者相關參考資訊，降低其投資不確定性。

投資人面對如此多的基金如何挑選優質且適合自己的基金成為重要課題，本研究以亞洲(不包括日本)基金及金磚四國基金為對象，探討如何篩選基金及探討不同區域基金之績效表現，其研究目的：

一、區域型基金是否具有績效持續性，過去的歷史績效是否可預測未來之績效。

二、使用決策樹和線性判別分析比較大盤法和晨星評級。

三、找出哪些屬性比較重要，門檻(分割點)為何？

1.3 研究設計

一、研究期間

本研究依據 2015 年晨星，亞洲(不包括日本)、金磚四國(俄羅斯、巴西、印度、中國)，分組方式分類的組別中，分成訓練組與測試組進行分析，為避免同一基金不同級別重複計算，只選擇「2012-2015 年」納入研究，若基金建立於 2015 年 5 月 15 號不滿三年之基金不列入研究範圍。[3]

二、研究範圍及樣本取樣

本研究主要是探討亞洲(不包括日本)、金磚四國(俄羅斯、巴西、印度、中國)，所選取樣本來自晨星為研究樣本。

三、研究限制

1. 本研究因考慮基金之中長期績效，樣本期間長達三年，為了資料分析完整性，故在 2015 年 5 月 15 號以前不滿三年成立基金或樣本期間內下市基金，不予納入樣本基金。
2. 本研究不考慮基金交易成本。
3. 以 2012 年 5 月 15 號至 2015 年 5 月 15 號為研究期間

四、研究設計

在研究設計中，本研究使用基金的歷史資料作為分析的屬性，以決策樹法(Decision tree method, D.T.)及線性判別分析(Linear Discriminant Analysis, LDA)進行判別，主要區隔成四個案例分析。

- (a) 將訓練組資料導入 WEKA 產生決策樹
- (b) 以 LDA 對訓練組資料進行分析
- (c) 將測試組資料導入 case(a)及 case(b)進行判別
- (d) 測試結果綜合討論

以上三種組合有助於了解基金篩選、探討晨星評比的結果，亦可進行對區域型基金之了解，同時也可以對基金投資有更深一層的認識。

1.4 研究流程

本研究共分成五大章節，各章內容如下：

第壹章 緒論：說明本研究之研究背景、研究目的、研究設計與研究對象範圍，並對本文有架構做概括地簡介。

第貳章 參考文獻：亞洲(不包括日本)、金磚四國區域型基金、決策樹、晨星評級、大盤法、LDA分析等相關文獻及實證整理作為本研究之參考依據。

第參章 研究方法：介紹本研究之研究架構、使用方法以及探討各項蒐集資料。

第肆章 實證研究：根據本研究架構，將研究資料導入研究方法，驗證晨星評級。

第伍章 結論：將本研究之實證結果彙整，並提出未來研究發展之建議及方向，以供後續研究及投資人參考。

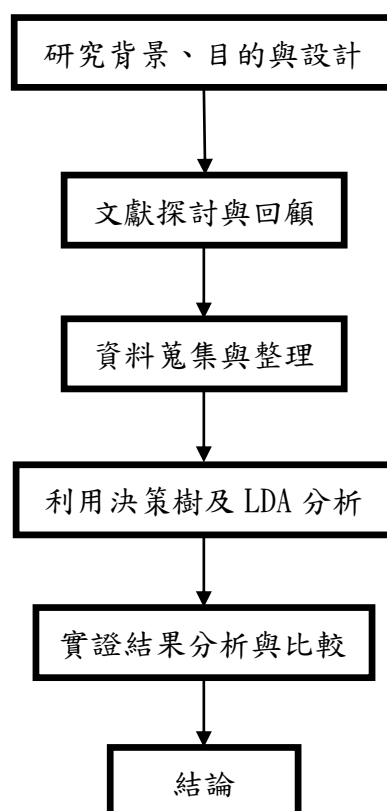


圖1.1 本研究之研究流程圖

第二章 文獻回顧與探討

2.1 相關資料

研究基金績效所參考的資料有許多種，但本研究探討「晨星評級」是否合理有效的相關資料方法，主要分為三種相關研究組合之類別：晨星評級、共同基金、決策樹、LDA 分析、大盤法。

梁碩修(2013) [2]，運用Blake and Morey (2000)分析架構，建立一個以虛擬變數為主的迴歸模型來檢定晨星的星號評級資訊是否對基金未來的投資績效有顯著的預測能力。本研究實證結果顯示：僅有在少數基金組別，高評級基金之未來一個月績效有顯著高於低評級基金；相對多數之基金組別則顯示評級資訊對基金未來三個月、六個月與十二個月的報酬率不具有預測能力。

簡好倫(2007) [4]，指出決定共同基金基效如大盤未來走勢、投資組合風險、基金經理團隊之優劣，而基金績效較高的投資組合，較易獲得投資人青睞。

林聲宇(2005) [5]，以應用決策樹與濾嘴法則於股票投資。本研究提出一結合濾嘴法則與決策樹的股票交易方法，首先利獮濾嘴法則產生交易資料，再運用決策樹C4.5演算法對這些交易資料進行分類，找出高報酬群體。與之前相關研究比較，本研究在分類時多加考慮了代表謝來性資訊的影響。透過對台灣以及NASDAQ市場的實證，顯示出本研究方法確實優於濾嘴法則與相關研究方法。

曾永富、黃嫻嫻、黃玄育(2014)[6]，以大盤法找出光電類股當月收盤價以及大盤當月收盤價計算出收盤價的漲跌幅，並比較其值計算出決策值。利用線性判別分析(Linear Discriminant Analysis, LDA)以及粒子群優化演算法(Particle Swarm Pptimization, PSO)方法計算出光電類股之平均報酬率，比較優化後以及無優化後的差別。

王秋萍(2009)[7]區域型新興市場股票基金整體樣本不具一個月(一季、半年)之績效持續性，僅具一年績效持續性。分區探討績效持續性下，新興亞洲與新興歐洲不具一個月(一季、半年、一年)績效持續性，而拉丁美洲僅具一年績效持續性。拉丁美洲基金一個月績效波動最大；新興歐洲基金次之；新興亞洲基金最小。

蔡毓翔(2003)[8]，將國內股票型基金分三年期、五年期、七年期以淨值報酬率 Treynor 指標、Sharpe 指標、Jense 指標進行績效衡量結果發現，股票型共同基

金的績效表現，在中長期無法顯著地優於大盤指數。

謝文璋(2012)[11]，以 2006 年 12 月到 07 年 12 月之各基金評比資料(晨星基金評比、台灣經濟新報基金評比、中華民國證券投資信託暨顧問商業同業公會基金評比)篩選出基金、建立投資組合，另已根據寬鬆 5050 法則建立一個以成本觀點所猜選之基金投資組合，比較各投資組合之未來一年、三年與五年之報酬率表現差異情形。

張有若(2002)[12]，利用風險值(VaR)取代標準差，並加入標竿市場的相對風險值，改善Sharpe Ratio在報酬成非常態分配時的偏誤，提高各類指標對風險的敏感度，應用在共同基金整體績效評估中。以全球平衡型基金、債券型基金、貨幣型基金及以上三類所組成的基金群組，所做的常態分配檢定，發現並非所有基金皆呈常態分配。

曾麗雪(2014)[13]，採用 Sharpe Ratio、Treyner Ratio、Jensen Ratio、以及 Return to Risk Ratio 樣本基金績效得出，在較長的觀察期間(二年與三年期)，各指標所衡量的績效結果較具一致性；在樣本期間，所有基金的 β 值均小於 1，顯示基金的系統風險比市場小；無論投資亞洲市場或全球市場基金與指數股票型基金的績效均有顯著差異；指數股票型基金係採被動式管理追求指數報酬，其績效優於全球市場及亞洲市場基金。

劉玟君(2006)[14]，指出基金排名之績效表現確實具有持續性，由相關性檢定、績效二分法與Cox-Stuart 檢定等方法評估之後，可知在短期時，基金前後期之排名確實具有持續性，故投資人可據此做為買賣基金之參考依據。基金排名持續性的現象具有差異性，台灣加權指數之波動，檢驗基金排名持續性的現象，發現在空頭市場時，基金排名持續性較不顯著，反之；在多頭行情時，基金排名持續性的現象為顯著，可知投資人可在利多行情時，利用基金排名進行投資買賣，可獲得較佳的績效。

陳秉洋(2011)[15]，將台灣50指數型基金(代號：0050)之歷史交易資訊換算為多項技術指標，再運用多變量皆析及類神經網路進行整合。使用因素分析從多項指標中萃取出具有高解釋能力之共同因素，配合集群分析及區別分析對於價格所屬趨勢建立鑑別函數。最後再搭配融資融券餘額、三大法人買賣超及期現貨基差之變化，整合市場各面向之資訊，提供本研究之買賣決策。結果顯示使用本研究之方法套用在一叢空頭行情中，能擷取出多個波段進行放空回補之動作而得到較高之投資報酬率。

2.2 專有名詞

本研究使用之屬性資料在以下進行說明。

晨星評級：為晨星公司自行訂定的標準，在比較基金與同組別的歷史報酬與風險後，所做出的綜合評判；從 1 顆星到 5 顆星代表基金由低至高的等級，愈多顆星代表過去表現愈好。同時，必須是該基金的績效資料多於連續 36 個月，以及同組別基金多於 5 檔，才會獲得評級。[9]

表 2.1 晨星 Morningstar 星號評級的意義

★★★★★ 前10%的基金
★★★★ 前10%-32.5%的基金
★★★ 前 32.5%-67.5%的基金
★★ 排名67.5%-90%的基金
★ 最後 10%的基金

總開支比(%為單位)：基金的總費用占總資產的比率，又稱為費用率。基金有管理成本與人事成本，必須先在總資產當中扣除這些費用成本，才是所有基金持有者的淨資產。費用並不會直接向投資者收取，而是從每日淨值中扣除，投資者不會感受到，所以很容易忽略。

最新價格：目前基金的淨值。

基金表現：該檔基金至 2015 年 5 月 15 號在該組別的績效表現，為大盤法決策值之來源依據，若績效高於該組別之平均則為 1，反之為 0(若紅線高於橘線則為 1，反之為 0)。

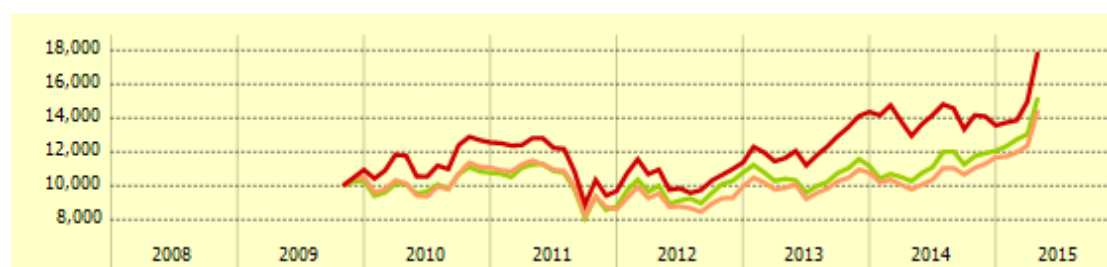


圖 2.1 GAM Star 中華股票基金-A USD 之基金淨值走勢圖

- 基金 GAM Star 中華股票基金-A USD
- 基金組別: 中國股票
- 指數: MSCI China NR USD

基金表現(三年)：該檔基金從 2012 年 5 月 15 號至 2015 年 5 月 15 號在該組別的績效表現，若績效高於該組別之平均則為 1，反之為 0。

總回報率(天)(%為單位)：每天投資而應返回的價值，企業從一項投資性商業活動的投資中得到的經濟回報。

總回報率(三年)(%為單位)：三年投資而應返回的價值平均，企業從一項投資性商業活動的投資中得到的經濟回報。

市盈率：投資組合的市盈率，為組合內各股票的市盈率的加權平均數。市盈率為公司股價與其過去十二個月每股盈利的比率。市盈率低，表示投資者對公司未來盈利增長信心不大；不過，投資者也可將此視為股份被遺忘或股價偏低，即股價有潛力上升。

市帳率：投資組合的市價帳面值率，為組合內各股票的市價帳面值率的加權平均數。市價帳面值率比較公司的股價及其帳面值。公司帳面值為扣除負債後的公司資產值。換言之，公司帳面值亦相等於公司結業時，在收回所有被拖欠項目和扣除所有債務後，可發還予股東的資產。

市銷率：投資組合的市銷率，為組合內各股票的市價對銷售率的加權平均數。市價對銷售率比較公司的股價及銷售額。較低的市帳率顯示投資者在該項投資所付的成本愈低。

股價對現金流：投資組合的市價現金流量率，為組合內各股票的市價現金流量率的加權平均數。市價現金流量率是其中一個常用作評估公司財政狀況的衡量，比較公司股價及其營運現金流量。計算方法是以公司的現值除以其營運現金額，再除以流動股數。市價現金流量率反映投資者願意為公司每\$1營運現金所付出的價格。這比率顯示該公司製造現金的能力，因此可用作審核公司的現金流動狀況。

股息率係數：投資組合的加權平均股息率每額外上升一個單位時(其他因素維持不變)，投資者可獲得的實現回報。

長線盈利增長預測：投資組合的長線盈利增長預期，為組合內各股票的長線盈利增長預測的加權平均數。長線盈利增長預測數字愈高，代表企業的盈利增長預期愈好。

往績盈利增長：投資組合的往績盈利增長，為組合內各股票的往績盈利增長的加權平均數。往績盈利增長數字愈高，代表企業的過往的盈利增長愈好。

銷售額增長：投資組合的銷售額增長，為組合內各股票的銷售額增長的加權平均數。銷售額增長數字愈高，代表企業的銷售額增長愈快。

現金流增長：投資組合的現金流增長，為組合內各股票的現金流增長的加權平均數。現金流增長數字愈高，代表企業的財務狀況愈佳。

帳面值增長：投資組合的帳面值增長，為組合內各股票的帳面值增長的加權平均數。帳面值增長數字愈高，代表企業的償付能力愈好。

三年回報(%為單位)：算術平均回報是基金每月回報的算術平均值，並用於計算基金回報的標準差。回報(算術平均)並不完全等同於基金三年年度化回報。

三年績效(標準差)(%為單位)：基金績效是基金某一段期間的淨值成長率，計算方式是期末基金淨值減去前一期期末淨值所得的差額，再除以上期期末淨值。

標準差：標準差統計一組數值自平均值分散的程度，描述基金回報在特定時間內的變動。投資者可利用基金歷來回報的標準差來預測基金將來最有可能的回報範圍。基金的標準差愈大，其預期回報範圍就愈廣，也意味基金的波動性愈高。

如果投資者只持有一檔基金，標準差即最適用於測量該基金的風險。但如果投資組合持有多隻基金，投資者並不能視個別基金標準差的總和為組合的標準差，原因是組合的標準差不僅取決於個別基金的標準差，還取決於基金之間的回報關聯性。

如果基金回報屬常態分佈，其回報則有約 68% 機會分佈在距離平均值的 1 個標準差範圍之內，以及有約 95% 機會分佈在距離平均值的 2 個標準差範圍之內。晨星以既定時間內的月度總回報計算標準差，繼而將所有月度標準差年度化。標準差亦為夏普比率的組成部份，該比率評估投資的風險調整後回報。

夏普比率：現代投資理論的研究表明，風險的大小在決定組合的表現上具有基礎性的作用。風險調整後的收益率就是一個可以同時對收益與風險加以考慮的綜合指標，以期能夠排除風險因素對績效評估的不利影響。夏普比率就是一個可以同時對收益與風險加以綜合考慮的三大經典指標之一。

Alpha (%為單位)指數：是指股票超額報酬表現的部份(相對於大盤指數或是一籃子股票而言)也就是指超出 beta 所能估計的超額報酬或是指不能被大盤指數所解釋的超額報酬。

風險 Beta(指數)：是用以度量一項資產系統性風險的指標，是資本資產定價模型的參數之一。指用以衡量一種證券或一個投資證券組合相對總體市場的波動性的一種證券系統性風險的評估工具。其計算的方式為以過去 12 個月或 24 個月之基金月報酬率對同期市場月報酬率做迴歸，估計斜率係數而得，當 $\beta > 1$ ($\beta < 1$)，表示基金波動度較指數為大，當指數上揚 10%(下跌 10%)，基金會上揚超過 10%(下跌超過 10%)；當 $\beta = 1$ ，表示指數漲跌多少，基金就跟著變動多少。

風險 R^2 (指數)：比如指數分佈，伽瑪分佈，逆高斯分佈，泊松分佈，二項式分佈等。這是從聯系投資回報和所有風險性資產回報的模型 Beta 係數直接得出的。

第參章 研究方法

3.1 研究架構

本研究主要研究內容為使用亞洲(不包括日本)及金磚四國基金導入 WEKA 軟體製作決策樹及套入 LDA 之公式，並針對晨星之星等評級進行檢驗，本研究之研究架構圖，如圖 3-1 所示。

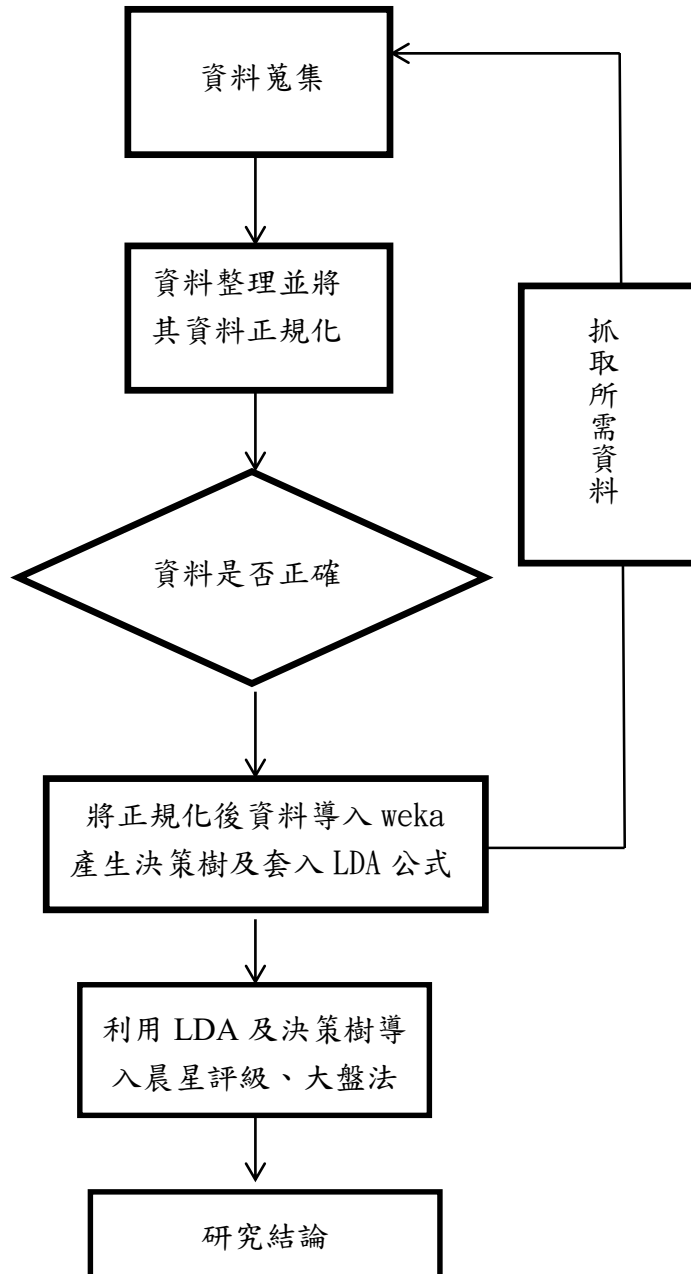


圖3.1 研究架構圖

3.2 決策樹(Decision Tree)

決策樹(Decision Tree)是指對所有已知的資料做分類或預測，並且利用資料中各個不同屬性值將資料分割成許多單一類別的子集合；是一功能強大且相當受歡迎的分類和預測工具；而且一次只觀察一個變數或屬性，以找出含最多資訊的變數。而我們唯一在意的東西，是分類或預測的正確性。

決策樹這種以樹狀圖為基礎的規則找尋法，其吸引人的地方在於它具有規則其產生出來的結果容易讓使用者了解與明白，因為使用者可以在不具備任何統計以及分析的知識下即可藉由決策樹來分析顧客以消費者的特質；決策樹的每一分支，就是一種對單一變數的檢驗，此檢驗會把空間分成兩塊或更多塊。一筆資料從根部的節點進入決策樹；在根部，應用一項測驗來決定這筆資料該進入下一層的哪一個子節點(Child Node)。選擇一開始的測驗有不同的演算法，但目的都是一樣的；這個過程一再重覆，直到資料到達葉部節點(Leaf Node)。所有到達某一個決策樹葉部的資料都以相同的方法來分類。從根部到每一個葉部都有一套獨特的路徑，這個路徑就是用來分類資料的規則的一種表達方式。

然而決策樹生成之後，可能會非常的複雜以及龐大，所以我們必須要對決策樹做修剪的動作；而修剪的最主要目的是為了移除葉節點與枝節的過程，其最終結果則是為了要改善決策樹的效能。

在決策樹中，比較著名的演算法有以下三種：ID3(C4.5)、CART、CHAID，其功能我們分別敘述如下：

一、ID3(Interactive Dichotometer 3)：

是指交互作用二分法的意思，是在 1983 年由 Quinlan 所提出的；由它的名稱，我們便可以清楚地知道 ID3 的邏輯特徵，意指如果在資料中有某項顯著特徵，它就會以此特徵將資料分成兩群，然後這兩群中如果又有一個為特殊特徵，就分為二，以此類推，反覆運作，直到所有同一特徵的資料都在一個類別裡為止。其做法是，首先選擇一個最佳的特徵作為根節點，由根節點開始，把所有資料依照此特徵分配到其所屬的分支，如果有某一個分支的訓練資料都已是同一個類別時，則此分支即成為一個樹葉，這個分支的推導也就算完成，可以結束了。同時這個分支也就是一個分類法則。否則只要在尚未找到樹葉時，分類的動作就得繼續下去，直至所有資料皆屬同一特徵、同一類別才能宣告完成。這個系統也之稱為觀念學習系統，是各種演算法中最典型的一個方法。

C4.5 是 ID3 的改良版，其運作的原理和過程皆和 ID3 相同，只不過 C4.5 可以處理遺漏的預測子和含有連續值的預測子；此外，C4.5 更加入了決策樹的修剪功能以及法則轉換功能。

二、CART(Classification and Regression Tree)：

分類迴歸樹：其最大的優點之一是演算法會自動檢驗模型，找出一個最佳的一般模型。這是由 Leo Breiman、Jerome Friedman、Richard Olshen 和 Charles Stone 在 1984 年所提出的演算法。CART 的方式是先建立一棵非常複雜的樹，然後再根據交互簡易或測試集檢驗的結果，將決策樹修剪成最佳的一般樹。在修剪之時，CART 是以整體錯誤率(Entire Error Rate)做為修剪依據，期望以最小的樹，也就是最少層的樹得到最有效的分類。CART 會針對任何節點先檢查節點中的資料是否屬於同一類別，若節點中的資料只屬於同一類別，則此節點不需要再分割。若節點中仍有二個以上的類別，則 CART 會測試所有資料屬性。

非連續屬性會檢查所有合併方式後，選擇最佳方法將屬性值合併為二組；連續屬性則會選擇最佳分離值將一為二。CART 和 C4.5 不同的是，C4.5 只根據訓練資料集來決定修剪結果；而 CART 則是根據各種版本對測試集資料的效能，來決定修剪的結果。

在這，必須有一個觀念，那就是最複雜的樹不一定是最好的樹，因為這樣有可能造成過度適應(Over fitting)的情形發生。此外 CART 在處理遺漏資料方面是相當拿手的，通常若某個屬性有遺漏值，就無法在建立樹的時候用來決定佳分割，事實上 CART 會儘量利用手上所有的資訊，將遺漏值透過替代(Surrogate) 的方式處理。而替代是樹中真正分割值和投入屬性的模擬，可以在投入屬性遺漏時使用。

三、CHAID(Chi-Square Automatic Interaction Detector)：

卡方自動互動偵測；此方法是三種決策樹演算法最古老的一種，是由 J.A.Hartigan 學者在 1975 年先提出的，而它是從 1964 年 J.A.Morgan 和 J.N.Sonquist 所提出的一套自動互動偵測系統 AID 所衍生出來的。其原理是偵測變數之間的統計關係，藉此建構出一棵決策樹。CHAID 與 CART 不同的是，CHAID 是利用連續卡方來測試並決定哪一類的預測子最不受預測值影響。

決策樹的建立以及演算法雖然複雜，但是在資料挖掘中卻是最常見的，因為它可以輕易地將結果表達，轉換法則使人清楚了解。

決策樹優缺點：

優點：

- 1.明白指出最佳變數的能力。
- 2.產生易於了解的規則。
- 3.節省分類時的計算時間。
- 4.可處理連續與類別變數。

缺點：

- 1.當類別太多時，資料量小，錯誤會增加的比較快。無法全面的看到各種變數的影響程度。
- 2.有時間順序的資料，需要很多預處理的工作。

決策樹範例(如圖 3-2 所示)：

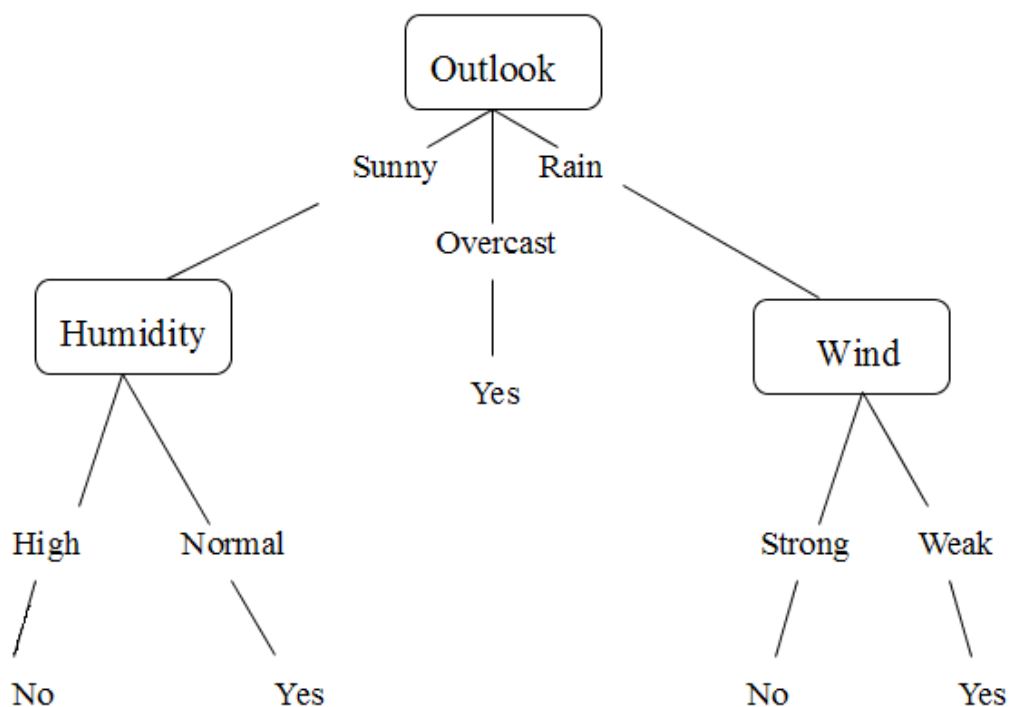


圖3.2 決策樹範例圖

這是一個典型的決策樹。這個決策樹根據天氣情況分類出(星期六上午是否適合打網球)，利用決策樹，可輕易的建立規則。

Rule 1： IF Outlook=Sunny AND Humidity=High THEN No

Rule 2： IF Outlook=Sunny AND Humidity=Normal THEN Yes

Rule 3： IF Outlook=Overcast THEN Yes

Rule 4： IF Outlook=Rain AND Wind=Strong THEN No

Rule 5： IF Outlook=Rain AND Wind=Weak THEN Yes

用 IF-THEN 規則來呈現知識。為每個根部到葉節點的途徑建立一條規則路徑上的每對屬性值形成關聯。葉節點具有類別預測。規則能夠簡易地被人類理解。

3.4 線性判別分析(Linear Discriminant Analysis, LDA)

線性判別分析為數學統計的方法，一般來說將待分析的 Raw Data 輸入這個方法，可以得到一條線性的方程式，容易了解，且分析快速。

然而由於科技的進步，衛星影像所得到的資訊與先進的輔助資訊，在測量科技上近十年來大幅進步，然而線性判別分析為非常傳統的方法，若將衛星影像所量測的雜訊代入，判釋的目標類別將引起大量的誤判和漏判。若可以在資料的萃取上和屬性刪減上改善，必對衛星影像在判釋科學上有很大的助益。

區別分析與迴歸分析概念相似之處在於，兩者都是在求得一組預測(自)變項的線性組合，其加權值在迴歸分析中稱為「迴歸係數」，在區別分析中則稱為「區別函數係數」，兩者都有原始(或未標準化)係數及標準化係數。不同的是，在迴歸分析中依變項通常是計量資料，但如果依變項是非計量的多分變項時，則應使用區別分析。因此區別分析的自變項(一般稱為預測變項)是計量性資料，而依變項(一般稱為分組變項)是非計量資料。

在解釋方面，區別分析主要在了解哪幾個變項可以區辨不同的組別，而這些組別間又有何不同；在預測方面，則是根據以往區別函數，如果研究者已經知道某一觀察體所有預測變項的數值，則可以對其進行分類，等事件發生後，再驗證分類的正確性。

其次，區別分析也與迴歸分析一樣，可以強迫將所有預測變項均投入分析，也可以採用逐步分析的方式，僅選擇較重要的變項。此外，迴歸分析會面臨的多元共線性問題，區別分析也會碰到。至於非計量的預測變項，在區別分析中可化為虛擬變項的方式處理。

線型區別函數式(Linear Discriminant Function)定義如下，利用此來進行判別(陳耀茂，2001)。

$$d_i = w_{i1}x_1 + w_{i2}x_2 + \dots + w_{ij}x_j + b_i \quad (3.2)$$

在公式(3.2)中， w_{ij} 表示判別係數，其中 x_j 代表輸入的變量， b_i 為 i 判別函數的常數項，這些判別函數可用來預測未知分群的資料。

然而 LDA 在分類上亦是有缺點的，由於線性判別分析為了將不同類別間的差距拉大，同時要把同類別內的的差距縮小，但是條件式需要足夠的訓練樣本才能達到最佳的分類效果。這就是知名的小樣本問題(small sample size problem)。

第四章 實證研究

4.1 基金篩選

此實證研究利用晨星之評級與大盤法來劃分決策值，再將屬性資料代入分類器，進而使用分類器產生新的決策值，對晨星評級進行比較研究，並以大盤法檢驗晨星之評級是否合理有效。

本研究將案例設計為以下兩種

案例 a：以決策樹分類-將訓練組資料導入決策樹中，產生新的決策值。

案例 b：以 LDA 分類-將訓練組資料導入線性判別分析公式中，產生新的決策值。

案例 c：將案例 a 與案例 b 之結果與晨星評級和大盤法互相比較。

4.2 訓練資料及測試資料

本研究蒐集晨星網站以 2012 年 5 月 15 號至 2015 年 5 月 15 號，以亞洲(不包括日本)基金為訓練組，金磚四國(俄羅斯、巴西、印度、中國)基金為測試組，在本研究中星等評級若為 1 或 2 以及 3 則決策值為 0，視為不可投資；若星等評級為 4 或 5 決策值則為 1，為可投資基金。如下表 4-1、4-2 所示：

表 4.1 訓練資料明細

基金名稱	晨 星 評 級	基 金 表 現	總回報率 (天)(%)	總回報率 (三年)(%)	市 盈 率	市 帳 率	市 銷 率	股 價 對 現 金 流	股 息 率 係 數	長 線 盈 利 增 長 預 測	往 績 盈 利 增 長	銷 售 額 增 長	現 金 流 增 長	帳 面 值 增 長	三 年 回 報 (%)	三 年 績 效 (標 準 差)(%)	夏 普 比 率	Alpha (%) 指數	風 險 Beta(指 數)	風 險 R ² (指 數)
GAM Star 亞洲股票基金累積單位-歐 元	2	1	0.71	9.13	13.7 9	1.95	0.6	0.09	1.67	15.35	12.16	10.06	20.41	12.16	0.82	15.12	0.65	0.12	1.01	65.74
NN (L) 亞洲高股息基金 P 股美元	3	0	0.15	5.23	10.1 8	1.09	1.06	0.16	3.72	9.79	9.13	3.51	6.72	9.13	0.5	13.07	0.45	-4.03	1.03	91.96
元大寶來新興亞洲基金	1	0	-0.19	7.42	12.9 9	1.81	0.72	0.12	1.61	11.93	13.77	11.17	22.76	13.77	0.65	10.81	0.59	-3.03	0.98	78.9
凱基亞洲四金磚基金	4	0	-0.33	14.47	12.9 2	1.71	0.7	0.1	2.39	15.09	9.65	12.2	21.41	9.65	1.18	11.19	1.15	4.31	0.88	60.27
凱基台商天下基金	4	0	-0.19	12.06	13.1 3	1.79	1.09	0.11	2.42	10.55	11.12	9.92	-11.87	11.12	0.98	8.22	1.26	5.87	0.47	31.52

匯豐環球投資基金-亞洲小型公司股票(日本除外)AC	5	1	0.73	22.69	8.31	1.05	1.62	0.14	2.51	21.43	1.47	3.1	10.84	1.47	1.82	15.92	1.37	13.09	0.9	47.58
匯豐環球投資基金-亞洲股票(日本除外)AD	3	1	0.99	6.18	9.42	1.26	0.72	0.12	3.01	10.48	9.29	4.95	-5.95	9.29	0.58	13.76	0.5	-3.55	1.09	92.11
天達環球策略基金 - 亞洲股票基金C 收益股份	4	1	0.95	9.04	10.55	1.31	0.65	0.14	2.64	9.11	10.64	8.01	7.6	10.64	0.79	12.3	0.76	-0.12	0.99	95.14
宏利環球基金-亞洲股票基金AA股	3	1	1.13	11.92	11.11	1.38	0.81	0.15	2.2	12.11	7.16	5.67	-11.14	7.16	1	12.08	0.99	3.28	0.9	82.77
富蘭克林坦伯頓全球投資系列-亞洲小型企業基金歐元A(acc)股	5	1	0.68	14.07	11.42	1.35	0.84	0.11	2.62	18.05	12.96	11.82	3.69	12.96	1.17	12.9	1.09	5.84	0.85	63.98

富蘭克林坦伯頓全球投資系列-亞洲 成長基金美元 A (acc)股	2	0	0.06	1.91	11.2 4	1.85	0.79	0.11	3.51	12.05	10.43	11.01	13.93	10.43	0.23	12.99	0.21	-5.7	0.87	66.23
富達基金-東南亞基金	3	0	0.5	5.75	14.4 3	1.95	0.42	0.09	2	10.51	12.28	4.25	9.48	12.28	0.53	12	0.52	-2.73	0.93	89.75
富達基金-亞洲成長趨勢基金(歐元累 積)	4	0	0.18	10.19	14.7 4	1.89	0.64	0.1	2.05	11.08	10.84	7.51	7.01	10.84	0.87	11.71	0.89	1.48	0.92	92
富達基金-新興亞洲基金(歐元)	3	0	0.54	9.58	15.4 8	2.31	0.46	0.08	1.82	11.81	15.76	10.03	4.5	15.76	0.84	13.51	0.74	0.91	0.95	72.84
富邦新興亞洲高成長基金	2	0	-0.09	6.6	13.7 7	1.81	0.74	0.08	1.67	11.58	9.64	11.12	9.78	9.64	0.59	12.11	0.48	-1.91	0.79	41.7
德意志亞洲中小型基金 FC	4	1	1.05	14.75	12.7 5	1.8	0.99	0.11	2.47	24.77	-1.62	24.45	1.2	-1.62	1.23	13.26	1.1	6.17	0.88	65.25
德盛老虎基金	4	1	1.06	11.38	14.1 4	1.91	0.48	0.08	1.71	14.58	10.02	10.11	5.97	10.02	0.97	12.83	0.9	2.22	0.97	85.06

德盛安聯中國東協新世紀基金	3	0	0.07	11.44	11.01	1.55	0.89	0.17	2.83	8.36	14.9	13.31	31.18	14.9	0.95	10.45	0.96	1.52	0.88	69.2
德盛亞洲總回報股票基金—A 配息類股 (美元)	4	1	1.29	10	15.98	2.38	0.4	0.07	1.69	17.41	10.01	10.92	5.64	10.01	0.86	12.71	0.81	1.12	0.95	83.48
德盛新興亞洲基金	4	0	0.5	13.18	14.99	2.03	0.73	0.1	1.29	19.68	10.19	10.1	14.69	10.19	1.1	12.46	1.06	4.14	0.94	83.36
摩根東方小型企業基金	2	0	0.24	15.77	14.58	1.84	0.73	0.08	1.86	18.58	6.02	5	2.34	6.02	1.29	12.28	1.26	7.02	0.87	74.63
摩根東方內需機會基金	3	1	0.43	9.56	12.5	1.7	0.8	0.11	1.97	16.08	13.83	13.43	13.97	13.83	0.82	11.18	0.75	-1.65	1.04	84.3
摩根東方基金	2	0	0.11	6.89	11.88	1.68	0.62	0.11	2.05	12.66	9.71	4.53	4.9	9.71	0.62	12.54	0.59	-2.17	0.99	93.07
摩根亞洲內需主題基金	2	1	0.78	8.41	12.9	1.67	0.7	0.1	2.01	12.91	9.94	5.35	-4.45	9.94	0.74	12.99	0.68	-0.8	1.01	88.55

摩根士丹利亞洲股票基金 A	2	1	0.96	6.19	13.44	1.68	0.48	0.09	2.25	10.2	11.64	8.19	6.42	11.64	0.57	12.44	0.54	-2.62	0.97	90.27
新加坡大華亞洲基金	3	0	0.28	6.47	12.24	1.61	0.89	0.14	2.68	9.18	10.9	10.89	19.4	10.9	0.59	12.31	0.57	-2.54	0.99	95.45
施羅德環球基金系列 - 亞洲小型公司 A1 類股份 - 累積單位	3	0	0.32	13.55	16.8	2.88	0.6	0.07	2.3	18.63	8.46	8.33	10.55	8.46	1.13	12.48	1.08	5.95	0.79	58.12
施羅德環球基金系列 - 亞洲優勢 A1 類股份 - 累積單位	3	1	0.89	6.82	15.73	1.77	0.4	0.07	2.01	12.7	10.96	4.16	3.53	10.96	0.61	11.54	0.63	-1.12	0.87	83.39
施羅德環球基金系列 - 新興亞洲 A 類股份 - 累積單位	4	1	0.64	8.39	12.03	1.6	0.71	0.1	2.46	11.24	10.99	5.8	3.62	10.99	0.73	11.9	0.73	-0.18	0.93	89.45
普信亞洲(除日本)股票型基金 A	3	0	0.62	6.21	15.12	1.8	0.49	0.08	1.96	11.84	9.42	3.18	-10.12	9.42	0.56	11.3	0.58	-1.87	0.88	89.93
景順亞洲機遇股票基金 A 股 美元	3	1	0.73	10.21	15.71	2.03	0.61	0.07	2.68	12.42	10.59	6.22	0.01	10.59	0.87	11.75	0.88	1.56	0.92	89.98

景順亞洲動力基金 A-年配息股 美元	4	1	1.26	9.43	11.6	1.24	0.95	0.1	2.17	11.52	8.65	2.68	-1.19	8.65	0.81	11.67	0.83	0.68	0.93	94.56
未來資產亞洲卓越消費股票基金-A(美元)	4	1	1.15	14.37	23.5	4.55	0.38	0.04	0.85	20.4	13.23	11.75	21.97	13.23	1.2	13.15	1.09	47.02	0.76	48.8
柏瑞環球基金-柏瑞亞洲(日本除外)股票基金 A	3	1	1.33	13.11	12.7 2	1.62	0.72	0.12	1.75	14.33	10.95	7.5	14.16	10.95	1.1	12.52	1.05	3.86	0.96	86.93
歐義銳榮新興亞洲股票基金 R	2	1	0.86	6.06	12.3 6	1.39	0.87	0.12	2.15	11.01	-11.78	-13.12	1.11	-11.78	0.56	12.95	0.51	-2.66	0.97	82.81
法儲銀富鑫新興亞洲基金 R/A EUR	2	0	0.34	5.99	12.5 7	1.79	0.62	0.11	2.19	10.49	13.22	9.19	15.15	13.22	0.55	12.39	0.53	-3.05	1	95.55
法巴百利達亞洲(日本除外)精選股票基金 C(歐元)	3	0	0.01	9.36	13.0 8	1.85	0.56	0.1	2.05	11.18	12.09	6.85	10.52	12.09	0.82	13.45	0.73	-0.32	1.05	90.76
瀚亞投資-亞洲股票基金 A(美元)	2	0	0.62	4.04	10.4 9	1.09	1.25	0.15	2.93	10.01	6.55	1.55	-1.78	-6.55	0.4	13.04	0.36	-5.35	1.05	95.95

第一金亞洲新興市場基金	2	1	0.27	5.91	16.75	2.65	0.52	0.1	1.81	17.96	12	15.77	19.62	12	0.54	12.15	0.42	-2.3	0.77	38.69
美盛百駿亞洲(日本除外)股票基金 A類股歐元累積型	2	0	0.42	7.03	11.51	1.42	0.83	0.16	4.24	10.04	11.32	5.39	10.32	11.32	0.64	12.9	0.59	-2.27	1.02	93.19
聯博-亞洲股票基金 AD 股歐元	2	1	0.99	7.93	9.01	1.09	0.97	0.16	3.51	12.59	10.07	12.31	3.31	10.07	0.73	14.69	0.59	-2.29	1.14	88.66
英傑華投資-亞洲股票收益基金	2	0	0.29	7.3	10.32	1.16	0.71	0.12	4.05	7.44	11.41	-0.5	-7.81	11.41	0.65	12.44	0.62	-1.9	1.01	96.55
貝萊德亞洲巨龍基金 A2	3	0	-0.05	13.19	12.55	1.42	0.69	0.09	2.32	11.22	-0.32	6.18	15.23	-0.32	1.11	12.9	1.03	3.44	1.02	91.85
鋒裕基金-亞洲股票(不含日本)A2	2	1	0.82	7.05	12.57	1.55	0.66	0.1	2.56	12.54	10.89	8.15	9.99	10.89	0.64	12.58	0.6	-2.01	0.99	92.39
霸菱亞洲增長基金-A類歐元配息型	3	1	1.48	8.77	18.4	2.99	0.4	0.06	1.32	16.34	15.07	13.48	18.06	15.07	0.77	12.5	0.73	-0.21	0.97	89.49

表 4.2 測試資料明細

基金名稱	晨 星 評 級	基 金 表 現	總回報 率 (天)(%)	總回報率 (三年)(%)	市 盈 率	市 帳 率	市 銷 率	股價對 現金流	股息 率係 數	長線盈利 增長預測	往績盈 利增長	銷售 額增 長	現金 流增 長	帳面 值增 長	三年回 報(%)	三年績效(標 準差)(%)	夏普 比率	Alpha (%) 指數	風險 Beta(指 數)	風險 R²(指 數)
匯豐環球投資基金-俄羅斯股票 AD	3	0	-1.59	-11.2	6.01	0.72	1.7	0.38	7.34	8.34	-17.73	-7.21	-10.02	-17.73	-0.63	29.62	-0.26	-2.54	0.95	98.04
宏利環球基金-俄羅斯股票基金 AA 股	3	0	-2.16	-8.97	5.57	0.95	1.08	0.38	10.33	7.49	-21.35	9.38	-6.57	-21.35	-0.45	28.3	-0.19	-0.7	-0.9	95.74
摩根俄羅斯基金	3	0	0	-10.19	7.89	0.95	1.29	0.51	6.79	12.8	-5.69	4.87	-2.26	-5.69	-0.6	26.71	-0.27	-2.83	0.84	92.42
法巴百利達俄羅斯股票基金 C (歐元)	4	1	0.87	-7.2	6.16	0.88	1.76	0.35	8.49	11.13	-24.33	5.76	11.7	-24.33	-0.32	27.03	-0.14	0.61	0.85	93.55
瑞銀(盧森堡)俄羅斯股票基金(美元)	3	1	1.46	-10.32	6.97	0.81	1.33	0.44	8.76	9.37	-17.12	-3.69	-15.98	-17.12	-0.56	29.04	-0.23	-1.88	0.92	95.33
霸菱俄羅斯基金-A 類美元累積型	3	1	1.14	-10.05	6.74	1.19	1.14	0.31	6.07	10.11	-51.76	0.8	2.48	-51.76	0.53	28.96	-0.22	-1.6	0.92	96.05
元大寶來巴西指數基金	4	1	1.68	-10.65	12.18	1.67	0.73	0.12	2.77	13.44	7.05	7.47	7.19	7.05	-0.7	23.64	-0.42	-0.68	0.94	98.82
匯豐環球投資基金-巴西股票 AC	2	1	2.59	-15.68	13.82	1.72	0.82	0.12	4.49	11.39	2.15	0.01	-3.74	2.15	-1.13	25.81	-0.53	-4.88	0.93	95.17
摩根巴西基金	3	0	0.16	-15.57	14.13	2.02	0.49	0.07	6.12	13.23	4.24	2	1.62	4.24	-1.16	23.96	-0.58	-5.88	0.86	94.48
法巴百利達巴西股票基金 C (美元)	3	0	-0.33	-13.24	12.94	1.16	1.09	0.17	2.99	12.02	-1.12	0.06	-4.8	-1.12	-0.9	25.95	-0.42	-1.93	0.94	96.97
瑞銀(盧森堡)巴西股票基金(美元)	2	0	-0.02	-15.72	18.83	2.61	0.48	0.07	8.91	13.22	15.77	15.43	13.24	15.77	-1.12	26.41	-0.51	-4.43	0.97	98.6

野村巴西基金	4	1	1.83	-12.01	12.37	1.57	1.03	0.14	2.34	14.85	3.57	2.68	2.02	3.57	-0.86	22.2	-0.52	-3.16	0.87	96.56
元大寶來印度指數基金	1	0	0.57	7.99	18.52	2.74	0.61	0.1	1.63	14.86	14.16	15.93	14.6	14.16	0.8	19.5	0.42	-2.14	0.96	96.57
元大寶來印度基金	2	0	0	8.23	20.59	3.29	0.51	0.08	1.07	17.85	15.46	15.66	19.3	15.46	0.8	18.4	0.45	-1.48	0.9	95.29
利安資金印度基金(美元)	2	0	0.53	7.78	16.38	2.14	0.72	0.1	1.32	17.3	15.27	9.89	1.55	15.27	0.78	19.66	0.48	-0.49	0.92	97.8
匯豐環球投資基金-印度股票 AD	1	0	0.92	7.2	13.65	1.22	0.71	0.17	1.91	14.52	5.28	8.75	19.33	5.28	0.91	28.77	0.38	-2.72	1.27	87.34
台新印度基金	3	0	0.63	11.74	19.88	3	0.56	0.11	1.09	16.91	12.21	22.97	8.53	12.21	1.1	20.45	0.58	1.41	0.97	89.32
安本環球 - 印度股票基金 A2	5	1	1.38	13.37	22.51	3.59	0.29	0.25	1.38	16.13	13.51	11.65	0.45	13.51	1.19	18.78	0.76	5.14	0.85	92.41
宏利環球基金-印度股票基金 AA 股	3	1	1.33	13.28	21.53	3.54	0.34	0.06	1.09	17.16	17.32	18.3	9.29	17.32	1.22	20.98	0.7	4.15	0.98	97.11
富蘭克林坦伯頓全球投資系列-印度基金 美元 A(acc)股	4	0	0.84	13.34	22.89	3.16	0.57	0.08	0.97	19.56	16.72	9.73	7.02	16.72	1.23	21.41	0.69	4.27	0.98	93.95
富達基金-印度聚焦基金	3	0	0.32	11.56	18.16	3.09	0.57	0.07	1.53	16.7	15.55	15.93	10.07	15.55	1.1	21.31	0.62	2.62	0.99	95.47
愛德蒙得洛希爾印度基金(A)	3	1	2.01	10.22	22.12	4.15	0.41	0.08	0.81	20.48	23.09	33.41	53.75	23.09	0.96	18.86	0.61	2.18	0.87	94.44

摩根印度基金	4	0	0.57	10.94	22.67	2.88	0.41	0.1	1.03	18.85	16.16	11.05	23.84	16.16	1.06	21.8	0.58	1.92	1.01	95.22
摩根士丹利印度股票基金 A	3	1	2.59	14.77	19.89	3.69	0.55	0.06	1.11	17.95	12.63	16.79	19.18	12.63	1.32	19.93	0.79	6.09	0.9	91.85
施羅德環球基金系列 - 印度股票 A1 類 股份-累積單位	3	0	0.96	11.98	18.79	3	0.44	0.06	1.1	17.45	14.94	18.39	10.47	14.94	1.1	19.22	0.68	3.58	0.89	96.14
木星全球系列-木星印度精選基金 L 類 股 美元年收益	3	1	2.32	10.66	14.98	1.85	1.38	0.11	1.44	17.15	8.03	15.03	30.44	8.03	1.07	23.38	0.55	1.63	1.04	88.77
柏瑞環球基金-柏瑞印度股票基金 A	4	0	0.98	16.87	25.58	4.62	0.5	0.1	0.93	24.47	15.25	20.06	47.46	15.25	1.49	21.52	0.83	7.52	0.97	90.64
法巴百利達印度股票基金 C (美元)	4	0	0.93	12.1	24.55	2.94	0.72	0.08	1.04	18.68	12.63	13.77	16.14	12.63	1.12	20.27	0.66	3.33	0.94	96.58
瀚亞印度基金	3	0	0.09	13.72	19.2	2.72	0.66	0.09	1.17	16.66	15.2	20.81	28.71	15.2	1.24	20	0.67	2.99	0.98	94.55
瀚亞投資-印度股票基金 A(美元)	2	1	1.48	9.02	17.59	2.81	0.53	0.08	2.64	16.84	15.23	13.55	12.74	15.23	0.9	20.83	0.51	0.37	0.97	96.16
群益印度中小基金	3	0	0.41	14.55	24.99	4.4	0.43	0.05	0.7	27.47	17.41	22.25	-1.75	17.41	1.32	21.28	0.68	4.27	0.95	79.32
聯博-印度成長基金 AX 股	4	1	1.13	12.27	19.37	3.08	0.4	0.09	1.06	20.33	21.4	10.34	17.64	21.4	1.19	23.59	0.6	2.66	1.09	94.62
貝萊德印度基金 A2	3	1	1.82	13.69	18.42	2.84	0.93	0.1	1.12	18.48	10.9	14.61	11.53	10.9	1.26	21.55	0.7	4.39	1	96.44

野村雙印傘型基金之印度潛力基金	3	0	0.27	11.13	19.14	3.47	0.53	0.08	1.14	17.63	15.33	9.69	19.98	15.33	1.04	19.93	0.56	0.74	0.97	94.06
首域環球傘型基金-首域印度次大陸基金	5	1	1.12	23.07	20.77	2.57	0.5	0.08	1.47	17.05	16.14	18.4	37.9	16.14	1.89	18.93	1.2	13.51	0.85	90.6
匯豐環球投資基金-中國股票 AD	3	0	1.21	16.29	9.52	1.39	0.78	0.12	3.03	10.76	10.03	7.7	0.36	10.03	1.4	18.63	0.9	0.71	1.04	94.15
國泰中國新興戰略基金	5	0	-0.77	23.57	11.43	2.05	0.94	0.07	1.63	14.06	13.83	15.35	-9.49	13.83	1.87	15.39	1.37	13.24	0.51	26.36
富達基金-中國聚焦基金(歐元累積)	4	0	0.41	17.19	12.64	1.51	0.76	0.12	2.58	9.73	11.56	5.41	1.96	11.56	1.45	17.34	1	2.5	0.96	92.77
富邦大中華成長基金	4	0	0	15.57	9.83	1.53	0.55	0.14	2.42	10.3	13.55	14.68	-17.64	13.55	1.31	15.56	0.92	0.09	0.92	84.31
德意志中國股票 FC	3	1	1.92	13.45	10.05	1.27	0.78	0.23	2.47	13.68	10.24	5.96	24.59	10.24	1.2	19.05	0.75	-1.81	1.04	90.89
德盛安聯中國策略增長基金	3	0	0	15.55	10.92	1.32	1.33	0.14	2.55	13.11	12.12	11.29	10.26	12.12	1.31	15.5	0.92	-0.65	0.97	93.89
摩根中國基金	3	0	1.01	13.84	11.34	1.44	0.82	0.14	2.45	12.74	12.36	6.74	5.17	12.36	1.22	18.33	0.8	-1.29	1.02	94.84
施羅德環球基金系列 - 中國優勢 A1 類 股份 - 累積單位	4	0	1.19	14.39	11.92	1.55	0.71	0.11	2.83	10.9	8.57	5.22	2.67	8.57	1.23	15.86	0.92	0.85	0.89	95.78
景順中國基金 A-年配息股 美元	3	1	1.41	13.26	12.97	1.7	0.51	0.09	3.03	10.01	14.27	9.33	7.77	14.27	1.18	18.29	0.77	-0.7	0.95	82.53
木星中國基金 (ACC)	2	1	1.62	15.08	10.27	1.53	0.72	0.11	2.18	10.72	11.73	9.17	17.84	11.73	1.3	17.21	0.9	1.65	0.89	81.77

歐義銳榮中國股票基金 R	2	0	0.61	12.87	10.46	1.38	0.86	0.14	3.1	10.2	8.85	6.65	4.2	8.85	1.15	18.45	0.75	-1.4	0.98	85.31
瀚亞中國基金	2	0	-0.14	11.63	8.17	1.05	0.83	0.17	4.11	9.73	15.17	12.51	5.11	15.17	1.01	14.84	0.72	-3.63	0.93	94.46
瀚亞投資-中國股票基金 A(美元)	2	1	1.55	10.28	10.16	1.26	0.72	0.13	3.6	9.91	11.83	4.69	-0.21	11.83	0.93	16.65	0.67	-3.46	0.94	96.71
瑞銀(盧森堡)中國精選股票基金(美元)	3	0	1.22	23.07	12.65	1.45	0.52	0.07	2.14	13.5	10.88	6.18	9.27	10.88	1.87	17.62	1.27	8.78	0.87	74.74
群益華夏盛世基金	4	0	-0.82	20.2	12.71	2.26	0.68	0.13	2.06	10.49	12.31	13.69	-15.33	12.31	1.64	15.86	1.16	9.01	0.6	34.95
聯博-中國時機基金 A 股	3	1	1.5	15.51	9.16	1.23	0.89	0.18	3.21	11.57	12.98	7.97	5.84	12.98	1.32	16.57	0.95	2.15	0.88	85.15
荷寶資本成長基金-荷寶中國股票 D 美元	4	0	1.11	17.82	10.31	1.19	0.8	0.21	3.28	10.54	0.81	7.03	24.37	0.81	1.5	17.98	1	2.45	1	94.38
貝萊德中國基金 A2 美元	4	0	1.11	17.62	10.31	1.19	0.62	0.21	3.28	10.54	0.81	7.03	21.36	0.81	1.5	17.98	1	2.45	1	93.27
鋒裕基金-中國股票 A2	3	1	1.32	15.28	10.3	1.48	0.71	0.14	2.66	11.25	9.16	5.26	3.49	9.16	1.31	16.98	0.92	0.75	0.96	96.82
霸菱香港中國基金-A類 歐元配息型	3	1	2.1	12.54	13.52	1.63	0.6	0.1	2.39	9.79	13	6.83	10.28	13	1.1	17.05	0.77	-1.61	0.96	95.39
首域環球傘型基金-首域中國核心基金-第四類股-USD	3	0	0.78	11.78	16.03	1.67	0.8	0.1	2.52	13.21	9.95	9.28	13.67	9.95	1.01	13.41	0.9	1.72	0.66	74.4
匯豐環球投資基金-新興四國市場股票 AC	3	0	0.75	1.58	7.58	0.84	1.22	0.28	4.43	9.84	-3.55	1.39	-11.82	-3.55	0.26	17.78	0.17	-2.13	0.99	96.67

富蘭克林坦伯頓全球投資系列-亞洲(不包括日本)基金美元 A (acc)股	1	1	0.93	-0.79	8.7	1.06	0.98	0.14	4	9.25	11.74	7.28	8.27	11.74	0.06	17.78	0.04	-4.37	0.96	91.47
德盛亞洲(不包括日本)基金-AT 累積類股 (歐元)	2	0	0.5	1.41	13.22	1.86	0.61	0.1	3.52	11.49	-14.7	1.95	-12.57	-14.7	0.25	17.71	0.16	-2.05	0.94	88.13
摩根新金磚五國基金	3	0	-0.49	2.14	10.49	1.63	0.86	0.14	4.45	11.14	7.72	10.86	2.25	7.72	0.27	14.88	0.12	-2.81	0.91	86.69
施羅德環球基金系列 - 亞洲(不包括日本) A 類股份 - 累積單位	4	1	1.28	2.62	12.36	1.67	0.81	0.12	4.05	11.66	6.54	2.5	-2.67	6.54	0.33	16.28	0.24	0.82	0.89	93.24
法巴百利達亞洲(不包括日本)股票基金 C (美元)	2	0	0.15	0.41	11.88	1.48	1	0.19	2.83	12.68	-19.29	9.97	9.52	-19.29	0.17	17.83	0.11	-3.15	0.97	91.66

4.3 晨星評級

4.3 部分是以晨星評級為決策值所做的分析。

4.3.1 決策樹分析結果(案例 a)

下圖為將訓練資料導入決策樹方法所畫出的決策樹圖，測試資料也需導入其中判斷，為了軟體執行順利，我們將導入資料之屬性名稱做了改變，B16 為 Alpha (%)指數、B6 為股價對現金流、B1 為總回報率(天)(%)。

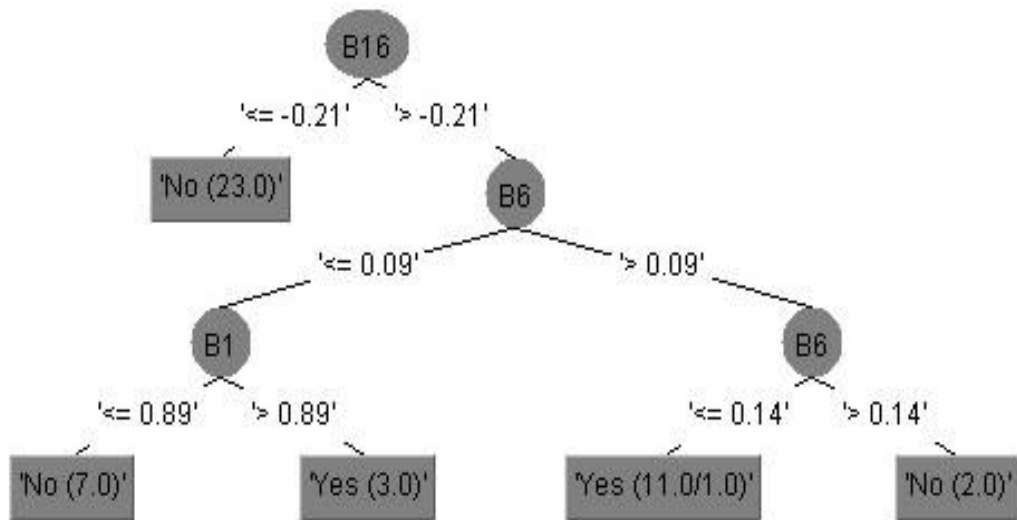


圖4.1 使用訓練資料產生決策樹圖

表 4.3 導入決策樹之訓練資料結果

編號	基金名稱	決策值	分類結果	結果比對
1	GAM Star 亞洲股票基金累積單位-歐元	0	0	
2	NN (L) 亞洲高股息基金 P 股美元	0	0	
3	元大寶來新興亞洲基金	0	0	
4	凱基亞洲四金磚基金	1	1	
5	凱基台商天下基金	1	1	
6	匯豐環球投資基金-亞洲小型公司股票(日本除外) AC	1	1	
7	匯豐環球投資基金-亞洲股票(日本除外) AD	0	0	
8	天達環球策略基金 - 亞洲股票基金 C 收益股份	1	1	
9	宏利環球基金-亞洲股票基金 AA 股	0	0	
10	富蘭克林坦伯頓全球投資系列-亞洲小型企業基金歐元 A (acc)股	1	1	
11	富蘭克林坦伯頓全球投資系列-亞洲成長基金美元 A (acc)股	0	0	
12	富達基金-東南亞基金	0	0	
13	富達基金-亞洲成長趨勢基金(歐元累積)	1	1	
14	富達基金-新興亞洲基金(歐元)	0	0	
15	富邦新興亞洲高成長基金	0	0	
16	德意志亞洲中小型基金 FC	1	1	
17	德盛老虎基金	1	1	
18	德盛安聯中國東協新世紀基金	0	0	
19	德盛亞洲總回報股票基金—A 配息類股 (美元)	1	1	
20	德盛新興亞洲基金	1	1	
21	摩根東方小型企業基金	0	0	
22	摩根東方內需機會基金	0	0	
23	摩根東方基金	0	0	
24	摩根亞洲內需主題基金	0	0	
25	摩根士丹利亞洲股票基金 A	0	0	
26	新加坡大華亞洲基金	0	0	
27	施羅德環球基金系列 - 亞洲小型公司 A1 類股份 - 累積單位	0	0	
28	施羅德環球基金系列 - 亞洲優勢 A1 類股份 - 累積單位	0	0	
29	施羅德環球基金系列 - 新興亞洲 A 類股份 - 累積單位	1	1	
30	普信亞洲(除日本)股票型基金 A	0	0	
31	景順亞洲機遇股票基金 A 股 美元	0	0	
32	景順亞洲動力基金 A-年配息股 美元	1	1	
33	未來資產亞洲卓越消費股票基金-A(美元)	1	1	

34	柏瑞環球基金-柏瑞亞洲(日本除外)股票基金 A	0	1	結果不符
35	歐義銳榮新興亞洲股票基金 R	0	0	
36	法儲銀富鑫新興亞洲基金 R/A EUR	0	0	
37	法巴百利達亞洲(日本除外)精選股票基金 C(歐元)	0	0	
38	瀚亞投資-亞洲股票基金 A(美元)	0	0	
39	第一金亞洲新興市場基金	0	0	
40	美盛百駿亞洲(日本除外)股票基金 A類股歐元累積型	0	0	
41	聯博-亞洲股票基金 AD 股歐元	0	0	
42	英傑華投資-亞洲股票收益基金	0	0	
43	貝萊德亞洲巨龍基金 A2	0	0	
44	鋒裕基金-亞洲股票(不含日本)A2	0	0	
45	霸菱亞洲增長基金-A類歐元配息型	0	0	
46	首域環球傘型基金-首域亞洲增長基金-第四類股-USD	0	0	
	可投資筆數			13 筆
	不可投資筆數			32 筆
	結果不符筆數			1 筆
	資料總計			46 筆
	資料吻合度			97.83%

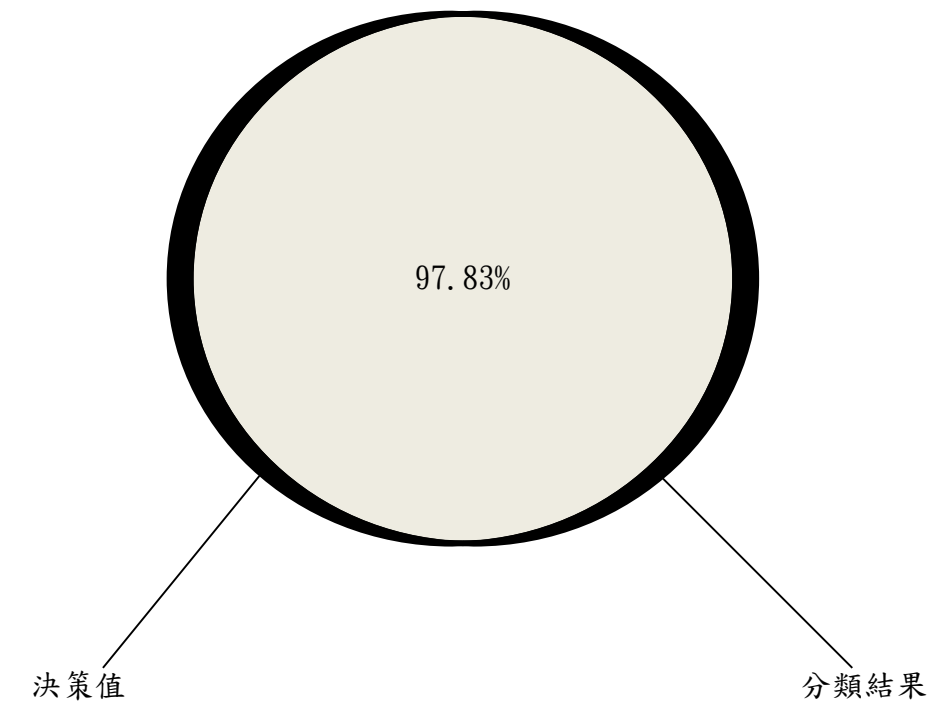


圖4.2 導入決策樹之訓練資料交集聯集圖

表 4.4 導入決策樹之訓練資料分類結果表

分類結果	基金編號
可投資	4、5、6、8、10、13、16、17、19、 20、29、32、33
不可投資	1、2、3、7、9、11、12、14、15、 18、21、22、23、24、25、26、27、 28、30、31、35、36、37、38、39、 40、41、42、43、44、45、46
結果不符	34

表 4.5 測試資料導入訓練資料之決策樹結果

編號	基金名稱	決策值	分類結果	結果比對
1	匯豐環球投資基金-俄羅斯股票 AD	0	0	
2	宏利環球基金-俄羅斯股票基金 AA 股	0	0	
3	摩根俄羅斯基金	0	0	
4	法巴百利達俄羅斯股票基金 C (歐元)	1	0	結果不符
5	瑞銀(盧森堡)俄羅斯股票基金(美元)	0	0	
6	霸菱俄羅斯基金-A 類美元累積型	0	0	
7	元大寶來巴西指數基金	1	0	結果不符
8	匯豐環球投資基金-巴西股票 AC	0	0	
9	摩根巴西基金	0	0	
10	法巴百利達巴西股票基金 C (美元)	0	0	
11	瑞銀(盧森堡)巴西股票基金(美元)	0	0	
12	野村巴西基金	1	0	結果不符
13	元大寶來印度指數基金	0	0	
14	元大寶來印度基金	0	0	
15	利安資金印度基金(美元)	0	0	
16	匯豐環球投資基金-印度股票 AD	0	0	
17	台新印度基金	0	1	結果不符
18	安本環球 - 印度股票基金 A2	1	0	結果不符
19	宏利環球基金-印度股票基金 AA 股	0	1	結果不符
20	富蘭克林坦伯頓全球投資系列-印度基金美元 A(acc)股	1	0	結果不符
21	富達基金-印度聚焦基金	0	0	
22	愛德蒙得洛希爾印度基金(A)	0	1	結果不符
23	摩根印度基金	1	1	
24	摩根士丹利印度股票基金 A	0	1	結果不符
25	施羅德環球基金系列 - 印度股票 A1 類股份-累積單位	0	1	結果不符
26	木星全球系列-木星印度精選基金 L 類股 美元年收益	0	1	結果不符
27	柏瑞環球基金-柏瑞印度股票基金 A	1	1	
28	法巴百利達印度股票基金 C (美元)	1	1	
29	瀚亞印度基金	0	0	
30	瀚亞投資-印度股票基金 A(美元)	0	1	結果不符
31	群益印度中小基金	0	0	
32	聯博-印度成長基金 AX 股	1	1	
33	貝萊德印度基金 A2	0	1	結果不符
34	野村雙印傘型基金之印度潛力基金	0	0	

35	首域環球傘型基金-首域印度次大陸基金	1	1	
36	匯豐環球投資基金-中國股票 AD	0	1	結果不符
37	國泰中國新興戰略基金	1	0	結果不符
38	富達基金-中國聚焦基金(歐元累積)	1	1	
39	富邦大中華成長基金	1	1	
40	德意志中國股票 FC	0	0	
41	德盛安聯中國策略增長基金	0	0	
42	摩根中國基金	0	0	
43	施羅德環球基金系列 - 中國優勢 A1 類股份 - 累積單位	1	1	
44	景順中國基金 A-年配息股 美元	0	0	
45	木星中國基金 (ACC)	0	1	結果不符
46	歐義銳榮中國股票基金 R	0	0	
47	瀚亞中國基金	0	0	
48	瀚亞投資-中國股票基金 A(美元)	0	0	
49	瑞銀(盧森堡)中國精選股票基金(美元)	0	1	結果不符
50	群益華夏盛世基金	1	1	
51	聯博-中國時機基金 A 股	0	0	
52	荷寶資本成長基金-荷寶中國股票 D 美元	1	0	結果不符
53	貝萊德中國基金 A2 美元	1	0	結果不符
54	鋒裕基金-中國股票 A2	0	1	結果不符
55	霸菱香港中國基金-A 類 歐元配息型	0	0	
56	首域環球傘型基金-首域中國核心基金-第四類股-USD	0	1	結果不符
57	匯豐環球投資基金-新興四國市場股票 AC	0	0	
58	富蘭克林坦伯頓全球投資系列-金磚四國基金美元 A (acc)股	0	0	
59	德盛金磚四國基金—AT 累積類股 (歐元)	0	0	
60	摩根新金磚五國基金	0	0	
61	施羅德環球基金系列 - 金磚四國 A 類股份 - 累積單位	1	1	
62	法巴百利達金磚四國股票基金 C (美元)	0	0	
	可投資筆數			10 筆
	不可投資筆數			31 筆
	結果不符筆數			21 筆
	資料總計			62 筆
	資料吻合度			66.12%

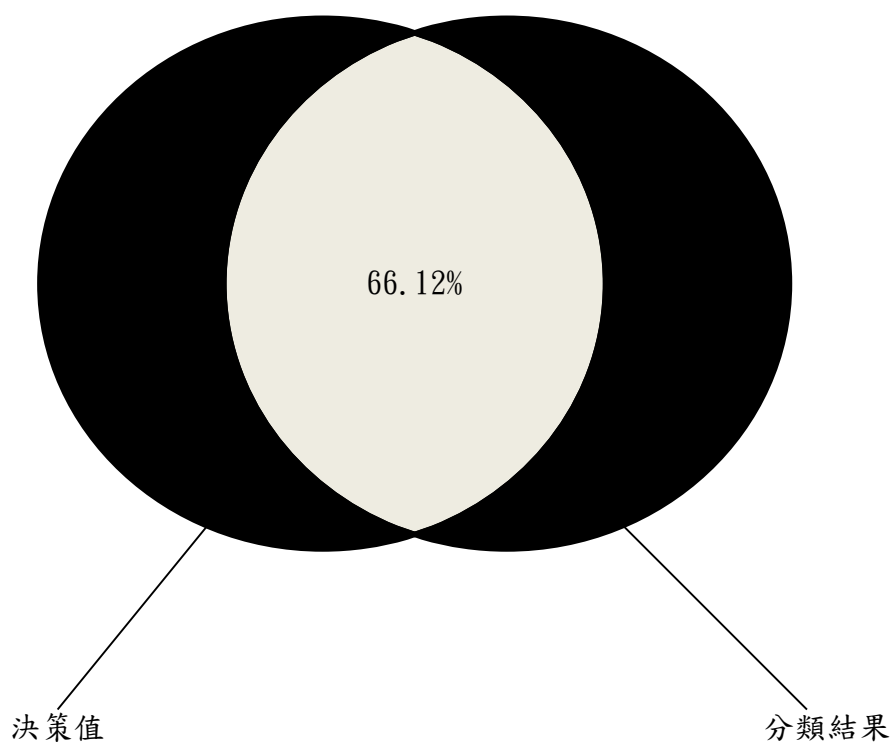


圖4.3 測試資料導入訓練資料之決策樹交集聯集圖

表 4.6 測試資料導入訓練資料之決策樹結果表

分類結果	基金編號
可投資	23、27、28、32、35、38、39、43、50、61
不可投資	1、2、3、5、6、8、9、10、11、13、14、15、16、21、29、31、34、40、41、42、44、46、47、48、51、55、57、58、59、60、62
結果不符	4、7、12、17、18、19、20、22、24、25、26、30、33、36、37、45、49、52、53、54、56

4.3.2 LDA 分析結果(案例 b)

表 4.7 測試資料導入訓練資料使用 LDA 產生之分類結果

編號	基金名稱	決策值	分類結果	結果比對
1	匯豐環球投資基金-俄羅斯股票 AD	0	1	結果不符
2	宏利環球基金-俄羅斯股票基金 AA 股	0	0	
3	摩根俄羅斯基金	0	1	結果不符
4	法巴百利達俄羅斯股票基金 C(歐元)	1	1	
5	瑞銀(盧森堡)俄羅斯股票基金(美元)	0	1	結果不符
6	霸菱俄羅斯基金-A 類美元累積型	0	0	
7	元大寶來巴西指數基金	1	0	結果不符
8	匯豐環球投資基金-巴西股票 AC	0	0	
9	摩根巴西基金	0	0	
10	法巴百利達巴西股票基金 C(美元)	0	0	
11	瑞銀(盧森堡)巴西股票基金(美元)	0	0	
12	野村巴西基金	1	1	
13	元大寶來印度指數基金	0	1	結果不符
14	元大寶來印度基金	0	1	結果不符
15	利安資金印度基金(美元)	0	0	
16	匯豐環球投資基金-印度股票 AD	0	0	
17	台新印度基金	0	0	
18	安本環球 - 印度股票基金 A2	1	0	結果不符
19	宏利環球基金-印度股票基金 AA 股	0	0	
20	富蘭克林坦伯頓全球投資系列-印度基金美元 A(acc)股	1	1	
21	富達基金-印度聚焦基金	0	0	
22	愛德蒙得洛希爾印度基金(A)	0	0	
23	摩根印度基金	1	1	
24	摩根士丹利印度股票基金 A	0	0	
25	施羅德環球基金系列 - 印度股票 A1 類股份-累積單位	0	1	結果不符
26	木星全球系列-木星印度精選基金 L 類股 美元年收益	0	1	結果不符
27	柏瑞環球基金-柏瑞印度股票基金 A	1	1	
28	法巴百利達印度股票基金 C(美元)	1	0	結果不符
29	瀚亞印度基金	0	0	
30	瀚亞投資-印度股票基金 A(美元)	0	0	
31	群益印度中小基金	0	0	
32	聯博-印度成長基金 AX 股	1	1	
33	貝萊德印度基金 A2	0	0	

34	野村雙印傘型基金之印度潛力基金	0	0	
35	首域環球傘型基金-首域印度次大陸基金	1	0	結果不符
36	匯豐環球投資基金-中國股票 AD	0	0	
37	國泰中國新興戰略基金	1	0	結果不符
38	富達基金-中國聚焦基金(歐元累積)	1	1	
39	富邦大中華成長基金	1	1	
40	德意志中國股票 FC	0	0	
41	德盛安聯中國策略增長基金	0	0	
42	摩根中國基金	0	1	結果不符
43	施羅德環球基金系列 - 中國優勢 A1 類股份 - 累積單位	1	1	
44	景順中國基金 A-年配息股 美元	0	0	
45	木星中國基金 (ACC)	0	0	
46	歐義銳榮中國股票基金 R	0	1	結果不符
47	瀚亞中國基金	0	0	
48	瀚亞投資-中國股票基金 A(美元)	0	0	
49	瑞銀(盧森堡)中國精選股票基金(美元)	0	0	
50	群益華夏盛世基金	1	0	結果不符
51	聯博-中國時機基金 A 股	0	0	
52	荷寶資本成長基金-荷寶中國股票 D 美元	1	1	
53	貝萊德中國基金 A2 美元	1	0	結果不符
54	鋒裕基金-中國股票 A2	0	1	結果不符
55	霸菱香港中國基金-A 類 歐元配息型	0	0	
56	首域環球傘型基金-首域中國核心基金-第四類股-USD	0	0	
57	匯豐環球投資基金-新興四國市場股票 AC	0	0	
58	富蘭克林坦伯頓全球投資系列-金磚四國基金美元 A (acc)股	0	1	結果不符
59	德盛金磚四國基金—AT 累積類股 (歐元)	0	0	
60	摩根新金磚五國基金	0	1	結果不符
61	施羅德環球基金系列 - 金磚四國 A 類股份 - 累積單位	1	0	結果不符
62	法巴百利達金磚四國股票基金 C (美元)	0	0	
	可投資筆數			10 筆
	不可投資筆數			32 筆
	結果不符筆數			20 筆
	資料總計			62 筆
	資料吻合度			67.74%

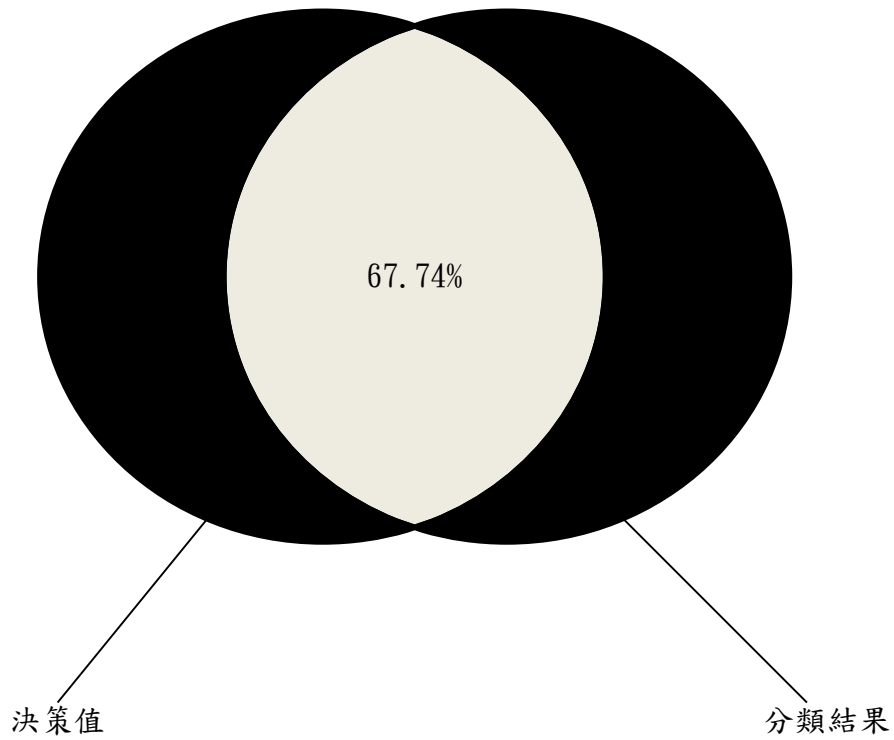


圖4.4 測試資料導入訓練資料使用LDA產生之分類交集聯集圖

表 4.8 測試資料導入訓練資料使用 LDA 產生之分類結果表

分類結果	基金編號
可投資	4、12、20、23、27、32、38、39、43、52
不可投資	2、6、8、9、10、11、15、16、17、19、21、22、24、29、30、31、33、34、36、40、41、44、45、47、48、49、51、55、56、57、59、62
結果不符	1、3、5、7、13、14、18、25、26、28、35、37、42、46、50、53、54、58、60、61

4.3.3 兩方法誤差矩陣表(案例 c)

表 4.9 決策樹誤差矩陣表

Kappa : 0.2404		實際分類			使用者精度
		可投資	不可投資	樣本點總和	
預測分類	可投資	10	13	23	43.48%
	不可投資	8	31	39	79.49%
	樣本點總和	18	44	62	
生產者精度		55.56%	70.45%		66.12%

表 4.10 線性判別分析誤差矩陣表

Kappa : 0.0634		實際分類			使用者精度
		可投資	不可投資	樣本點總和	
預測分類	可投資	39	15	54	72.22%
	不可投資	5	3	8	37.5%
	樣本點總和	44	18	62	
生產者精度		88.64%	16.67%		67.74%

4.4 大盤法

4.4 部分是以大盤法為決策值所做的分析。

4.4.1 決策樹分析結果(案例 a)

下圖為將訓練資料導入決策樹方法所畫出的決策樹圖，測試資料也需導入其中判斷，為了軟體執行順利，我們將導入資料之屬性名稱做了改變， B1 為總回報率(天)(%)、B10 為銷售額增長。

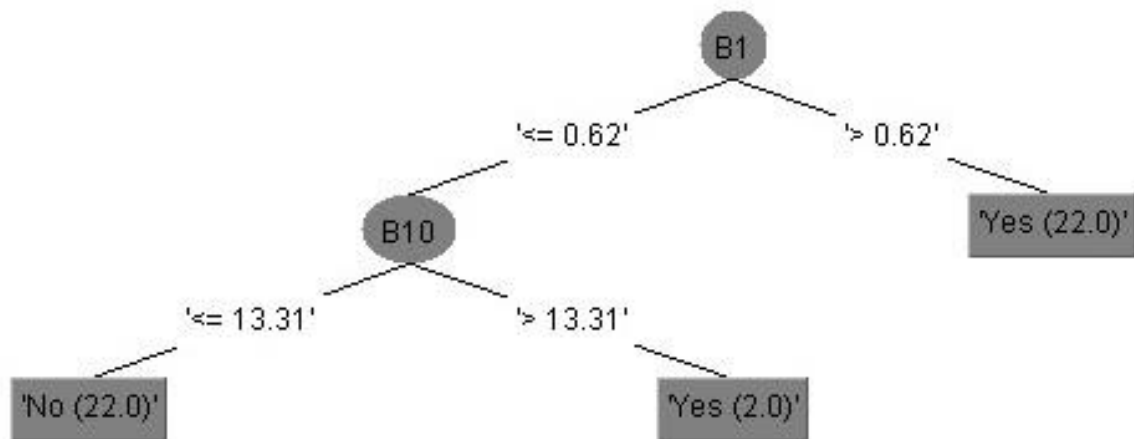


圖4.5 使用訓練資料產生決策樹圖

表 4.11 導入決策樹之訓練資料結果

編號	基金名稱	決策值	分類結果	結果比對
1	GAM Star 亞洲股票基金累積單位-歐元	1	1	
2	NN (L) 亞洲高股息基金 P 股美元	0	0	
3	元大寶來新興亞洲基金	0	0	
4	凱基亞洲四金磚基金	0	0	
5	凱基台商天下基金	0	0	
6	匯豐環球投資基金-亞洲小型公司股票(日本除外) AC	1	1	
7	匯豐環球投資基金-亞洲股票(日本除外) AD	1	1	
8	天達環球策略基金 - 亞洲股票基金 C 收益股份	1	1	
9	宏利環球基金-亞洲股票基金 AA 股	1	1	
10	富蘭克林坦伯頓全球投資系列-亞洲小型企業基金歐元 A (acc) 股	1	1	
11	富蘭克林坦伯頓全球投資系列-亞洲成長基金美元 A (acc) 股	0	0	
12	富達基金-東南亞基金	0	0	
13	富達基金-亞洲成長趨勢基金(歐元累積)	0	0	
14	富達基金-新興亞洲基金(歐元)	0	0	
15	富邦新興亞洲高成長基金	0	0	
16	德意志亞洲中小型基金 FC	1	1	
17	德盛老虎基金	1	1	
18	德盛安聯中國東協新世紀基金	0	0	
19	德盛亞洲總回報股票基金—A 配息類股 (美元)	1	1	
20	德盛新興亞洲基金	0	0	
21	摩根東方小型企業基金	0	0	
22	摩根東方內需機會基金	1	1	
23	摩根東方基金	0	0	
24	摩根亞洲內需主題基金	1	1	
25	摩根士丹利亞洲股票基金 A	1	1	
26	新加坡大華亞洲基金	0	0	
27	施羅德環球基金系列 - 亞洲小型公司 A1 類股份 - 累積單位	0	0	
28	施羅德環球基金系列 - 亞洲優勢 A1 類股份 - 累積單位	1	1	
29	施羅德環球基金系列 - 新興亞洲 A 類股份 - 累積單位	1	1	
30	普信亞洲(除日本)股票型基金 A	0	0	
31	景順亞洲機遇股票基金 A 股 美元	1	1	
32	景順亞洲動力基金 A-年配息股 美元	1	1	
33	未來資產亞洲卓越消費股票基金-A(美元)	1	1	

34	柏瑞環球基金-柏瑞亞洲(日本除外)股票基金 A	1	1	
35	歐義銳榮新興亞洲股票基金 R	1	1	
36	法儲銀富鑫新興亞洲基金 R/A EUR	0	0	
37	法巴百利達亞洲(日本除外)精選股票基金 C(歐元)	0	0	
38	瀚亞投資-亞洲股票基金 A(美元)	0	0	
39	第一金亞洲新興市場基金	1	1	
40	美盛百駿亞洲(日本除外)股票基金 A 類股歐元累積型	0	0	
41	聯博-亞洲股票基金 AD 股歐元	1	1	
42	英傑華投資-亞洲股票收益基金	0	0	
43	貝萊德亞洲巨龍基金 A2	0	0	
44	鋒裕基金-亞洲股票(不含日本)A2	1	1	
45	霸菱亞洲增長基金-A 類歐元配息型	1	1	
46	首域環球傘型基金-首域亞洲增長基金-第四類股-USD	1	1	
	可投資筆數			24 筆
	不可投資筆數			22 筆
	結果不符筆數			0 筆
	資料總計			46 筆
	資料吻合度			100%

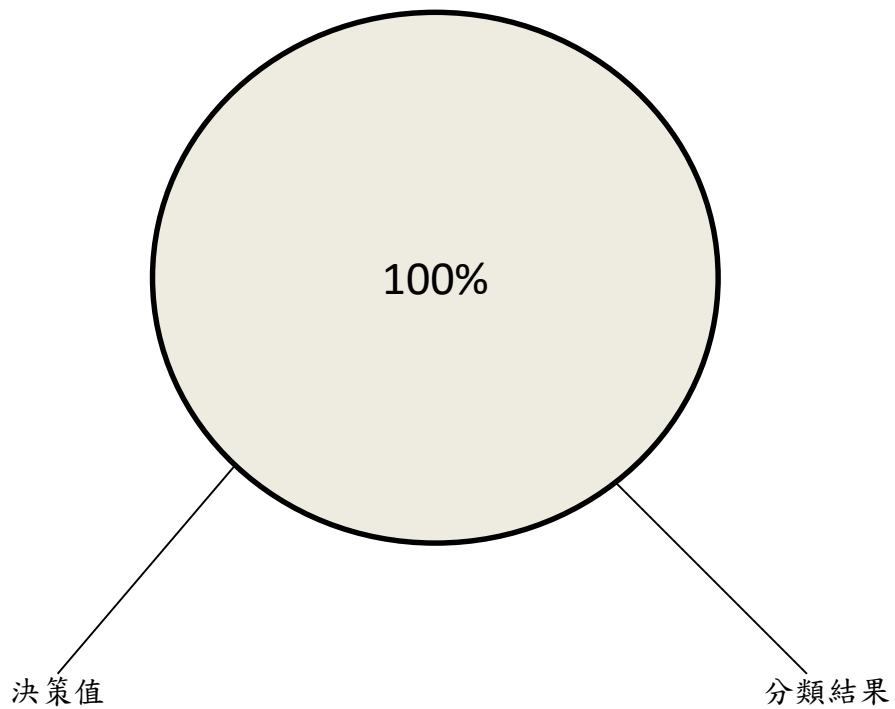


圖4.6 導入決策樹之訓練資料交集聯集圖

表 4.12 導入決策樹之訓練資料分類結果表

分類結果	基金編號
可投資	1、6、7、8、9、10、16、17、19、 22、24、25、28、29、31、32、33、 34、35、39、41、44、45、46
不可投資	2、3、4、5、11、12、13、14、15、 18、20、21、23、26、27、30、36、 37、38、40、42、43
結果不符	

表 4.13 測試資料導入訓練資料之決策樹結果

編號	基金名稱	決策值	分類結果	結果比對
1	匯豐環球投資基金-俄羅斯股票 AD	0	0	
2	宏利環球基金-俄羅斯股票基金 AA 股	0	0	
3	摩根俄羅斯基金	0	0	
4	法巴百利達俄羅斯股票基金 C(歐元)	1	1	
5	瑞銀(盧森堡)俄羅斯股票基金(美元)	1	1	
6	霸菱俄羅斯基金-A 類美元累積型	1	1	
7	元大寶來巴西指數基金	1	1	
8	匯豐環球投資基金-巴西股票 AC	1	1	
9	摩根巴西基金	0	0	
10	法巴百利達巴西股票基金 C(美元)	0	0	
11	瑞銀(盧森堡)巴西股票基金(美元)	0	1	結果不符
12	野村巴西基金	1	1	
13	元大寶來印度指數基金	0	1	結果不符
14	元大寶來印度基金	0	1	結果不符
15	利安資金印度基金(美元)	0	0	
16	匯豐環球投資基金-印度股票 AD	0	1	結果不符
17	台新印度基金	0	1	結果不符
18	安本環球 - 印度股票基金 A2	1	1	
19	宏利環球基金-印度股票基金 AA 股	1	1	
20	富蘭克林坦伯頓全球投資系列-印度基金美元 A(acc)股	0	1	結果不符
21	富達基金-印度聚焦基金	0	1	結果不符
22	愛德蒙得洛希爾印度基金(A)	1	1	
23	摩根印度基金	0	0	
24	摩根士丹利印度股票基金 A	1	1	
25	施羅德環球基金系列 - 印度股票 A1 類股份-累積單位	0	1	結果不符
26	木星全球系列-木星印度精選基金 L 類股 美元年收益	1	1	
27	柏瑞環球基金-柏瑞印度股票基金 A	0	1	結果不符
28	法巴百利達印度股票基金 C(美元)	0	1	結果不符
29	瀚亞印度基金	0	1	結果不符
30	瀚亞投資-印度股票基金 A(美元)	1	1	
31	群益印度中小基金	0	1	結果不符
32	聯博-印度成長基金 AX 股	1	1	
33	貝萊德印度基金 A2	1	1	
34	野村雙印傘型基金之印度潛力基金	0	0	
35	首域環球傘型基金-首域印度次大陸基金	1	1	

36	匯豐環球投資基金-中國股票 AD	0	1	結果不符
37	國泰中國新興戰略基金	0	1	結果不符
38	富達基金-中國聚焦基金(歐元累積)	0	0	
39	富邦大中華成長基金	0	1	結果不符
40	德意志中國股票 FC	1	1	
41	德盛安聯中國策略增長基金	0	0	
42	摩根中國基金	0	1	結果不符
43	施羅德環球基金系列 - 中國優勢 A1 類股份 - 累積單位	0	1	結果不符
44	景順中國基金 A-年配息股 美元	1	1	
45	木星中國基金 (ACC)	1	1	
46	歐義銳榮中國股票基金 R	0	0	
47	瀚亞中國基金	0	0	
48	瀚亞投資-中國股票基金 A(美元)	1	1	
49	瑞銀(盧森堡)中國精選股票基金(美元)	0	1	結果不符
50	群益華夏盛世基金	0	1	結果不符
51	聯博-中國時機基金 A 股	1	1	
52	荷寶資本成長基金-荷寶中國股票 D 美元	0	1	結果不符
53	貝萊德中國基金 A2 美元	0	1	結果不符
54	鋒裕基金-中國股票 A2	1	1	
55	霸菱香港中國基金-A 類 歐元配息型	1	1	
56	首域環球傘型基金-首域中國核心基金-第四類股-USD	0	1	結果不符
57	匯豐環球投資基金-新興四國市場股票 AC	0	1	結果不符
58	富蘭克林坦伯頓全球投資系列-金磚四國基金美元 A (acc)股	1	1	
59	德盛金磚四國基金—AT 累積類股 (歐元)	0	0	
60	摩根新金磚五國基金	0	0	
61	施羅德環球基金系列 - 金磚四國 A 類股份 - 累積單位	1	1	
62	法巴百利達金磚四國股票基金 C (美元)	0	0	
	可投資筆數			24 筆
	不可投資筆數			15 筆
	結果不符筆數			23 筆
	資料總計			62 筆
	資料吻合度			62.9%

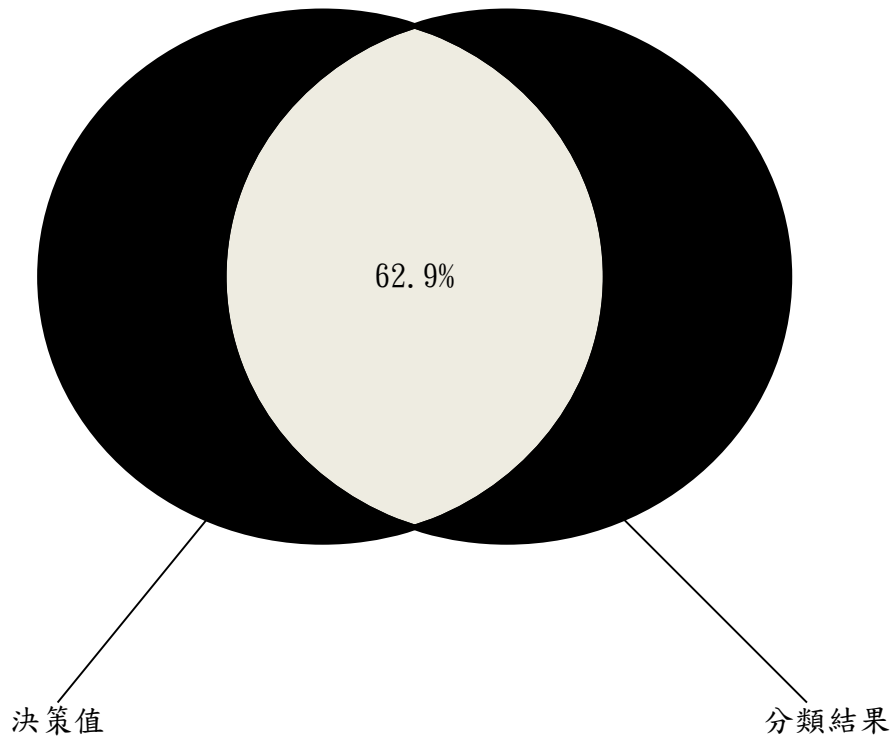


圖4.7 測試資料導入訓練資料之決策樹交集聯集圖

表 4.14 測試資料導入訓練資料之決策樹結果表

分類結果	基金編號
可投資	4、5、6、7、8、12、18、19、22、 24、26、30、32、33、35、40、44、 45、48、51、54、55、58、61
不可投資	1、2、3、9、10、15、23、34、38、 41、46、47、59、60、62
結果不符	11、13、14、16、17、20、21、25、 27、28、29、31、36、37、39、42、 43、49、50、52、53、56、57

4.4.2 LDA 分析結果(案例 b)

表 4.15 測試資料導入訓練資料使用 LDA 產生之分類器結果

編號	基金名稱	決策值	分類結果	結果比對
1	匯豐環球投資基金-俄羅斯股票 AD	0	0	
2	宏利環球基金-俄羅斯股票基金 AA 股	0	0	
3	摩根俄羅斯基金	0	0	
4	法巴百利達俄羅斯股票基金 C(歐元)	1	1	
5	瑞銀(盧森堡)俄羅斯股票基金(美元)	1	0	結果不符
6	霸菱俄羅斯基金-A 類美元累積型	1	1	
7	元大寶來巴西指數基金	1	1	
8	匯豐環球投資基金-巴西股票 AC	1	0	結果不符
9	摩根巴西基金	0	0	
10	法巴百利達巴西股票基金 C(美元)	0	0	
11	瑞銀(盧森堡)巴西股票基金(美元)	0	0	
12	野村巴西基金	1	1	
13	元大寶來印度指數基金	0	1	結果不符
14	元大寶來印度基金	0	0	
15	利安資金印度基金(美元)	0	0	
16	匯豐環球投資基金-印度股票 AD	0	0	
17	台新印度基金	0	1	結果不符
18	安本環球 - 印度股票基金 A2	1	0	結果不符
19	宏利環球基金-印度股票基金 AA 股	1	1	
20	富蘭克林坦伯頓全球投資系列-印度基金美元 A(acc)股	0	0	
21	富達基金-印度聚焦基金	0	1	結果不符
22	愛德蒙得洛希爾印度基金(A)	1	0	結果不符
23	摩根印度基金	0	1	結果不符
24	摩根士丹利印度股票基金 A	1	0	結果不符
25	施羅德環球基金系列 - 印度股票 A1 類股份-累積單位	0	0	
26	木星全球系列-木星印度精選基金 L 類股 美元年收益	1	1	
27	柏瑞環球基金-柏瑞印度股票基金 A	0	0	
28	法巴百利達印度股票基金 C(美元)	0	0	
29	瀚亞印度基金	0	0	
30	瀚亞投資-印度股票基金 A(美元)	1	1	
31	群益印度中小基金	0	1	結果不符
32	聯博-印度成長基金 AX 股	1	1	
33	貝萊德印度基金 A2	1	1	

34	野村雙印傘型基金之印度潛力基金	0	0	
35	首域環球傘型基金-首域印度次大陸基金	1	1	
36	匯豐環球投資基金-中國股票 AD	0	1	結果不符
37	國泰中國新興戰略基金	0	0	
38	富達基金-中國聚焦基金(歐元累積)	0	0	
39	富邦大中華成長基金	0	0	
40	德意志中國股票 FC	1	0	結果不符
41	德盛安聯中國策略增長基金	0	0	
42	摩根中國基金	0	0	
43	施羅德環球基金系列 - 中國優勢 A1 類股份 - 累積單位	0	1	結果不符
44	景順中國基金 A-年配息股 美元	1	1	
45	木星中國基金 (ACC)	1	0	結果不符
46	歐義銳榮中國股票基金 R	0	0	
47	瀚亞中國基金	0	0	
48	瀚亞投資-中國股票基金 A(美元)	1	1	
49	瑞銀(盧森堡)中國精選股票基金(美元)	0	1	結果不符
50	群益華夏盛世基金	0	0	
51	聯博-中國時機基金 A 股	1	1	
52	荷寶資本成長基金-荷寶中國股票 D 美元	0	0	
53	貝萊德中國基金 A2 美元	0	0	
54	鋒裕基金-中國股票 A2	1	1	
55	霸菱香港中國基金-A 類 歐元配息型	1	1	
56	首域環球傘型基金-首域中國核心基金-第四類股-USD	0	1	結果不符
57	匯豐環球投資基金-新興四國市場股票 AC	0	0	
58	富蘭克林坦伯頓全球投資系列-金磚四國基金美元 A (acc)股	1	1	
59	德盛金磚四國基金—AT 累積類股 (歐元)	0	1	結果不符
60	摩根新金磚五國基金	0	0	
61	施羅德環球基金系列 - 金磚四國 A 類股份 - 累積單位	1	1	
62	法巴百利達金磚四國股票基金 C (美元)	0	0	
	可投資筆數			17 筆
	不可投資筆數			28 筆
	結果不符筆數			17 筆
	資料總計			62 筆
	資料吻合度			72.58%

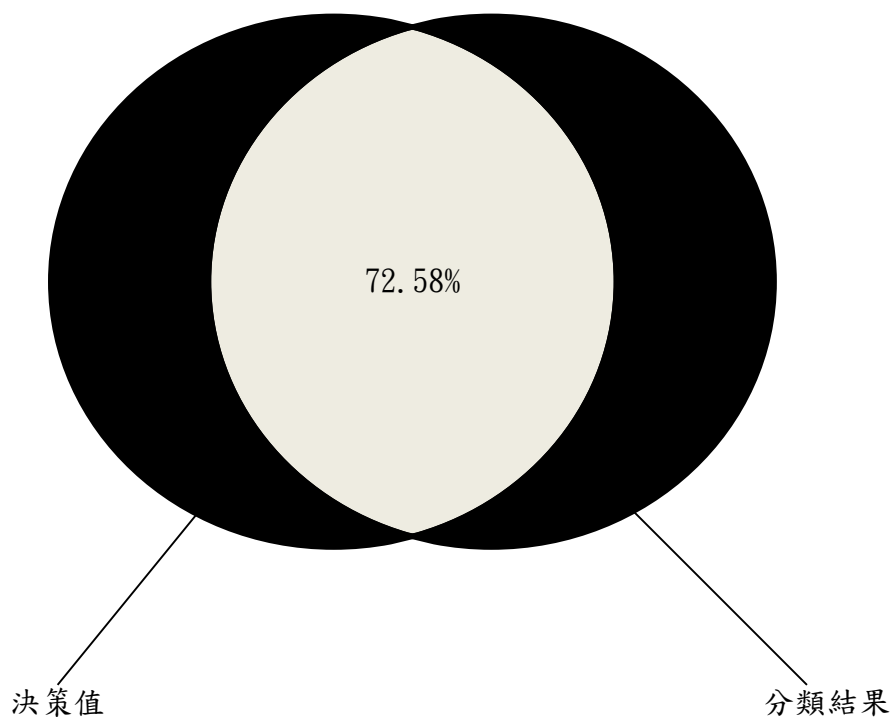


圖4.8 測試資料導入訓練資料使用LDA產生之分類交集聯集圖

表 4.16 測試資料導入訓練資料使用 LDA 產生之分類結果表

分類結果	基金編號
可投資	4、6、7、12、19、26、30、32、 33、35、44、48、51、54、55、58、 61
不可投資	1、2、3、9、10、11、14、15、16、 20、25、27、28、29、34、37、38、 39、41、42、46、47、50、52、53、 57、60、62
結果不符	5、8、13、17、18、21、22、23、 24、31、36、40、43、45、49、56、 59

4.4.3 兩方法誤差矩陣表(案例 c)

表 4.17 決策樹誤差矩陣表

Kappa : 0.3355		實際分類			使用者精度
		可投資	不可投資	樣本點總和	
預測分類	可投資	24	23	47	51.06%
	不可投資	0	15	15	100%
	樣本點總和	24	38	62	
生產者精度		100%	39.47%		62.9%

表 4.18 線性判別分析誤差矩陣表

Kappa : 0.3470		實際分類			使用者精度
		可投資	不可投資	樣本點總和	
預測分類	可投資	37	16	53	69.81%
	不可投資	1	8	9	88.89%
	樣本點總和	38	24	62	
生產者精度		97.37%	33.33%		72.58%

最後我們將以晨星評級為決策值和以大盤法為決策值在兩方法分類模擬之結果整理成綜合結果表。

表 4.19 綜合結果表

決策值來源	晨星評級		大盤法	
	決策樹	LDA	決策樹	LDA
Kappa 值	0.2404	0.0634	0.3355	0.3470
整體精度	66.12%	67.74%	62.9%	72.58%

第五章 結論

本研究中所收集的各種屬性資料、決策資料，都是隨手可得的，簡而言之經過研究之後，民眾可以利用其中的工具與方式來當成晨星評級以及大盤法以外的投資參考工具。

由案例a中經決策樹分析後，判斷基金為可投資或不可投資，重要的財務屬性有總回報率(天)、股價對現金流、銷售額增長、Alpha指數；然而使用大盤法為決策值經過決策樹分析出有62.9%、LDA分析出有72.58%的基金具有績效持續性，意指過去的歷史績效可預測未來績效。

由表4.19 綜合結果表中可以得知，以大盤法為決策值使用兩方法所模擬出之分類器Kappa值略高於使用晨星評級為決策值之Kappa值，意指大盤法為決策值所模擬出之分類器可信度高於晨星評級。

由於本研究中資料量較小而資料屬性較多，在這樣的條件下決策樹容易產生計算誤差，因此LDA是較客觀的分類器。在表4.19 綜合結果表中，LDA分類器的整體精度大盤法略高於晨星評級，因此可以由這個結果得知晨星評級尚有改善空間。

參考文獻

- [1]維基百科(金磚四國簡介) ,
<http://zh.wikipedia.org/wiki/%E9%87%91%E7%A0%96%E5%9B%9B%E5%9B%B>
[D](#)
- [2]梁碩修,「晨星評級與基金績效:基金評級是否有助於預測未來績效?」,逢甲大學金融碩士在職專班碩士論文,民國一百零二年。
- [3]摩根資產管理,
https://www.jpmmrich.com.tw/wps/portal/!ut/p/b0/04_Sj9CPykssy0xPLMnMz0vMAfGjzOK9AkIDjEJcDQz83XycDIxczIyd3TzcDAwsDfULsh0VAQkMEU4!/?WCM_GLOBAL_CONTEXT=/wps/wcm/connect/JPMRich/eportal/LB_B2C_L02/B2C_L02P400/B2C_L02P420/B2C_L02P420_01/B2C_L02P420_01_01/B2C_L02P420_01_002
- [4]簡好倫,「國內股票型基金篩選法則之研究」,國立臺北大學國際金融碩士在職專班碩士論文,民國九十七年。
- [5]林聲宇,「以應用決策樹與濾嘴法則於股票投資」,國立交通大學工業工程與管理學系碩士論文,民國九十四年。
- [6]曾永富、黃佩嫻、黃玄育,「以優化式線性判別分析股票篩選以光電股票為例」,嶺東科技大學資訊管理系大學部專題,民國一百零三年。
- [7]王秋萍,「區域型新興市場股票基金績效持續性」,國立中正大學企業管理研究所碩士論文,民國九十八年。
- [8]蔡毓翔,「台灣股票型共同基金績效之研究」,淡江大學財務金融研究所碩士論文,民國九十二年。
- [9]晨星,
<http://tw.morningstar.com/ap/FundTools/help/glossary.aspx?tab=portfolio&ismpf=0>
- [10]Our Family (WEKA), <https://sites.google.com/site/yzuimprojectzi/shi-xi-zhu-ti-jian-jie/weka>
- [11]謝文瑋,「台灣基金績效評比之驗證」,東吳大學經濟學系碩士論文,民國一百零一年。
- [12]張有若,「全球共同基金群組風險與績效評估」,中原大學企業管理學系碩士學位論文,民國九十一年。
- [13]曾麗雪,「不同投資區域之共同基金績效評估」,國立雲林科技大學財務金融系碩士論文,民國一百零三年。
- [14]劉玟君,「探討共同基金之最適操作策略」,國立暨南國際大學國際企業研究所碩士論文,民國九十五年。

[15]陳秉洋，「結合技術指標、籌碼變動及期貨基差建構整合性投資策略－以台灣 50 指數型基金為例」，國立成功大學統計學研究所碩士論文，民國一零一年。

附錄

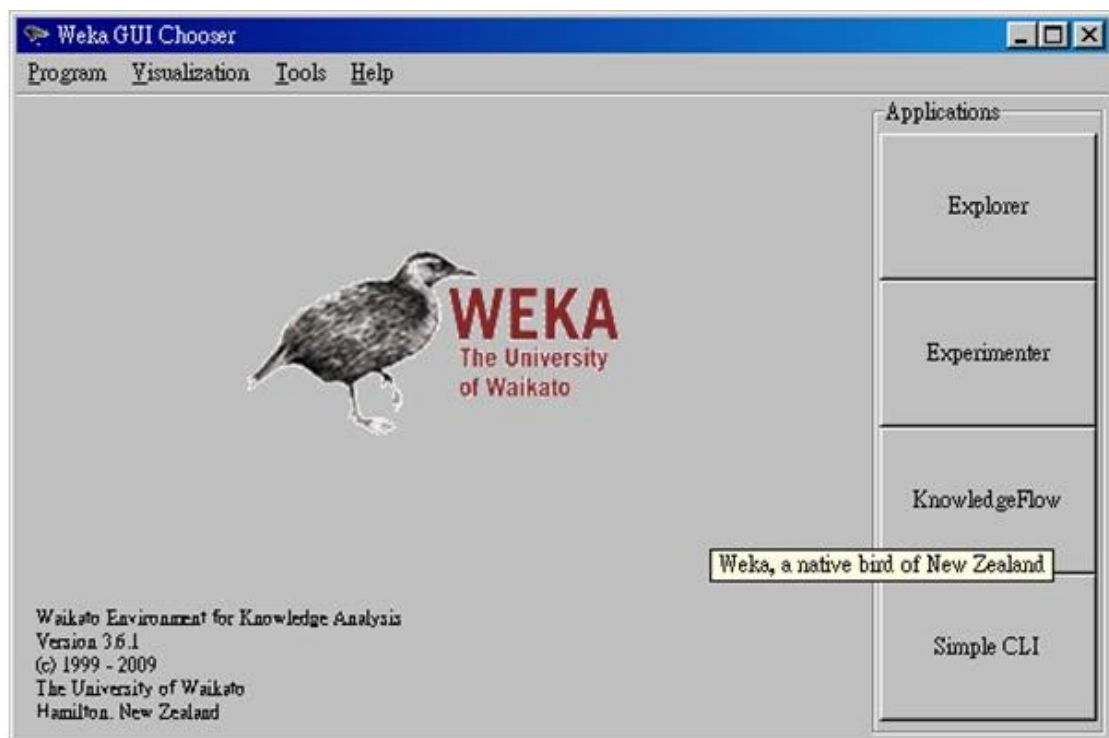
附錄 A WEKA

WEKA[10]的全名是懷卡托智慧分析環境(Waikato Environment for Knowledge Analysis)，是一款免費的，非商業化(與之對應的是 SPSS 公司商業資料採擷產品 -Clementine)的，基於 JAVA 環境下開源的機器學習(Machine learning)以及資料採擷(Data Mining)軟體。它和它的原始程式碼可在其官方網站下載。

該軟體的縮寫 WEKA 也是紐西蘭獨有的一種鳥名，而 WEKA 的主要開發者同時恰好來自紐西蘭的 the University of Waikato。WEKA 作為一個公開的資料採擷工作平台，集合了大量能承擔資料採擷任務的機器學習演算法，包括對資料進行預處理、分類、回歸、聚類、關聯規則以及在新的互動式介面上的視覺化。

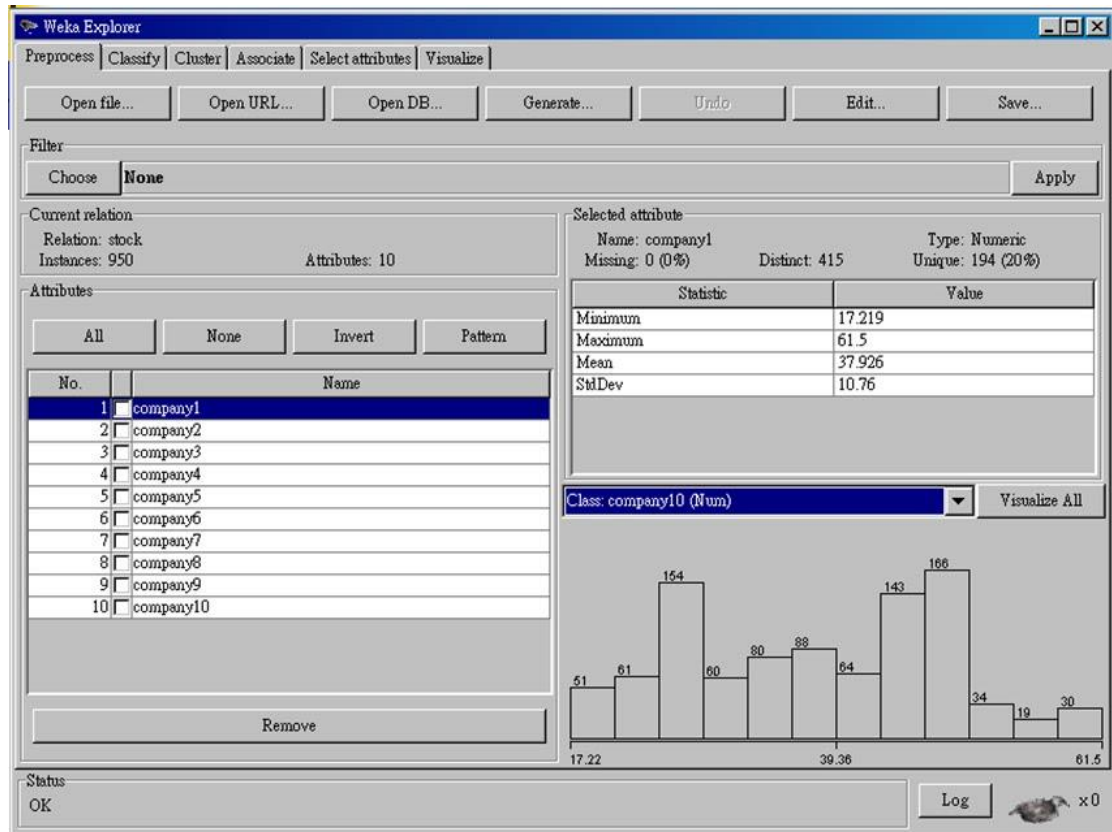
而開發者則可使用 Java 語言，利用 WEKA 的架構上開發出更多的資料採擷演算法，可以通過 Java 編程和命令列來調用其分析組件。同時，WEKA 也為普通用戶提供了圖形化介面，稱為 WEKA Knowledge Flow Environment 和 WEKA Explorer。在 WEKA 中集成自己的演算法甚至借鑒它的方法自己實現視覺化工具並不是件很困難的事情，是目前 Data Mining 非常廣泛使用的工具。基本操作：

點選 WEKA 圖示之後，便會出現一個小型的 GUI,提供了四種操作環境：CLI, Explorer, Experimenter, Knowledge Flow。(如圖 3-3 所示)



圖A.1 進入WEKA之介面

首先介紹 Explorer 介面(如圖 3-4 所示)，點選之後在窗口的頂部，標題列下是一排標籤。當 Explorer 首次啟動時，只有第一個標籤頁是活動的；其它均是灰色的。這是因為在探索資料之前，必須先打開一個資料集(可能還要對它進行預處理)。



圖A.2 Explorer介面

所有的標籤頁如下所示：

1. Preprocess 選擇和修改要處理的資料。
2. Classify 訓練和測試關於分類或回歸的學習方案。
3. Cluster 從數據中學習聚類。
4. Associate 從資料中學習關聯規則。
5. Select attributes 選擇資料中最相關的屬性。
6. Visualize 查看資料的互動式二維圖像。這些標籤被啟動後，點擊它們可以在不同的標籤頁面上進行切換，而每一個頁面上可以執行對應的操作。不管位於哪個頁面，視窗的底部區域(包括狀態列、log 按鈕和 Weka 鳥) 仍然可見。

Log 按鈕：

點擊這個按鈕，會出現一個單獨的視窗，包含一個可拖動的區域。文本的每一行被加了一個時間標記，顯示了它進入日誌的時間，一旦在 WEKA 中執

行操作時，該日誌就會記錄發生了什麼。

WEKA 圖示：

狀態列的右邊是 WEKA 狀態圖示。當不運行任何進程時，WEKA 鳥會坐下並打一個小盹。X 符號旁的數字顯示了正運行的併發進程的數量。當系統閒時，它是零，而當進程的數量增長時，它也會增長。任意進程啟動後，小鳥會站起來並到處活動。如果它仍然是站著的，但是很長時間內不動，那麼它生病了：某個地方出錯了！在這種情況下，應該重新啟動 WEKA Explorer。

Explorer 資料預處理：

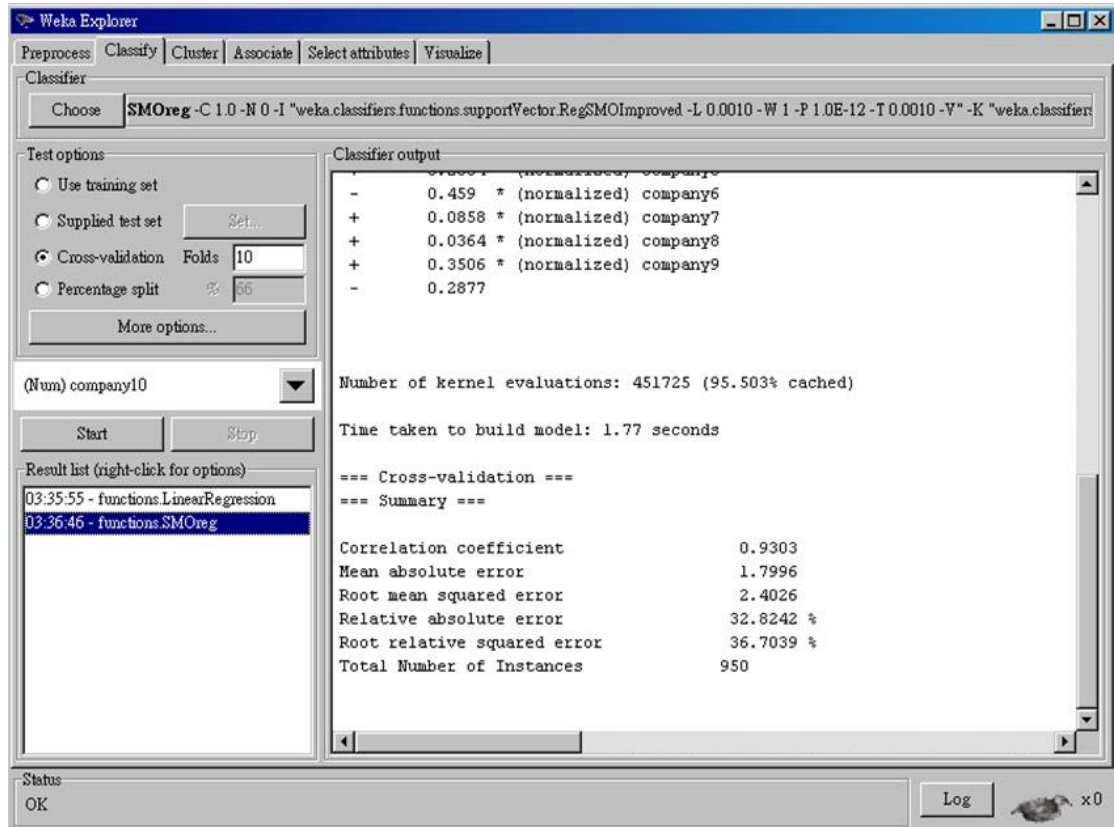
使用 Preprocess 功能選擇和修改要處理的資料，載入的資料可以使用下面四種格式：

1. 檔案：可以直接重本端所擁有的檔案直接載入。
2. URL：可以輸入資料所存在的 URL 直接載入。
3. DB：可使載入事先儲存在資料庫的資料來用。
4. Generate：資料可以由資料產生器人工產生。

再輸入資料後，右邊即會出現資料的分布圖及其基本的統計量，可以利用 Remove 按鈕移除其此次實驗不需要的屬性，而 WEKA 本身有提供很多的 filter，可以讓使用者對載入的資料做預先的處理，例如將資料離散化...等等的功能。

Explorer 資料分類(預測)：

而我們實驗主要使用的是 WEKA 的 Classify 功能對於我們的資料進行訓練及相關預測，WEKA 本身有內建一些大家常用的分類器，例如 NaiveBayes、LibSVM、Adaboost...等等。除此之外我們還能夠設定一些參數，例如我們要使用的 training set 是不是全部的範圍，要不要把測驗的範圍切割成多份，一些當 training set 一些當 testing set。或者是可以取其中的百分之幾當作要 training 的資料，剩下的拿來當 testing。此外每次跑完一次後，軟體會顯示出相關分類的結果，以及訓練好分類的那個 model，還有詳細的結論跟一些準確度的分析。



圖A.3 Knowledge Flow 介面

應用選定的分類器後得到的結果會根據 Test Option 一欄中的選擇來進行測試。共有四種測試模式：

1. Using training set：根據分類器在用來訓練的實例上的預測效果來評價它。
2. Supplied test set：從檔案載入的一組實例，根據分類器在這組實例上的預測效果來評價它。點擊 Set... 按鈕將打開一個對話方塊來選擇用來測試的檔。
3. Cross-validation：使用交叉驗證來評價分類器，所用的折數填在 Folds 文字方塊中。
4. Percentage split：從資料集中按一定百分比取出部分資料放在一邊作測試用，根據分類器這些實例上預測效果來評價它。取出的資料量由 % 一欄中的值決定。

Knowledge Flow 是 WEKA 這個軟體一個很特殊的地方，它是一個可以讓使用者方便快捷建立所想要的程式的方式，可以藉由直接拖拉上面所擁有的每個原件，每一個原件代表者一項功能，將不同的原件組合起來，然後依造順序，可以將使用者的資料依序經過訓練然後產生結果。這個環境本質上和 Explorer 所支持的功能是一樣的，而它的優勢，就是支持增量學習(incremental learning)。

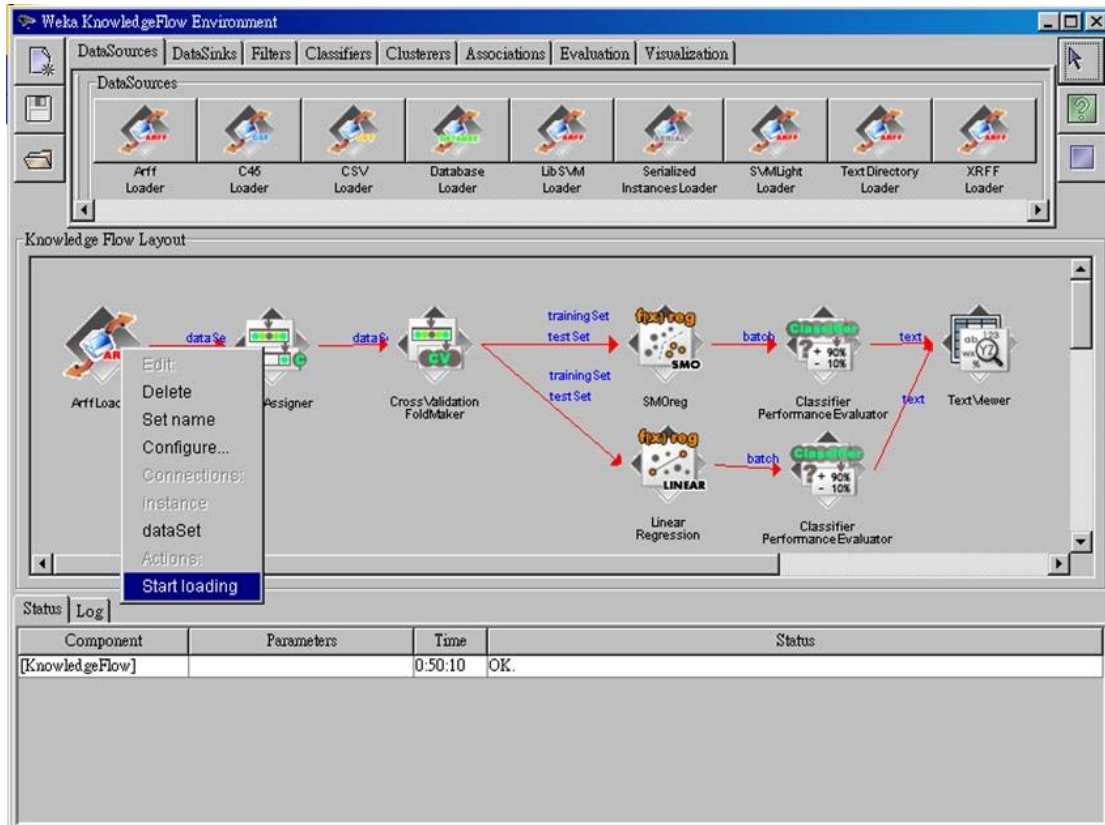


圖 A.4 KF

Experimenter 介面：

運行演算法試驗、管理演算法方案之間的統計檢驗的環境使用 experimenter 我們也可以去建立、執行、修改、分析我們之前所跑出來的結果，根據不同的結果我們可以馬上修改互相比較。此外，他還可幫我們將計算出來的結果存成.arff 的檔案格式。

Simple CLI 介面：

使用者除了使用視窗的介面外，WEKA 還另外提供了一個簡單的 command line 的輸入方式，使用者可以直接透過下命令的方式去執行其中的操作，而不一定要使用 command line 的方式。提供了一個簡單的命令列介面，從而可以在沒有自帶命令列的作業系統中直接執行 WEKA 命令。

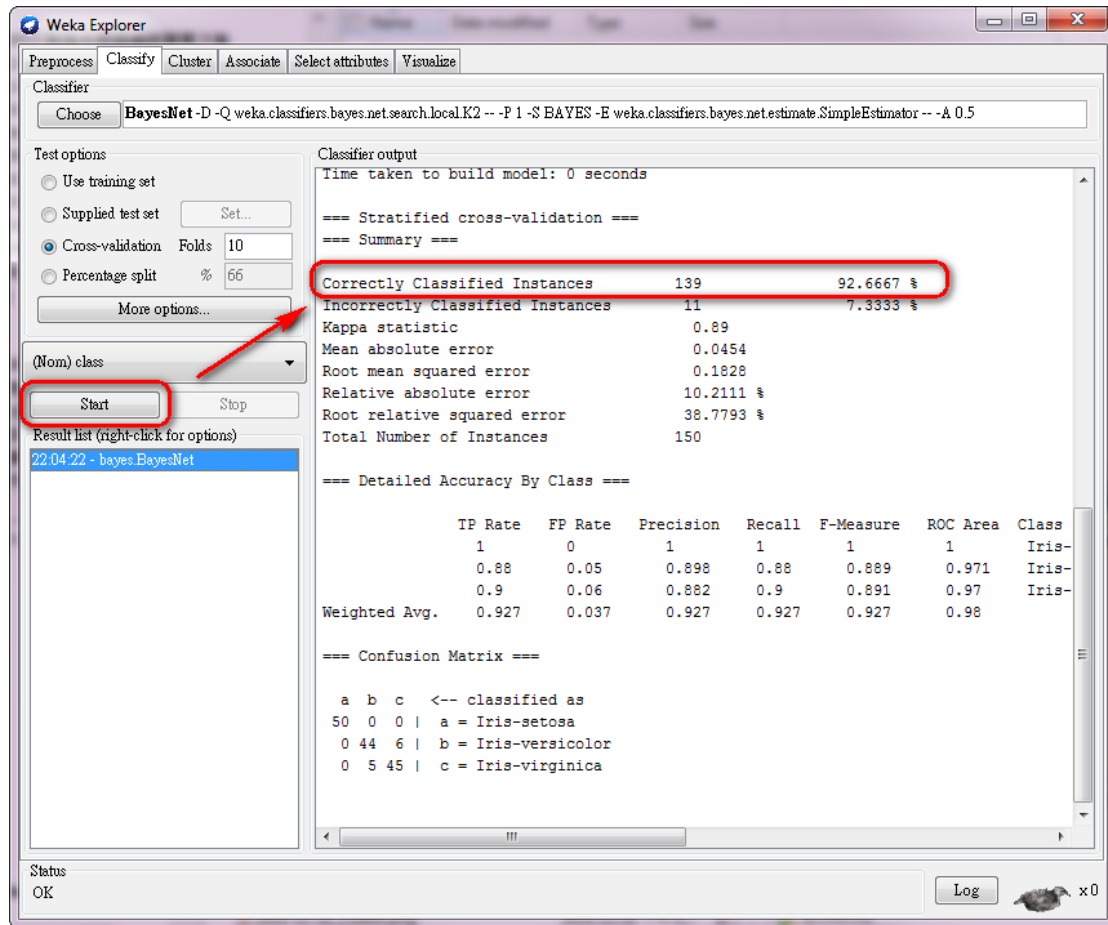
交叉驗證：

有時亦稱循環估計，是一種統計學上將數據樣本切割成較小子集的實用方法。於是可以先在一個子集上做分析，而其它子集則用來做後續對此分析的確認及驗證。一開始的子集被稱為訓練集。而其它的子集則被稱為驗證集或測試集。

K 次交叉驗證(K-fold cross-validation)，初始採樣分割成 K 個子樣本，一個單獨的子樣本被保留作為驗證模型的數據，其他 K-1 個樣本用來訓練。交叉驗證重複 K 次，每個子樣本驗證一次，平均 K 次的結果得到一個單一估測。這個

方法的優勢在於，同時重複運用隨機產生的子樣本進行訓練和驗證，每次的結果驗證一次，10次交叉驗證是最常用的。

預設的 BayesNet 探勘結果 / BayesNet Mining Result with Default Options

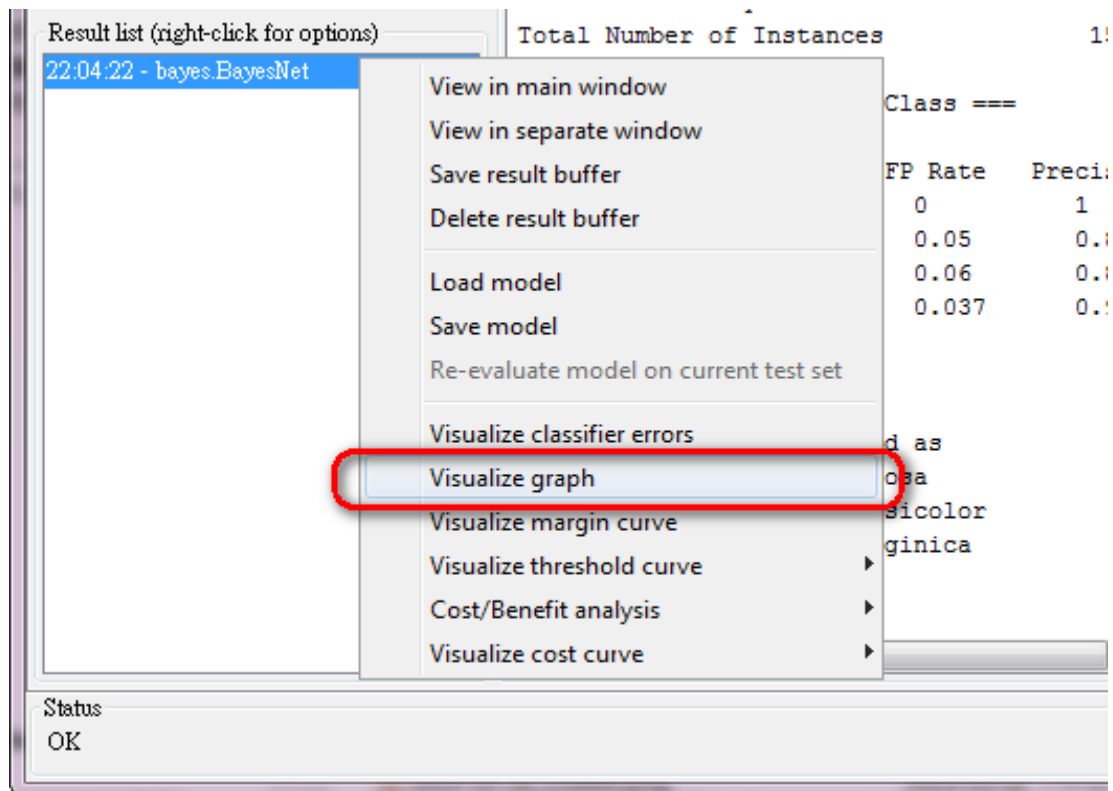


圖A.5 探勘結果

測試方式用預設的交互驗證(Cross-validation) 10層，然後分類的目標是 (Nom) play，按下「Start」。最後結果看可以看到其中一行：

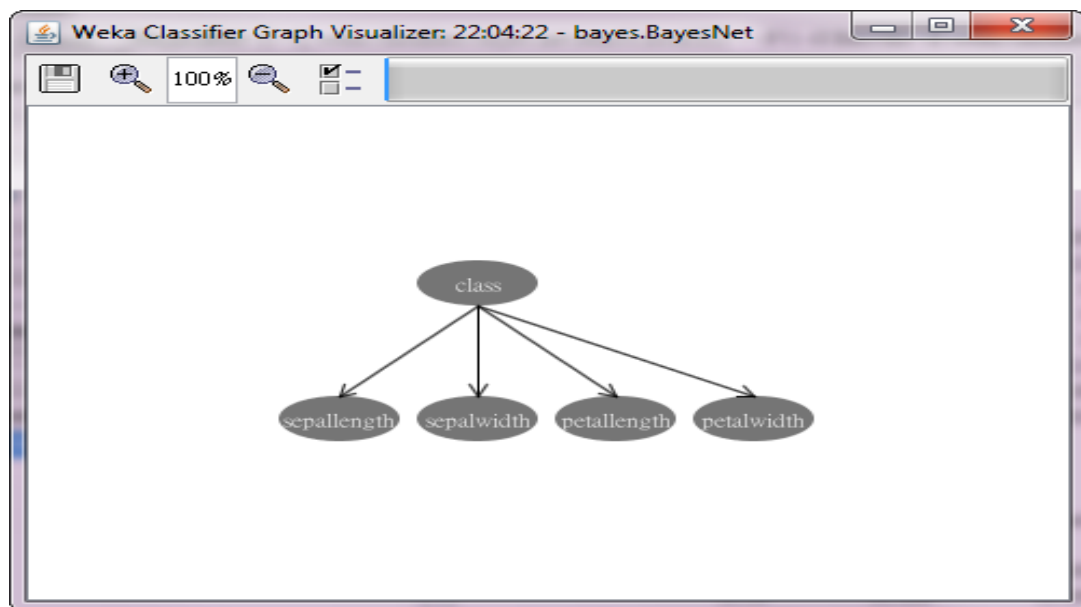
Correctly Classified Instances 139 92.6667%

這意思是，預設的 BayesNet 分類法下來，正確率只有 92% 左右。



圖A.6 BayesNet分類法

我們在結果集上按下右鍵，選擇「Visualize graph」。



圖A.7決策樹成品圖

以上是基本的操作。

附錄 B 誤差矩陣說明

投資理財已然成為現代人非常重要的功課，在許多網站也能找到投資理財工具的相關資料，在本研究的研究方法中均假設方法的分析處理為百分之百準確，此假設會造成分類準確度評估的誤差產生。因此，參考資料的真實性和方法評估的準確度，對於使用者或研究者來說，是具有一定程度重要性的因素。方法分類之精準度評估為方法分類完成後，依據所蒐集的真實情況參考資料，評估分類後方法的精準度。是以每一基礎單位為一檔基金表示，再利用隨機方式抽取檢核的檢核點後，逐一的檢查每一個檢查點，可產生一統計表，稱為誤差矩陣。

在誤差矩陣中，能產生使用者精度(user accuracy)、生產者精度(producer accuracy)、整體精度(overall accuracy)及Kappa統計值。以下就本研究所採用之精準度評估加以介紹。

1.生產者精度(producer accuracy)：

生產者精度，是指屬於某一真實資料的屬性資料之檢核點，有部份的檢核點被錯誤分類，而被正確分類的基金檔數之百分比；亦即將誤差矩陣中對同一種分類，對角線的數值，除以行中所有數值相加之和，所產生的百分位數。

$$\text{生產者精度} = \frac{X_{ii}}{\sum X_{i+}} \times 100\%$$

2.使用者精度(user accuracy)：

使用者精度是指分類後的各決策值預測，對應到真實資料決策值時，真正為該種類別的基金檔數之百分比；由下列公式表示，是依據每一種分類，將對角線的數值，除以列所有數值相加之和，所產生的百分位數。

$$\text{使用者精度} = \frac{X_{ii}}{\sum X_{i+}} \times 100\%$$

3.整體精度(overall accuracy)：

整體精度代表分類後正確的決策值分類的檢核點數除於總基金檔數的檢核點數，所產生的百分比；由下列公式表示，是在誤差矩陣中對角線的所有數值加總除以全部樣本的總和，所產生的百分位數。

$$\text{整體精度} = \frac{\sum_{i=1}^n X_{ii}}{\sum_{i=1}^n \sum_{j=1}^n X_{ij}} \times 100\%$$

4.Kappa統計值：

為了更能表現整體分類的誤差性，可藉由誤差矩陣間的相互運算而產生之統計值，它同時考慮漏判與誤判等因素，表示分類成果比隨機分類優良多少的指標。

$$\text{Kappa 統計值} = \frac{N \sum_{i=1}^n X_{ii} - \sum_{i=1}^n (X_{i+} \cdot X_{+i})}{N^2 - \sum_{i=1}^n (X_{i+} \cdot X_{+i})}$$

