

遺傳演算法在發展股市投資專家知識規則之研究*

邱昭彰 李安邦
元智大學資訊管理系學

摘要

一般而言證券市場投資者常運用技術指標來作為買進賣出的依據，然而技術指標常存在著一些問題，比如指標的選擇、鈍化現象、相似性或互為矛盾性等限制；除非是專業人士，針對報紙證券版上每日證券行情所提供的常用技術指標。一般大眾誠難從中窺其堂奧，進而有效利用這些資訊作為投資決策上之參考。人工智慧技術中之遺傳演算法是採用字串演算方式。在操作過程中，可以適當編碼與字串定義來表示投資策略，並尋求最適買入點與賣出時機等。本研究在考量實用性之前提下，試圖應用遺傳演算法極佳化的搜尋能力及其易於發展交易模式的特性，並藉諸於常用之技術指標作為導引變數，從歷史資料中找出最適投資策略。

關鍵字：人工智慧、遺傳演算法、技術指標、投資決策

壹、緒論

從美國道瓊公司將台灣股市納入道瓊世界股票指數（DJWSI），摩根史坦利公司將台股納入摩根史坦利資本國際（Morgan Stanley Capital International, MSCI）新興市場自由指數（Emerging Markets Free, EMF），及芝加哥商業交易所（Chicago Mercantile Exchange, CME）與新加坡國際金融期貨交易所（Singapore Interbank Monetary Exchange, SIMEX）將台股指數期貨合約納入期貨市場上的趨勢來看，國內股市已無可避免的受到國際金融與經濟的影響。因此，在國際化、自由化與提昇競爭力的同時，商機的掌握將

是促成經濟穩定成長原動力之主要關鍵因素。所以當企業或個人在面對金融市場國際化與自由化的競爭時，思考如何運用基本面與技術面分析來分散風險並提高投資報酬率，以有效的決策來掌握投資契機，已成為學術界與實務界所關心的重要議題。

證券價格常受人為因素、政治因素、經濟因素、競爭條件、及突發事件等多種因素影響。縱觀股價的研究分析如基本面分析、技術分析、或心理面分析等，無不試圖從股價交易的行為中探索其變化規則。事實上證券交易在諸分析中，時常出現某些可資買賣之參考依據，然而由於證券市場是一開放及

* 接受本篇論文的評審委員： 陳文華

動態的環境，任何規則或交易經驗都會隨著時間、環境變動而產生誤差。因此，如何建構一操作機制，能隨時間及環境變動等因素來動態調整其交易規則，並據此指導進出股市的最適時機，便愈顯其重要性。

學術界在投資學方面研究一直未曾間斷，自 Markowitz (1952) 建立其平均-變異數 (mean-variance criteria) 投資組合理論，奠定投資組合理論基礎之後，陸續亦有不同修正模型提出，如平均數-吉尼係數準則 (Mean-Gini Criteria)、隨機優勢準則 (stochastic dominance criteria) 等 (Levy, 1992)。這些模式經常需要用大量數學運算從許多限制條件中來找出最適 (optimal) 的解。一些最適化技術如線性規劃 (linear programming)、動態規劃 (dynamic programming)、非線性規劃 (nonlinear programming) 及最適控制理論 (optimal control theory) 等應用均是以定量方式來解決最適化的問題。然而，這些技術需針對不同的問題來使用不同的方法，也各有其優缺點 (Yeh, 1996)。從定量 (quantitative) 觀點來看，最適化問題常透過數學公式以尋求解答，然而交易市場中資料變化多端，若欲從中找出其變化規則，傳統的計量技術在面對問題本身的線性化、條件限制、記憶體空間及運算時間等問題下尋求最適解實非易事 (Tayler, 1995)。此外，財務相關理論的應用偏重於學術領域的研究，對於一般大眾投資人的實際交易行為亦難有實質上的指引。鑒於這些

理論模式實用性的限制，本研究乃從定性 (qualitative) 的方式，試圖從市場交易資料中發掘一些能協助買賣的知識規則，而所採用的遺傳演算法主要在於萃取此交易的知識規則。

遺傳演算法 (Genetic Algorithms, GAs) 係 Holland (1975) 根據自然界生物系統演化過程所開發出來的，其目的在於設計一「計算系統 (computing system)」從模擬自然界「適者生存、不適者淘汰」的機制中來具體找出最適解。由於 GA 速度快、可靠性高、具彈性、及容易作為模擬與模式的介面，已成為近年來最流行的演化計算方法 (Goldberg, 1994)。近年來 GA 在財務上的實際應用日漸受重視，無論其應用領域為線性或非線性問題，由於它具有在一個多變數函數的數學空間中搜尋最適解的能力，例如預測非線性時間序列以輔助買賣股票的決策 (Davis, 1994)、建立動態證券市場技術模型 (金必煌, 1995)、及發展財務模型以預測每股盈餘 (范饒耀, 1996) 上已有不錯的成果，而這些問題解決模式是以定量方式為基礎。然而，GA 也非常適合發展法則與交易的策略 (Tayler, 1995)，例如在協助識別領先指標並發展交易法則及策略 (Colin, 1994)、建立金融機構的客戶債信評估標準及債券市場最佳交易之決策支援系統 (Bauer, 1994) 等均顯示 GA 在法則應用上之適用性。不過，上述學者應用範圍大多針對公司財務債信等基本資料作預測或評比，對於資料不易取得或分

析的一般投資大眾而言，這些方法卻不易推廣與普及。因此本研究採用 GA 以法則模式來找出最適投資規則，並以常用之技術指標為指引，及人性化方式提出買賣建議，期能於實際投資環境中，提供投資大眾另一可行之參考模式。

本研究是以各股股票的技術指標為輸入參數以發展規則庫。此規則庫並非一成不變，例如技術指標的買入賣出的判斷值，在多頭市場與空頭市場中可能有所變動；同時各家公司股票性質不同時也有不同的指標可供選擇。因此，本研究目的主要在於透過股價理論與相關文獻探討，結合技術指標與 GA 來搜尋一可行投資決策模式，並針對不同股票而選擇不同技術指標，以發展其投資決策規則。此外，這些投資規則亦可配合市場資料的變化，而動態調整其最適投資策略。

貳、文獻探討

一、股價理論

股價理論主要在於分析與討論股價定價與漲跌成因，由於各派學說的時代背景與經濟環境不盡相同，所持看法亦眾說紛紜。不過，一般投資分析主要可分為基本面分析與技術面分析；基本面分析主要概念是對於上市公司獲利能力、產業前景等正確的反應，此外如國內政治、經濟景氣、利率匯市、國際情勢、原油等物資價格波動等因素，皆屬基本面分析範疇。

技術面分析係基於統計科學的方法，根據市場過去循環的軌跡探索未來股價的變動（鄭超文，1996）。技術分析理論流派甚多，大致上可分為價格型態、趨勢遵循、市場特性與結構分析。無論理論流派為何，其基本假設是歷史會不斷的重演，由於不斷的重演，才能研究過去歷史資料，用以預測未來股價變化。但技術分析並非依據統計圖表就可精確研判，還需將各種不同原理、市場面、政治面、經濟面及社會面等資訊加以綜合研判，以提高準確度。技術分析中最普遍的方法之一就是運用技術指標，所謂技術指標就是將股市過去的價量關係轉換成可供分析與決策的統計量。近年來利用技術指標來探討股市投資績效的相關研究相當多，這些研究結果顯示，應用技術指標確實可以改善投資績效，但其所運用指標如乖離率、威廉指標、移動平均線及隨機指標等等，因個人研究的對象不同而有不同的指標應用，至於進一步如何有效利用指標，使用何種技術指標，來加以組合運用提高報酬率卻多語意模糊。

近年來相關人工智慧技術在證券市場上應用亦有相當成就，不過對投資者的指引仍有限，例如以一非線性模型，來預測股價走勢（金必煌，1995）；結合套利理論，以基本面因素作為類神經網路輸入端，找出台灣股市投資組合績效（洪新原，1995）；以公司財務報表為分析對象，作為證券投決策參考（湯玉珍，1994）；亦有以技術指標為研

究對象，建構一模糊系統協助挑選具有潛力的股票（許顯榮，1995）；黃金生（1996）探討經濟決定因子對台灣保險業股票風險貼水（risk premiums）的影響，並以類神經網路中之倒傳遞網路（Back-Propagation Network, BPN）與加入前期條件變異數之自我回歸異質條件變異數模式（GARCH）兩種模式來預測保險股票之風險貼水¹，其預測結果顯示 BPN 預測效果優於 GARCH 統計模式；Bauer（1994）以總體經濟因素為變數，先以回歸方式找出各變數與股市報酬率的相關係數，然後從中過濾出較重要的前十個因素，再以 GA 找出較佳的規則組合，並探討與股市報酬率的關係；Graf（1994）則以類神經網路作為應用工具，以道瓊、Nikkei 指數、美國庫券、German Bund-Future、馬克/美元（DM/USD）匯率及外部技術指標等來建立德國股價中、長期預測；Tayler（1995）基於數學分析在建構經濟模式時，需將問題做有限範圍之假設而無法反映複雜的現實的經濟，乃以 GA 來建構智慧型股市模式，他用股利及配息等時間序列為變數來建立一個股價行為預測模式；Chenoweth（1996）運用指數差離指標（moving average convergence-divergence, MACD）作為類神經網路的輸入變數來產生買賣的訊號，並以 S&P 500 股市作為實證對象，根據該系統資料顯示在 54 個交易次數中年報酬率可達 15.99%。上述所舉之例均顯示 GA 及人工智慧相關技術在財務的問題研究及實務應用上有日漸增加的趨勢，然而在技術指標的應用方面，也因各個指標具有其

獨特性，因此使用時自然無法趨向一致。這些文獻中就運用 GA 技術而言，金必煌（1995）是建構一非線性股價預測模型；范饒耀（1996）是以上市公司的財務比率，來預測每股盈餘的模型；就上述的應用，大多是運用人工挑選的指標，作投資決策分析，而不是從歷史資料中，藉由學習機制以汰選相關重要影響指標而進一步作推理決策。本研究之投資規則乃由系統從眾多技術指標中自行挑選，而指標值亦由系統自行產生。

二、技術指標的影響

技術分析的指標種類相當多，其主要是依據交易價格、成交量來分析交易行為的數學式，大約分為價指標、量指標與價量指標三類。本文所採用之技術指標係依據各報紙的證券行情版中常用的技術指標，其簡要說明如下：

移動平均線（MA）

利用統計學上「移動平均」的原理，將每日價格以移動平均計算，求出一個平均值，各日平均值相連而成平均線。價格如採用股價，稱之為股價平均線，如以成交量平均之，則稱為成交量平均線；依採用平均時間，可簡單區分為短、中、及長期，報紙所載以 10 日與 30 日為參考依據。

- 短期：以 6 日及 10 日為準，代表一週的平均價，做為短線進出的依據。
- 中期：以 24 日或 30 日為主，稱為月

¹ 保險股票風險貼水 = 保險股發行量加權報酬 - 一月期定期存款利率

移動平均線，又稱為月線。

- 長期：歐美多以200日為準，至於國內則以144日線（半年線）288日線（年線）最為常用，此外尚有所謂的季線。

一般採用的是葛南薇（Granville）移動平均線八大法則以作為買賣依據，例如所謂「買進時機」是指：當平均線從下降逐漸轉為盤局或上升，而股價從平均線的下方突破平均線時；或股價雖跌破平均線下，但立刻又回到平均線上，此時均線仍持續上升；或股價趨勢線走在平均線之上，股價突然下跌，但未跌破平均線，股價又上升時；或股價趨勢線低過平均線，突然暴跌，遠離平均線時。而「賣出時機」則：當平均線走勢從上升逐漸轉為盤局或下跌，而股價從平均線的上方往下跌破平均線時；或股價雖上升突破平均線，但又立刻回復到平均線之下，且平均線仍然繼續下跌時；或股價趨勢線在平均線下，股價上升但未達平均線又告回落時；或股價趨勢線在上升中股價趨勢線且走在平均線之上，突然暴漲，遠離平均線時。

相對強弱指標 (RSI)

RSI 為目前市場上使用最普遍的技術指標，係依據市場上的供需原理所產生的，其主要是計算某一段時間內買賣雙方力道，作為超買或超賣參考，以避免過早賣出買進，造成賺少賠多的損失。RSI採用的時間週期、原理與移動平均線一樣，隨採用者的喜

好來決定。市場上比較盛行的有5天、14天等，惟採用的週期愈短，敏感度愈高，但變動性太大。雖然週期愈長，趨勢愈穩定，但敏感度相對降低，國內所採用的是6日RSI。至於以RSI值為研判標準，其值亦隨個人喜好而有所不同，簡單之研判例如以RSI值上升至80以上或20以下，作為超買與超賣之判定。

乖離率 (BIAS)

乖離率是以當日平均數值為準，先計算出股價與平均數值間的差距，稱為乖離程度，乖離程度再除以平均數值的百分比便為乖離率。換言之，乖離率為當日指數或個別股當日收盤價與平均數之間的差距。乖離率可分為正乖離與負乖離，若股價在平均數值之上，則乖離為正，反之則為負，股價與平均數值相同，則乖離率為零。BIAS值之研判標準，其值隨個人喜好而有所不同，例如當BIAS達+4%以上或-4%以下，分別為超買與超賣現象，作為買入與賣出時機，不過這些值因應時空環境而有所改變。

心理線 (PSY)

心理線是人氣指標的一種，利用一段時間內股票漲日與跌日的比率，以測試投資人內心趨向買方或是賣方的心理現象，而作為買賣股票的指標。一般投資專家多以13日、24日為短、中期投資指標，國內採用13日為投資指標。一般是以心理線指標界於25%~75%為合理變動範圍，通常是隨大多頭

行情或大空頭行情而有所調整。

三、技術指標的迷思

技術分析的指標很多，無論是從價格觀點、成交量觀點或價量兼具的觀點，其所陳述的數學式，難免在性質上有所重複、或甚至無法反映實際市場變化的指標。一般情況下，使用指標是一種百分率的遊戲，價格變動的隨機性常是指標失效的重要原因之一。因此，一個成功的技術分析師通常使用多種指標作為進出場的判斷，以提高成功率（Chande, 1996）。縱觀各種技術指標的運用，可能會產生下列問題：

- 技術指標的具某種程度相似性，因為指標大多從價、量變化的不同數學式中求得，僅是形式上的不同而已，相對的，其程度上亦存有某種相關性。
- 指標雖具有某種相關性，但由於出處各有不同，彼此相關性並不明顯，如何搭配使用指標便是一個值得思考問題。
- 指標取自過去的價或量的一種資料，指標只是一種買賣的參考依據，每種指標皆有鈍化的時候；在多頭市場上，可能難以碰觸指標的買入訊號；相同地，在空頭市場中，亦有可能是持續的買入訊號而造成虧損，投資者不可不慎。
- 指標計算期間的引用常隨個人經驗而有不同，不同個股亦有不同特性，不可

一概使用。比如，相對強弱指標以 6 日、14 日較常用，對於買入點、賣出點說法不盡相同，以六日 RSI 買入點為例，有以 80 為界、或以 90 以上為超買點，都有或多或少的差異。

基於上述考量投資者指標資訊之來源與取材之方便性，本研究之技術指標採用報紙所刊載的 6 日 RSI、6 日 BIAS、10 日與 30 日價 BIAS、13 日 PSY 作為輔助指標。為符合各股特性與避免指標鈍化問題，本研究採用 GA 並依據大盤歷史資料來自行選擇組合使用的指標及推論最適當的買賣判斷點。

參、遺傳演算法理論

過去的幾年裡，科學家發展出許多數學方法來解決最適化問題。然而這些方法中卻沒有一個模式能解決所有工程學上最適化的問題。這些方法通常是在從設定的起始點中開始尋找最適值，如果問題不止只有一個局部最適值（local optimum），則所找的最適值可能不是總體最適值（global optimum），其結果端視起點選擇而定。而統計方法有時雖可避免上述問題的發生，但仍需兼顧諸多理論之基本條件及假設。除此，隨著資料複雜度的增加，其處理時間也相對增加。人工智慧領域中，類神經網路雖擁有強大的學習能力，但就像黑盒子欠缺解釋的能力，無法使投資人信服；模糊專家系統雖然容易納入專家的知識，並可提供投資人適度的解釋，但基本上卻缺乏學習的能力，除非加入複雜的

學習機制，不然無法隨時間的變遷而作自我修正 (Deboeck, 1994)。因此，本研究採用遺傳演算法其動態資料學習的能力，以發展專家知識規則，以作為投資者參考。

一、遺傳演算法基本架構

GA 是由 Holland 根據自然生物系統演化過程開發出來的，運用達爾文「適者生存、不適者淘汰」理論 (Darwinian survival of the fitness theory)，從舊的族群中得到較好進化物種，然後彼此隨機交換資訊以期產生優良的下一代 (Holland, 1975)。Holland 的基本遺傳演算法 (Simple Genetic Algorithm, SGA) 其基本運算方式為再生 (reproduction)、交配 (crossover) 和突變 (mutation) 三個運算單元。如圖 1 所示，其過程包含了初始族群 (initialize the population) 的決定、適合度 (fitness) 之計算、再生、交配、突變與收斂判別等。在實際電腦運作中，是以一群位元來模擬染色體中的遺傳基因 (gene)，並加以隨機編碼而組成一群字串，此字串意味著自然界生物的染色體，一組字串代表一個生物個體；字串與字串 (父代) 模擬自然界生物的交配過程以產生一組新字串 (子代, offspring)，含有上一代特性的新字串經世代推演以產生更優秀的下一代，這種運算一直持續至達成目標為止，整個過程中最優秀的子代隱喻為問題的解答。

從 SGA 模擬生物演化的概念與實作上

了解，其計算過程中是使用字串運算，所謂字串運算是將參數直接用在編碼上，編碼可以二進位制編碼，也可以實數方式編碼，是一種 context free 的演算法，其所代表的意義與各種題意上的限制由染色體 (chromosome) 和適合度 (fitness) 來決定，減少了繁複的數學運算處理，比一般搜尋方式更具彈性 (Birgchenhall, 1995)。以重複循環演算與蒐尋過程來看，GA 搜尋方式是以資訊的交換進行的，但他不是隨機漫步式 (random walk) 的交換資訊，而是一個有效率地利用前一代資訊來搜尋最適解 (Goldberg, 1989)；GA 並以平行方式處理，同時搜尋空間的多組解，以多點的搜尋來替代單點搜尋，所以降低了落入局部最適解的機率，並加快獲致最適解的速率；為避免某些演化收斂過速或所謂落入局部最適解 (local optimum) 時，GA 中的突變機制，會使搜尋有機會跳脫局部解至全域最適解 (global optimum)。同時，運用 GA 並不需要太多的背景知識，比如各變數間的相依或獨立，或預期的結果，完全以目標函數 (適應函數) 作為評估的考量。此外，由於它是以盲目搜尋 (blind-search) 的隨機方式找出可行的結果，所以，結果好壞取決於適應函數選取的適當與否 (Greene & Smith, 1987)。

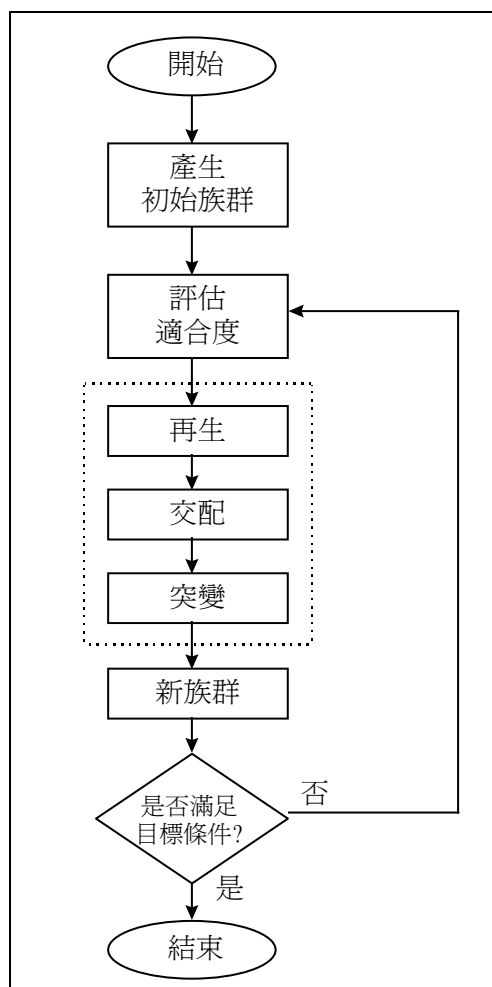


圖 1、SGA 流程圖

二、 模擬前置作業

編碼 (coding) 與解碼 (decoding)

爲了制定參數搜尋空間，設計者必先對參數的可能變動範圍做一衡量，以固定長度的字串來加以編碼，最簡單且廣泛使用的方法是二進位編碼法。例如有一參數變動範圍在[-4,4]之間，若以五位元編碼，則可作 $2^5=32$

個的區隔，那麼 $00110_2=6_{10}$ 代表實際值的意義爲

$$-4+[4-(-4)]*6/32=-2.5$$

編碼與解碼乃針對問題模式的不同而有不同的定義；就問題精確度而言，碼的長度愈長則變動範圍內區隔愈多，解析度愈佳；但相對地卻會導致執行速度減緩、及大量記憶體耗費，並需由設計者從中調配。

定義適合度函數 (fitness function)

根據問題形式與需求定義此族群生存的條件，適合度函數用以決定每個個體生存與否或優先次序，以此成爲選取的依據。

三、 模擬作業

產生初始族群：

完成編碼與解碼程序後，即可清楚藉由染色體代表問題涵意，經由電腦亂數編碼方式建立之第一代族群，即意味著初始族群各個體擁有不同特性。亂數編碼的字串，其本身與問題並無相關性，僅在解碼與計算適合度值時才顯現與原問題的相關性。在實際演算模擬過程中，需考量族群大小 (population size) 的取決，一般而言字串的長度較長者均需採用較多的初始族群，以達均勻分布於搜尋空間。族群大小太少則較無法獲致問題真正收斂的條件；反之，若太過則易造成計算量增加而較耗時。

評估適合度

評估適合度目的在於作為選取的標準，所謂選取 (selection)，為自然界適者生存的現象，是指依據目標函數 (objective function) 經過解碼所計算出的適合度值以作為選取標準，較大適合度的父代有優先或較大的機會產生子代。GA 雖是採字串運算方式，然計算適合度值時，需由解碼原則將字串加以解碼為實際值，較趨向解決問題模式者有優先或生存的機會，適合度較低者被淘汰或降低被選取機會。較常用之選取方式，諸如轉輪盤選取法 (roulette wheel selection) (Goldberg, 1989)、排序法 (rank-based selection) 及競賽法 (tournament selection) 等。所謂轉輪盤選取法先計算出同代染色體之適合度總和，然後再求出染色體適合度與總和之比值以做為被選取之機率，此機率累計為一輪盤，因此較佳適合度之染色體之被選取機率較高，相對在輪盤上所佔之面積較大，被選取到的機率也較高。排序法是將母代染色體之適合度做排序來形成機率圓盤，排序高者有較高之選取機會。競賽法是以較高適合度染色體作為交配之選取對象。

再生

所謂再生，就是將父代依個別的適合度值高低，作為被選取機率的高低比率，由選擇機制從中挑選個體，將之複製至交配池 (mating pool)，並待下一步驟進行。此運算單元的目的在於將適應環境較佳、具有較強生存能力之染色體，有較高機率將其特性傳衍至下一代。

交配 (Crossover)

這一步驟是將交配池中所有染色體做兩兩遺傳互換重組的動作。以產生新的個體。主要目的是期望藉由所挑選出的父母親個體做結合後，能產生更優秀的兩個子代，其交配的過程如下：

- 從交配池中選取兩個個體進行配對，選取方式可以是隨機式的或指定的。
- 在這兩個體中決定交配點，以單點交配為例 (如圖 2 所示)，將位於交配點的左側或右側的字串互換。常用的交配方法包括單點、兩點、多點交配、uniform-point、partially matched crossover (Goldberg, 1989)，order or linear order crossover (Falkenaue, 1991) 等多種方式。

父代	子代
父1: 1001 <u>10</u> 11	子1: 1001 <u>11</u> 01
父2: 0100 <u>11</u> 01	子2: 0100 <u>10</u> 11

圖 2、單點交配

模擬過程中需考量交配比率 (crossover rate) 的多寡，比率過高會造成父代較佳特性的流失，太低則會失去進化的目的。所謂交配比率是指兩個染色體出現交配的機率。

突變 (Mutation)

所謂突變是指隨機選取字串，將其中某

一位元或數個位元改爲它值，如圖 3 所示，突變率即爲一個新染色體內的值被隨機改變的機率。突變的意義在於產生差異性較大的子代。在空間搜尋中，當收斂至某區間時，有可能是收斂至局部最適解，而掉入局部解陷阱中。因此藉由突變期望能避免過早收斂的缺失，進而能找出新的搜尋區域。雖然藉突變方式僅可能避免過早收斂，然而亦有可能造成系統的不收斂。因此，系統的突變率（mutation rate）在扮演搜尋最適解的過程中即顯得格外重要，而一般系統之預設突變率都設的很小或不准予發生，例如 0.05 爲一較常用之設定值。

新族群：

父代經再生、交配、及突變過程而產生新的子代，並根據預定方式來完全或部分取代原族群，成爲一個新的循環點。

子代	突變後
子1 : 10011101	子1: 00011101
子2 : 01001011	子2: 01001111

圖 3、遺傳突變

停止條件 (Stopping Criteria)

GA 程式正常的執行下，族群演化過程中乃趨向全域最適解發展。至於在判斷系統是否已經收斂、何時才停止演化、或目標是否達成，常用之簡易方式，就是直接設定演算

法演化幾個世代以後便停止；如遺傳代數爲 500 代或 1000 代，至於幾代才合適則需要看問題的複雜程度或資料量而定。另一方式是以時間來決定停止演化的條件，例如 10 分鐘或 1 小時。上述兩種停止條件的缺點是無法確定是否已找到最適解。此外，亦有以每代進步的幅度大小來決定是否停止演算。常見的有兩種判斷方式，首先是以進步情形爲依據，它又可分爲以每代族群的適合度之平均值、及以目前搜尋所得之最適適合度兩種爲基準。當適合度平均或最適適合度在演化過程中已無法進步或進步有限時，表示已經收斂。例如最近 100 代的演化中進步在 0.01% 以內；另一種方式是以適合度平均值與搜尋過程中之所得最佳適合度的誤差值百分比爲依據，達到某種程度如百分比來代表已經收斂。

四、系統發展與測試

本研究所採用之歷史資料爲民國八十四年一月起至民國八十五年八月三十一日的台灣股價加權指數，資料來源爲台灣經濟新報資料庫。系統建構係以針對一般投資者在不考量融資融券及資本額的限制條件下，就股價加權指數作買入賣出點之建議。此遺傳投資模型運用 GA 進行搜尋，以建立一套整合技術指標的模型。本研究採用常見的四種技術指標，其投資策略表達方式，可如下列規則表示：

假如

滿足（條件一）與（條件二）與（條件三）
與（條件四）與（條件五）與（條件六）時
則
買入（或賣出）並持有（或放空）
（一段期間）

其中（條件）的敘述內容為：技術指標的
值大於或小於某一個值

至於本系統對於規則之發展係以 GA 從
過去之股價加權指數的變動中，比較與觀察
各技術指標之間變動的共通性。從中找出影
響較大的技術指標，進而滿足最適的投資報
酬率。因此，四個條件式經 GA 搜尋後，可
能採用的條件式至多四個，其結果將視資料
內容及結構而定。舉例而言，一個經由搜尋
後所獲致的結果可如下列規則所示：

- 規則一：假如 6 日 RSI 小於 20 且 6 日
BIAS 小於 -4% 時
則 買入並持有 12 日
- 規則二：假如 6 日 RSI 大於 80 且 6 日
BIAS 大於 5% 時
則 賣出並放空 10 日

系統建構：

在發展規則過程中，GA 以部分歷史資料
為學習對象，並以其它時段資料驗證結果。
本研究以民國 84 年 1 月 5 日起至 85 年 8 月

31 日止為研究期間，將前 286 日為第一區間
之空頭趨勢作為學習期間，後 184 日之多頭
趨勢為測試期間。本研究之系統建構、參數
定義與操作過程說明如下：

編碼與解碼：

編碼與解碼主要根據上述規則來編排
染色體，並以其所代表涵意來作為適合度的
定義。本研究所採用技術指標、指標選擇、
及大於或小於的邏輯判斷各六組，另外包括
持有或放空期間一組，買入賣出共兩類合計
34 組。其規則組合變化，以單一條件敘述內
容之可能組合為 $2*100$ ，四種技術指標組合為
 $(2*100)^4$ ，整個規則之可能組合變化共為
 $(2*100)^4 * 2*30$ 。在如此龐大空間搜尋，一
般的方法不易在有限時間內找到一組最適投資
規則，本研究乃採用 GA 之強大搜尋功能來
完成投資規則之搜尋。定義 GA 基因值域之
範圍如表 1，其中技術指標的值域範圍在演
算時，需經過轉換至該指標之適當值。

適合度：

運用 GA 來整合技術指標並完成投資規
則的搜尋，主要是期望建立最適報酬率，所
謂最適報酬率乃以持有期間（或放空期間）
計算所得之報酬率作為適合度指標。亦即針
對買入規則時，期望持有期間之報酬率為最
大；賣出規則時則期望放空期間之報酬率為
最小。

表 1、編碼方式及其涵意

名稱	基因個數	值域	值域形式	說明
技術指標	4	[0,100]	整數	6 日 RSI、6 日價 BIAS、6 日量 BIAS、13 日 PSY
指標選擇	6	[0,1]	整數	0 表該指標不用，1 表選用該指標。
邏輯判斷	6	[1,2]	整數	1 表小於，2 表大於
持有期間 或 放空期間	1	[0,30]	整數	以營業日數為單位

操作條件：

染色體的編碼方式是以實數形式表示，值的來源採用隨機方式產生；所謂族群大小是指所有染色體的集合。適當的族群大小有利於搜尋效率，根據 Srinivas (1994) 建議 30-100 為合適區間，本實驗設定族群大小為 50；交配率與突變率以一般經驗值設定分別為 0.5 與 0.06；本研究中之交配方式採用一致交配方式；結束條件是以最近 100 代的演化中適合度進步在 0.01% 以內，就停止其搜尋，其所得解即為最適組合方式。詳細參數設計值請參照表 2 所示。

學習方式

本研究運用 GA 找出最適組合解，整個搜尋空間區分為買入與賣出兩部分。

買入的搜尋策略與目標：

策略：**找出學習期的技術指標組合與值域並**

計算持有期間（測試期）的報酬率

目標：**持有期間（測試期）的最大報酬率**

賣出搜尋策略與目標：

策略：**找出學習期的技術指標組合與值域並**

計算持有期間（測試期）的報酬率

目標：**放空期間（測試期）的最小報酬率**

表 2、GA 參數

參數名稱	說明
染色體長度	34
族群大小	50
交配機率	0.5
突變機率	0.06
交配方式	Uniform

績效評估

績效衡量指標所採用之報酬率定義前已述及，正確的評估投資策略中，買入並持有

之期望報酬率應為正值，賣出放空期間之期望報酬率應為負值。其中持有期間或放空期間之總報酬率（R）定義如下：

$$R = \sum_{i=1}^m \sum_{j=1}^n r_i$$

上列計算式中之 n 表示為持有期間或放空期間，m 為系統執行交易策略之總次數，r_i 為當日報酬率。由於交易成本的存在，交易次數的不同會造成績效衡量之失真，舉例而言，如以兩條不同規則作為投資策略，其結果分別為 10 次與 20 次交易次數且皆有 10% 的報酬率，在交易成本考量下，績效應以 10 次交易為佳。此外，在一個大多頭市場中，若以總報酬率為績效衡量，其結果易傾向買入過於頻繁來達成總報酬最大，反之在空頭市場中亦可能有賣出過多之類似情形。因此本系統以平均每筆交易之報酬率（AR）作為系統執行交易績效衡量標準，其算式如下：

$$AR = \frac{1}{m} R$$

然而，如以每筆交易之報酬為績效衡量標準，則易造成單次交易而總報酬率偏低之可能，與實際投資策略不合。因此本系統以每筆交易報酬率最大為績效目標，同時限定投資次數在總交易日之 5% 以上，以接近實際投資情形；同時設定檢查程式，排除同時出現買入與賣出訊號。整個投資包含買入與賣

出，因此研究期間之報酬定義為整體報酬（TR），其計算方式如下：

$$TR = (AR * m)_{\text{買入部分}} - (AR * m)_{\text{賣出部分}}$$

肆、結果分析與討論

本研究係經由模擬實際買賣持有日期或放空期間找出最適之買入或賣出投資規則。操作方式是當出現符合買入搜尋所得規則時，則買入股價加權指數，並持有該指數至規則庫所顯示持有期間才賣出；若是符合賣出規則的訊號時則賣出股價加權指數，並放空一段期間。

買入部分

表 3 為 GA 模擬之搜尋結果。表四中所謂傳統判斷是指依據 RSI、價、量 BIAS、PSY 等四種指標之一般經驗判斷方式，從買 1 區可以發現，總投資數為 1 次；當放寬買入判斷規則如買 2 區顯示，此時出現 11 次買入訊號，投資率偏低，且每筆報酬率為負值，表示以此持有一日情況下，意味著第二日上漲的機會僅佔 36% 而已，其成效不甚良好。當放寬判斷規則並增加持有日如買 3 及買 4 區時，投資次數增加，但每筆報酬率仍為負值；買 5 區與買 4 區以相同買入規則與持有期間，發現其報酬率則為 62.9% 之正值。觀察買 1 至買 4 區，其應用期間為空頭趨勢，買 5 區則為多頭趨勢，可見傳統買入判斷不利於空頭趨勢。

表 3、模擬結果

1 資管評論 第八期 一九九八年十二月

模擬方式		傳統判斷					GA	
買入類別		買 1	買 2	買 3	買 4	買 5	訓練	測試
應用期間(起)		84.1.5				85.1.4	84.1.5	85.1.4
應用期間(迄)		84.12.30				85.8.31	84.12.30	85.8.31
總次數		286	286	286	286	184	286	184
買入規則	6日RSI	<20	<20	<30	<20	<20	>87	
	6日股價BIAS	<-4%	<-4%	<-2%	<-2%	<-2%	>-4%	
	6日成交量BIAS	<-4%						
	13日心理線	<25						
買入結果	持有期間	1日	1日	6日	23日	23日	23日	23日
	投資次數	1	11	45	32	19	15	26
	投資率	0.3%	3.8%	15.7%	11.2%	10.3%	5.2%	14.1%
	正確數	0	4	16	9	14	8	17
	正確率	0.0%	36.4%	35.6%	28.1%	73.7%	53.3%	65.4%
	每筆報酬率	-0.7%	-1.4%	-1.5%	-2.3%	3.3%	0.4%	3.7%
	總報酬率	-0.7%	-15.4%	-68.4%	-72.6%	62.9%	5.6%	95.4%
賣出類別		賣 1	賣 2	賣 3	賣 4	賣 5	訓練	測試
賣出規則	6日RSI	>80	>80	>70	>80	>80	>14	
	6日股價BIAS	>+4%	>+4%	>+2%	>+2%	>+2%		
	6日成交量BIAS	>+4%					<2.8%	
	13日心理線	>75						
賣出結果	放空期間	1日	1日	6日	12日	12日	12日	12日
	投資次數	0	225	200	216	102	21	5
	投資率	0.0%	78.7%	69.9%	75.5%	55.4%	7.3%	2.7%
	正確數	0	131	118	129	46	15	3
	正確率	-	58.2%	59.0%	59.7%	45.1%	71.4%	60.0%
	每筆報酬率		-0.3%	-0.9%	-1.6%	1.0%	-2.8%	-0.7%
	總報酬率		-60.8%	-174.0%	-343.4%	106.1%	-58.4%	-3.6%
整體報酬		-0.7%	45.4%	105.6%	270.8%	-43.2%	63.9%	99.0%

然而，從 GA 訓練結果發現，將 84 年間訓練 結果應用於 85 年測試期間，其成效相當良

好。當交易策略如下：

假如 6日RSI 大於 87
且 6日股價BIAS 大於 -4% 時
則 買入並持有 23日

以該訓練期間所得規則作為測試期間投資策略，GA 投資次數為 26 次，投資率達 14.1%，比較相同應用期間之買 5 區，其正確率（65.4%）雖不及買 5 區（73.7%），但每筆報酬率（3.7%）較佳，顯示 GA 規則優於傳統判斷方式，買入總報酬率自然大幅領先傳統方式。

賣出部分

放空部分的報酬率應為負值，從表三傳統判斷方式之賣 1 至賣 4 區中發現，總報酬率均為正值，顯示空頭趨勢適合放空操作，然而放空投資率高達七成以上（如賣 2、賣 3 與賣 4 區），不符合實際投資情形。若根據賣四區之賣空訊號，在賣 5 期間之多頭趨勢放空時，將遭受相當損失，也就是說以傳統判斷方式恐無法找到較適當放空時機；以 GA 訓練所得之放空策略為下列所示：

假如 6日RSI 大於 14
且 6日成交量BIAS 小於 2.8% 時
則 今日賣出並放空 12日

值得探討的是放空策略中，BIAS 為「小於」一個值，似與傳統判斷的「大於」某一

正值的觀點背道而馳；不過，此 BIAS 需與另一條件相互搭配方為適當。表四賣出中 GA 訓練結果的正確率為 71.4%；將此規則用於測試期間，正確率為 60%，每筆報酬率為負值，此說明 GA 訓練結果在不利放空之多頭趨勢下，仍可以找到適當的放空時機。由此可見 GA 所獲致訓練結果均較直接觀察技術指標的決策方式為佳。從整體報酬來看，傳統方式在 2、3、4 區中可以倍數獲利，然其獲利主要來自於賣空部分，在多頭趨勢之 5 區未能獲利。反觀 GA 所學習之專家規則，以空頭趨勢所學習規則應用於多頭之測試期間，其整體報酬幾近 1 倍，可見 GA 專家知識規則在股市投資上具有穩定報酬。

伍、結論

由上述結果可以了解，就買入部分而言，以相同持有期間之 GA 規則每筆報酬率（3.7%）高於傳統判斷方式（3.3%）；賣出部分之 GA 規則（-0.7%）亦比傳統判斷方式（1.0%）能提供正確的賣出時機，可見以 GA 方式建立所獲致的交易規則明顯優於傳統判斷方式。從模擬過程中發現，學習方式之搜尋策略與目標可能會形成交易次數太低或頻繁的結果，進而影響結果之精確性；或是否一個交易日中同時產生買入與賣出信號。本系統架構已融入限制式與檢查式等防弊措施，檢驗表四結果，其交易模式應屬合理，檢查式中亦無同時產生買賣情形。顯見以技術指標之 GA 專家知識規則在股市投資上有

其適用性。

再從研究過程中所發現的 GA 兩個特點來看，GA 的確比一般模式具有較多的彈性。這兩個特點是：一、能在空間做全域搜尋而得近似最適解，同時又不需強大的背景知識便能發展合理的投資決策，亦可彌補數學計算之繁瑣過程。二、依據 GA 的字串運算特性，在應用上只要定義適當的編碼、解碼條件與適應函數，都可以趨向一個近似最適的解。本研究雖僅對大盤作分析，若欲針對各家股票作分析與預測，其作法亦是相近的。本研究所採用之資料為常見的技術指標，未來將可考量納入其他的技術指標，以作為發展交易規則之候選指標。另外，由於各股有其不同的交易行為週期性，因此各股技術指標使用之週期應有所不同；本文僅對固定期間如 6 日 RSI 或 13 日心理線等為對象，後續研究者可以浮動週期的技術指標為搜尋方式，來表現各股不同的週期性。股市的影響因子不只技術面的指標，在考量台股市場易受其它因素影響之特性，系統可朝動態調整方向發展，比如適度結合基本面與消息面指標的綜合研判，惟消息面如何分類以及納入系統，應是未來值得深思的課題。

參考文獻

鄭超文，「點線賺錢術技術分析詳解」，財

訊出版社，民國八十五年七月修定九版二刷。

洪新原，「結合套利定價理論及類神經網路以支援投資組合分析之研究」，高雄市資訊發展協會研討會 pp. 25-39，民國八十四年六月。

范饒耀，「基因演算法在財務預測之應用」，政治大學資訊管理研究所碩士論文，民國八十五年六月。

金必煌，「運用基因演算法建立動態證券模型市場技術模型」，交通大學資訊管理研究所碩士論文，民國八十四年六月。

許顯榮，「使用模糊專家系統之智慧型選股決策支援系統」，台灣大學資訊工程研究所碩士論文，民國八十四年六月。

戴敏育，「應用基因演算法發展模糊預測之研究 - 以銷售預測為例」，淡江大學資訊管理研究所碩士論文，民國八十四年六月。

湯玉珍，「模糊理論應用於證券投資決策之研究」，大葉工學院事業經營研究所碩士論文，民國八十三年六月。

黃金生、施東河、劉建利等，「類神經網路在台灣人壽保險業股票風險溢酬預測的應用」，資訊管理學報，第三卷第一期，1996。

Bauer, A. J., "Genetic Algorithms and Investment Strategies," *John Wiley & Sons*, N.Y., 1994。

Birgchenhall, C. R., "Genetic Algorithms, Classifier System and Genetic Programming and Their use Models of Adaptive Behavior and Learning," *The Economic Journal*, Royal

- Economic Society, 1995, pp. 788-795.
- Chande, T. S., and Kroll, S., "The New Technical Trader Boost Your Profit By Plugging into The Latest Indicators," *John Wiley & Sons, Inc.*, 1996.
- Colin, A. M., "Genetic Algorithms for Financial Modeling," in G. J. Deboeck, (Eds.), *Trading on the Edge*, John Wiley & Sons, NY., 1994.
- Chenoweth, T., et al., "Embedding Technical Analysis into Neural Network Based Trading System," *Applied Artificial Intelligence*, Vol. 10, No. 6, 1996, pp. 523-541.
- Davis, L., "Genetic Algorithms and Financial Applications," in G. J. Deboeck, (Eds.), *Trading on the Edge*, John Wiley & Sons, N.Y., 1994.
- Deboeck, G. J., "Why Use Fuzzy Modeling?" in G. J. Deboeck, (Eds.), *Trading on the Edge*, John Wiley & Sons, NY., 1994.
- Falkenauer, E. and Bouffouix, S., "A Genetic Algorithm for Job Shop," *Proc. of the 1991 IEEE Int'l conference on Robotics and Automation Sacramento, California, U.S.A.*, April 1991, pp. 824-829.
- Graf, J., "Application of Learning Algorithms to Predicting Stock Prices," in V. L. Plantamura et al., (Eds.), *Frontier Decision Support Concepts*, John Wiley & Sons, NY., 1994.
- Greene, D. P. and Smith, S. F., "A Genetic System for Learning Models of Consumer Choice," *Proceeding of the Second International Conference on Genetic Algorithms*, J. J. Grefenstette (Eds.), Lawrence Erlbau, Associates, Hilldale, NJ., 1987.
- Goldberg, D. E., "Genetic Algorithms in Search, Optimization and Machine Learning," *Addison-Wesley Publishing Company*, 1989.
- Goldberg, D. E., "Genetic and Evolutionary Algorithms Come of Age," *the Communication of ACM*, March 1994, pp. 113-119.
- Holland, J. H., "Adaptation in Natural and Artificial System," *University of Michigan Press*, 1975.
- Levy, H., "Stochastic Dominance and Expected Utility,: Survey and Analysis," *Management Science*, April 1992, pp. 555-593.
- Mott, S., "Case-Based Reasoning : Market, Applications and Fit with Other Technologies," *Expert Systems with Applications*, Vol. 6, 1993, pp. 97-104.
- Wong, F and Lee, D., "Hybrid Neural Network For Stock Selection," *International Conference on AI Application on Wall Street*, 1993, pp. 294-301.
- Srinivas, M., "Genetic Algorithms : A Survey," *Computer*, June 1994, pp. 17-26.
- Tayler, P., "Modeling Artificial Stocks Markets Using Genetic Algorithms," in S. Goonatilake & P. Treleaven (Eds.), *Intelligent Systems for Finance and Business*, John Wiley & Sons, NY., 1995, pp. 271-288.
- Yeh, C. H., " A Literature Review on the Components Improvement of Genetic Algorithms," *Journal of Taiwan Water Conservancy*, Vol. 44, No. 1, March 1996, pp. 92-105.