

國立臺灣大學電機資訊學院資訊工程學系

碩士論文

Department of Computer Science and Information Engineering

College of Electrical Engineering and Computer Science

National Taiwan University

Master Thesis

使用增強式學習法改善一個簡易的臺灣股價指數期貨

當沖交易系統

Using Reinforcement Learning to Improve a Simple
Intra-day Trading System of Taiwan Stock Index Future



林敬斌

Ching-Pin Lin

指導教授：呂育道 博士

Advisor: Yuh-Dauh Lyuu, Ph.D.

中華民國 98 年 7 月

July, 2009

國立臺灣大學碩士學位論文

口試委員會審定書

使用增強式學習法改善一個簡易的臺灣股價指數期貨

當沖交易系統

Using Reinforcement Learning to Improve a Simple
Intra-day Trading System of Taiwan Stock Index Future

本論文係林敬斌 君 (R96922100) 在國立臺灣大學資訊工程學所完成之碩士學位論文，於民國 98 年 6 月 25 日承下列考試委員審查通過及口試及格，特此證明

口試委員：



(簽名)

(指導教授)

| | |
|-------|-------|
| _____ | _____ |
| _____ | _____ |
| _____ | _____ |
| _____ | _____ |

系主任、所長

(簽名)

誌謝

兩年的研究所生涯倏忽即逝，轉眼間就接近了尾聲。在此首先感謝指導教授呂育道博士的諄諄教誨，帶領我們走入財務工程的領域，讓我們將所學知識與金融計算相互結合；傳道授業解惑之餘，也讓我們適性發展，並協助我們將論文圓滿完成。

接下來要感謝實驗室的學長姊、同學以及學弟妹，在有疑惑時不吝指導，彼此討論教學相長，並且分享資訊傳授經驗，使我對未來有更多的了解及規劃。

最後要感謝我的家人以及女友，在這段期間內的鼓勵與陪伴。大家的支持與指導都是我完成此篇論文的動力與助力。



摘要

本論文應用增強式學習法中的 Q-learning 於改善一個簡易的臺灣股價指數期貨當沖交易系統，使用歷史資料模擬原本策略的績效，以及訓練與檢測改善後的績效。

研究標的為臺灣股價指數期貨(以下簡稱臺股期貨)，訓練資料為 2003 年到 2007 年每個交易日的每筆成交資訊，檢測期間為 2008 年 1 月到 2009 年 5 月。

被改善的策略為一作順勢交易的通道突破系統，增強式學習法的訓練結果用以判斷每次交易是否該改作逆勢交易，以期取得較大的獲利機會。

關鍵詞：臺灣股價指數期貨、程式交易、當日沖銷、機器學習、增強式學習法、Q-learning



Abstract

This thesis applied Q-learning algorithm of reinforcement learning to improve a simple intra-day trading system of Taiwan stock index future. We simulate the performance of the original strategy by back-testing it with historical data. Furthermore, we use historical information as training data for reinforcement learning and examine the improved achievement. The training data are the tick data of every trading day from 2003 to 2007 and the testing period is from January 2008 to May 2009. The original strategy is a trend-following channel breakout system. We take the result of reinforcement learning to determine whether to do trend following or countertrend trading every time the system plans to make position.

Key words: Taiwan stock index future, programming trading, intra-day trading, machine learning, reinforcement learning, Q-learning



目錄

| | |
|---|-----|
| 口試委員會審定書 | i |
| 誌謝..... | ii |
| 摘要..... | iii |
| Abstract..... | iv |
| 第一章 緒論 | 1 |
| 1-1 簡介..... | 1 |
| 1-2 論文架構..... | 1 |
| 第二章 背景知識 | 2 |
| 2-1 臺灣股價指數期貨 (Taiwan Stock Index Future)..... | 2 |
| 2-2 當日沖銷 (Intra-day Trading)..... | 2 |
| 2-3 K 線 (Candlestick)..... | 3 |
| 2-4 機器學習與人工智慧 (Machine Learning and AI)..... | 3 |
| 2-5 增強式學習法 (Reinforcement Learning)..... | 4 |
| 2-6 Q-learning..... | 5 |
| 第三章 研究設計 | 6 |
| 3-1 研究背景..... | 6 |
| 3-2 研究目的..... | 6 |
| 3-3 研究方法..... | 7 |
| 3-3-1 假設..... | 7 |
| 3-3-2 資料來源..... | 7 |
| 3-3-3 學習方法..... | 7 |
| 3-3-4 系統實作..... | 8 |
| 第四章 實驗結果 | 9 |
| 4-1 方法一..... | 9 |
| 4-2 方法二..... | 11 |
| 4-3 方法三..... | 13 |
| 第五章 數據分析及結論 | 15 |
| 5-1 數據分析..... | 15 |
| 5-1-1 方法一..... | 15 |
| 5-1-2 方法二..... | 15 |
| 5-1-3 方法三..... | 16 |
| 5-2 結論..... | 16 |
| 5-3 未來展望..... | 16 |
| 參考文獻..... | 17 |

圖目錄

| | |
|---|------------------------|
| (圖一) K線..... | 3 |
| (圖二) 增強式學習法..... | 4 |
| (圖三) 通道突破系統示意圖 | (圖四) 較長線的壓力支撐示意圖.....7 |
| (圖五) 方法一 2008 年未扣成本前，改善前計 1833 點，改善後計 3116 點..... | 9 |
| (圖六) 方法一 2008 年已扣成本後，改善前計-1296.5 點，改善後計-7.5 點.... | 9 |
| (圖七) 方法一 2009 年未扣成本前，改善前計 539 點，改善後計 406 點..... | 10 |
| (圖八) 方法一 2009 年已扣成本後，改善前計-343 點，改善後計-476 點..... | 10 |
| (圖九) 方法二 2008 年未扣成本前，改善前計 1056 點，改善後計 1681 點..... | 11 |
| (圖十) 方法二 2008 年已扣成本後，改善前計-118 點，改善後計 284 點..... | 11 |
| (圖十一) 方法二 2009 年未扣成本前，改善前計 295 點，改善後計 664 點..... | 12 |
| (圖十二) 方法二 2009 年已扣成本後，改善前計-125 點，改善後計 188.5 點..... | 12 |
| (圖十三) 方法三 2008 年未扣成本前，改善前計 939 點，改善後計 1531 點.... | 13 |
| (圖十四) 方法三 2008 年已扣成本後，改善前計 572.5 點，改善後計 1165 點..... | 13 |
| (圖十五) 方法三 2009 年未扣成本前，改善前計-66 點，改善後計 7 點..... | 14 |
| (圖十六) 方法三 2009 年已扣成本後，改善前計-207 點，改善後計-134 點.... | 14 |



表目錄

| | |
|-----------------------|----|
| (表一) 方法一的改善前後比較表..... | 15 |
| (表二) 方法二的改善前後比較表..... | 15 |
| (表三) 方法三的改善前後比較表..... | 16 |



第一章 緒論

1-1 簡介

增強式學習法在機器學習的領域中扮演了重要的角色，其已被證實應用在許多不同的領域中皆有不錯的成效，本論文將其 Q-learning 學習法應用於改善一個簡易的台股期貨當沖交易系統，並檢驗其成效。

訓練資料為由臺灣期貨交易所提供的 2003 年到 2007 年每筆成交資訊。基本上，系統使用的是由每筆成交資訊得到的一分鐘 K 線，意即每一分鐘的行情都會被畫成一根 K 線。每筆成交資訊只是用於較精確的交易時間點，以及要將系統改成接收盤中即時資訊時可以較為方便。

原本的簡易策略為一短線的通道突破系統，其交易原則為順勢交易，意即突破壓力就作多、跌破支撐就作空，並有其特定的停利停損策略。但有時突破會回擋、跌破會反彈，使得短線上逆勢交易會比順勢交易來得有利。本論文便以增強式學習法從訓練資料的模擬交易中統計報償，藉以決定在檢測期間中同樣狀態下出現的交易訊號，應作順勢交易或逆勢交易才會有較大的獲利機會。

1-2 論文架構

第一章 緒論

簡介並且闡述系統的特性，以及列出論文的架構。

第二章 背景知識

概略性地介紹本研究使用到的一些名詞或是知識。

第三章 研究設計

介紹本研究的一些目的與相關細節。

第四章 實驗結果

列出三種方法於檢測期間的實驗結果。

第五章 數據分析及結論

對實驗數據的分析與比較，以及做出結論和未來展望。

第二章 背景知識

2-1 臺灣股價指數期貨 (Taiwan Stock Index Future)

臺灣證券交易所股價指數期貨契約由臺灣期貨交易所(TaiFex)推出，於1998年7月21日開始交易。交易標的為臺灣證券交易所發行量加權股價指數，契約價值為台股期貨指數乘上新臺幣200元，每日最大漲跌幅為前一交易日結算價的上下百分之七。

近年來，台股期貨已成為我國流動性極高的金融商品，不同的市場參與者，依其交易目的概分為避險者(hedgers)、投機者(speculators)及套利者(arbitrageurs)。他們規避或承受著不同的風險，卻同樣地使得期貨市場越來越繁榮；搭配上從2001年12月24日開始交易的臺灣證券交易所股價指數選擇權契約，使得整個市場更加活絡且不易受到單方面的操控。

台股期貨提供較單一股票宏觀的觀點來看待整個市場，且比股票容易交易，不僅易於放空，當沖的限制也較小；從2007年10月8日開始，更實施當沖保證金減半的措施，使得入場的所需資金大幅減少。

台股期貨採用保證金交易制度，因此槓桿比股票高出許多，約在10到20倍之間；若使用當沖保證金減半，槓桿更高達20到40倍。高報酬的特性，吸引了許多交易人入市；其背後的高風險，也讓大部分的人血本無歸[3]。人們總是高估了自己對於虧損的承受度，殊不知想要在這市場長久地存活下去，最重要的便是資金控管及嚴格的紀律。程式交易在這方面做得不錯，因為它不像人，交易時不帶情緒，面對停利出場或是虧損認賠，都不會皺一下眉頭。

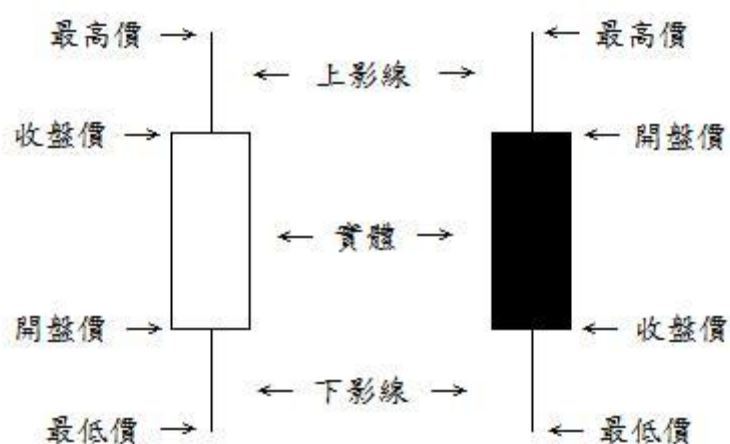
2-2 當日沖銷 (Intra-day Trading)

當日沖銷(簡稱當沖)意指交易人於同一交易日買進賣出，賺取其中的差價。不留倉到隔天，避開隔日跳空的風險。但由於當日的波動度有限，損益較小，因此交易成本相對上比波段交易高出許多。

台股期貨的當沖保證金減半制度，使交易人需要的投資本錢更少。不留倉、每日漲跌幅限制以及交易日最後15分鐘強制平倉，更降低了許多風險，對於資金較小的交易人，或許是個不錯的選擇。

2-3 K線 (Candlestick)

K線源自於十八世紀日本米市商人本間宗久記錄米市行情時所繪製的圖形，其包含了一段時間內的開盤價、收盤價、最高價以及最低價等四個關鍵價位，由於易讀易懂，現已成為記錄價格的普遍使用工具。其詳細圖解如下圖：



(圖一) K線

若收盤價比開盤價高，則為收紅，實體以白色表示，稱為陽K線。若收盤價比開盤價低，則為收黑，實體以黑色表示，稱為陰K線。

不同的K線圖，其隱含了不同的意義，代表了一段時間內價格的關鍵變動。不僅可以由單一K線來研判，還可以由雙K線甚至多K線的組合型態來推斷，因此K線為代表價格的一重大指標。

2-4 機器學習與人工智慧 (Machine Learning and AI)

電腦的發明著實大大地改變了人們的生活，不斷推陳出新的硬體規格，雖無法讓電腦像人類般思考與行動，卻可以使它的運算能力以及記憶能力遠遠超出人類所及。

人們利用其強大的運算記憶能力來設計系統，由系統代替人類從以往的經驗或資料中學習與統計，不僅省下了大量的時間，也解決了不少問題。在股價或期

貨市場方面，不管是傳統的符號式人工智慧(symbolic AI)或是較新的統計式人工智慧(statistical AI)，都有不少文獻將之使用於股價或期貨市場。

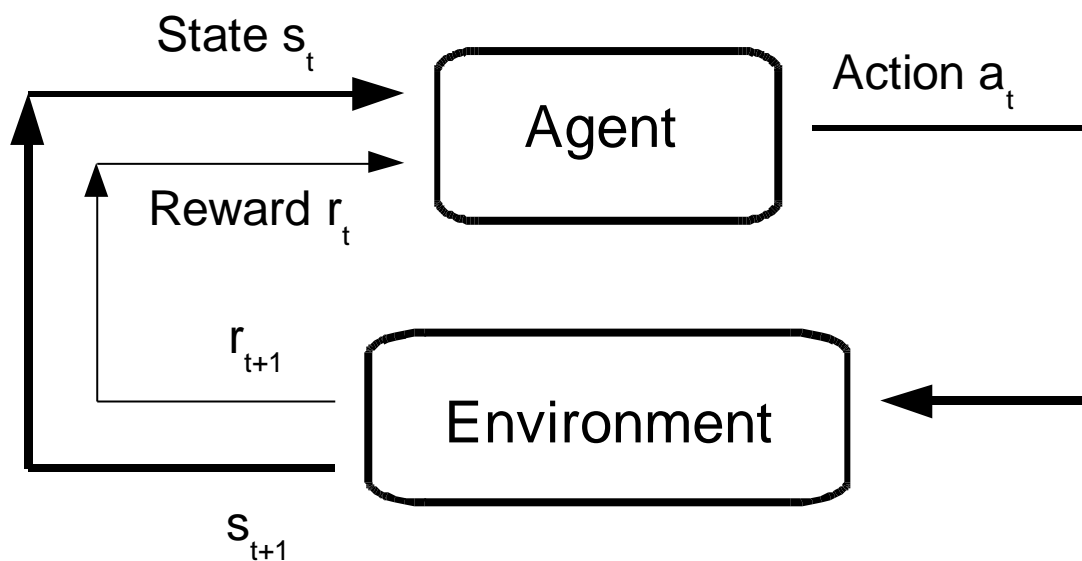
本研究使用的是增強式學習法(reinforcement learning)，屬於統計式人工智慧方法，底下將對其做簡略的介紹。

2-5 增強式學習法 (Reinforcement Learning)

如同一個新生兒在探索環境時，藉由每次動作後，環境所提供的回饋訊息，慢慢地自行採掘出他應該採取怎麼樣的動作才能得到他想要的結果。

在增強式學習法中，學習代理人必須去嘗試各種可能的動作，而我們設定的價值觀會使環境針對不同的動作會傳回不同的報償。好的動作會被獎勵，壞的動作會被懲罰，環境藉由如此的增強訊號來訓練學習代理人下次遇到類似的情形時該如何反應。

學習代理人(learning agent)執行一個動作(action)之後，評估函數(evaluating function)或稱獎勵函數(reward function)會根據學習代理人所執行的動作產生報償(reward)，影響價值函數(value function)，以調整學習代理人遇到此狀態(state)的因應措施。藉由與環境的互動，學習代理人的目標是找出一組原則(policy)，來將整體收益(return)給最大化；意即在每一種狀態下去執行最有價值的動作。增強式學習法的架構示意圖如下：



(圖二) 增強式學習法

價值函數用於估測某個狀態的良好程度，或是某個狀態下採取某個動作的良好程度，而良好程度可視為未來預期能得到的報償。因此定義了兩種價值函數，狀態-價值函數(state-value function)和動作-價值函數(action-value function)，可依照不同的環境或系統特性來選擇使用[2][4]。

2-6 Q-learning

Q-learning 為一種實現增強式學習的演算法，其記錄了每個狀態下所有可能動作執行後預期能得到的報償，而系統在每個狀態下會採取預期報償最高的動作。

每個狀態下，每次執行動作後，系統會將執行動作後得到的立即報償和執行動作後到達下一狀態在未來預期能得到的報償來更新此狀態-動作配對的預期報償。

$$Q(s_t, a_t) \leftarrow Q(s_t, a_t) + \alpha[r_{t+1} + \gamma \max_a Q(s_{t+1}, a) - Q(s_t, a_t)]$$

時間點 t ：狀態 s_t ，執行動作 a_t 。

時間點 $t+1$ ：剛剛執行動作後的立即報償 r_{t+1} ，並且到達狀態 s_{t+1} 。

$Q(s_t, a_t)$ ：狀態 s_t 下，採取動作 a_t 的預期報償。

γ ：折扣率。

α ：控制收斂速度的學習率(要給剛得到的立即報償多大的權重)。

Q-learning 學習了所有動作的預期報償，使系統擁有較佳的探索能力，因為它不只被限制於遵循目前的策略，而可選擇作所有可能執行的動作，因此有機會探索到比目前更佳策略。

另一方面，傳統實作上是以一個表格(Q-table)去儲存每一個狀態-動作配對的預期報償，系統必須耗費不少的記憶體去儲存資訊；且當狀態-動作空間非常巨大或是連續時，此方法會變得難以實踐[2][4]。

第三章 研究設計

3-1 研究背景

期貨交易是一場艱困的遊戲，其零和遊戲的特性代表著有人獲利就會有人虧損。由於參與者大多是人[3]，難免會有心理情結，因此造就了市場行為，形成了壓力或支撐這類實務上存在的傾向。情緒會使人無法自己，但不論是新手，或只是一時草率行事，還是風險入迷者，甚至是無法自拔的賭徒，市場一向不會手軟，總是會對情緒性的交易殘酷地給予重擊。

因此，程式交易逐漸地變得吸引人，因為它恰好避開了人性這方面的弱點。而期貨市場在研究的領域中，大多是以類神經網路(neural network)、灰理論(grey theory)、模糊集合(fuzzy set)、向量自我迴歸(vector auto-regression)為大宗；另外也不乏有支援向量機(support vector machines, SVM)、AdaBoost等等。但增強式學習法(reinforcement learning)在國內文獻中，尚無應用在臺股期貨市場上的先例。

3-2 研究目的

一般研究都是直接對這個市場學習，而最終得到一個策略原則。既然增強式學習法是從一個隨意的原則開始，那本研究便從一個自訂的策略原則開始，利用增強式學習法的特性，來改善原有的策略，使其能夠更貼近市場。

市場的八二法則說明了能在這市場賺錢的人實在少之又少[3]；換句話說，只要有獲利，就已經是贏了不少人。想要一夕致富之類這種貪婪的想法，往往只會更適得其反。剛進市場的新手，也幾乎都會繳上一筆「學費」來體驗市場的殘酷。本系統目標針對較小額的資金，希望提供交易人在可接受的風險下，能夠獲得些許報酬的機會。

此外，期貨是一個極其複雜的市場，許多人往往被繁雜的市場變化以及分歧的資訊搞得六神無主。個人則是希望能將這系統盡量保持簡單，一方面是此系統的目的並不在於完整地預測整體走勢，而只是短期的波動；另一方面，如此在盤中訊號出現時，系統能夠快速地決定要作順勢或是逆勢交易。

3-3 研究方法

3-3-1 假設

1. 交易成本(包含手續費與交易稅)為買賣一口雙邊共 1.5 點。
2. 市場具完整的流動性，意即我們皆可以成交在想要的價格。
3. 可以直接且快速地向期交所下單，且我們的下單不對市場造成影響。
4. 所有交易皆為當日沖銷。
5. 每個交易日從 8:45 開始即可建倉，12:20 後不再建倉，13:20 強制平倉。

3-3-2 資料來源

本研究採用由臺灣期貨交易所提供的每筆成交資訊，並且即時在系統中轉換成一分鐘K線，以模擬實際交易時的情形。訓練期間為 2003 年 1 月 2 日到 2007 年 12 月 31 日共 1241 個交易日。檢測期間為 2008 年 1 月到 2009 年 5 月，共 342 個交易日。

3-3-3 學習方法

使用增強式學習法中的 Q-learning 演算法，起始策略為一自訂的短線通道突破系統，採用順勢交易，示意圖如(圖三)。另外有一較長線的壓力或支撐，作為狀態的判定，示意圖如(圖四)。而增強式學習的要素列舉如下：

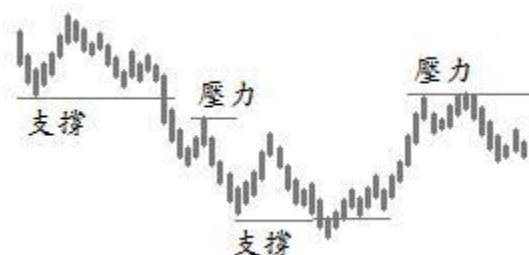
- 狀態(state)：
1. 作多價位減前一長線壓力(前一長線支撐減作空價位)
 2. 原本交易(多、空)與目前長線方向(上、下)是否一致

動作(action)： 作多、作空。

報償(reward)： 該筆交易後的獲利或虧損點數。



(圖三) 通道突破系統示意圖



(圖四) 較長線的壓力支撐示意圖

一般而言，狀態的選定與價值函數的設定對於增強式學習法的影響最大。在此，價值函數便是未來的預期報償；而狀態的選定就如同監督式學習法 (supervised learning) 中的特徵(feature)選定一般，極其重要。在希望保持系統簡潔以及 Q-learning 的特性影響下，我們決定盡量減少狀態個數，只留下少數具影響力的狀態。

3-3-4 系統實作

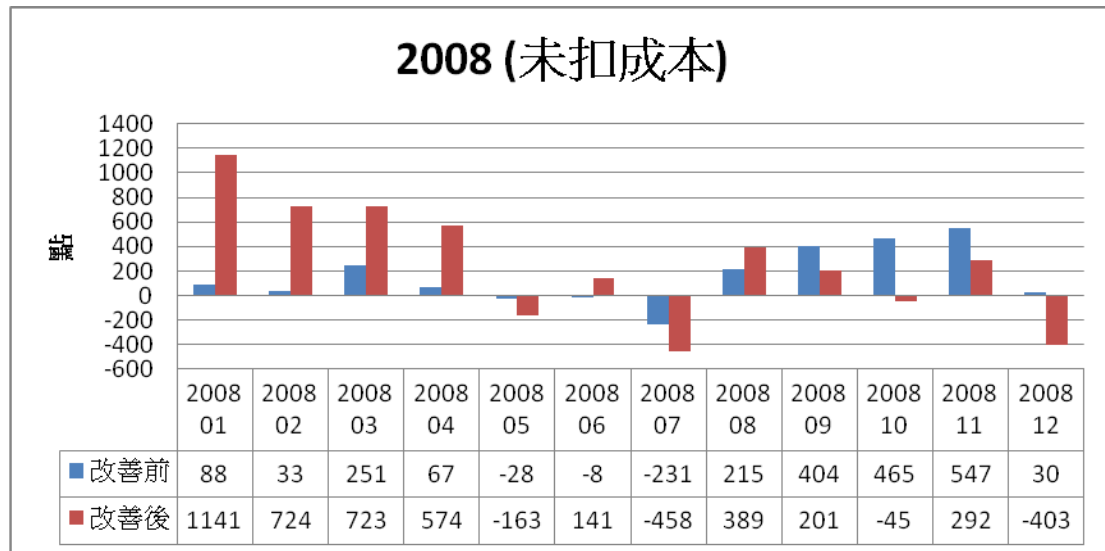
系統使用 C#語言撰寫，可切換選擇歷史回測或是盤中接收資訊。對於歷史資料用檔案載入的方式，對於即時盤中資訊則利用 DDE (Dynamic Data Exchange) 來接收，並可利用券商提供的 API 來進行下單動作，實現程式交易。使用一分鐘 K線做策略判定，並監測 tick 資訊，在有突破線的成交價產生後，馬上依照策略進行動作。



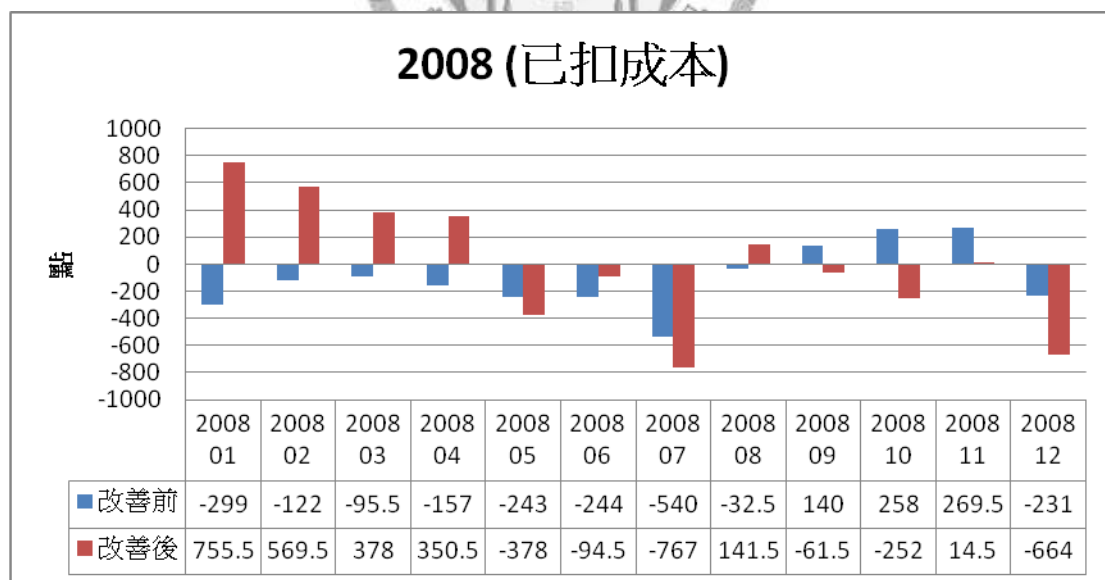
第四章 實驗結果

4-1 方法一

方法一為與訓練方式相同的停利停損設定，不斷交易。

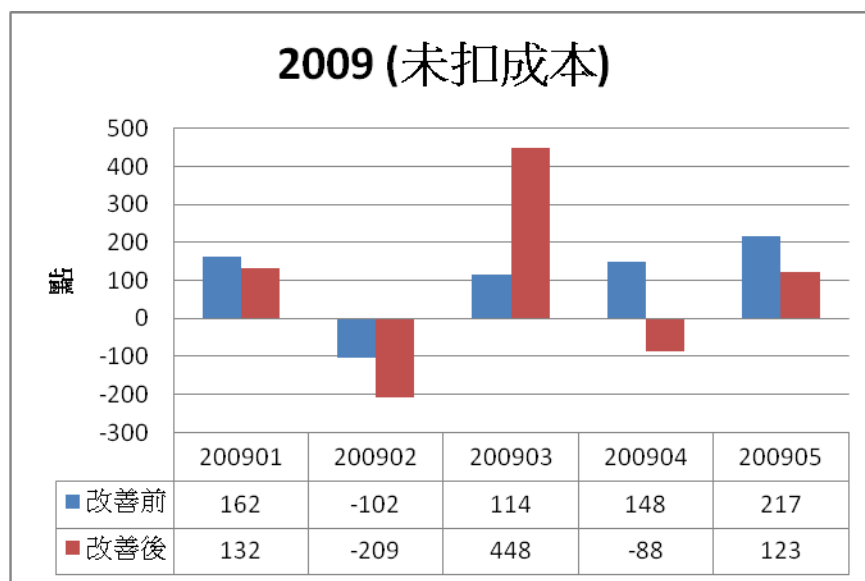


(圖五) 方法一 2008 年未扣成本前，改善前計 1833 點，改善後計 3116 點

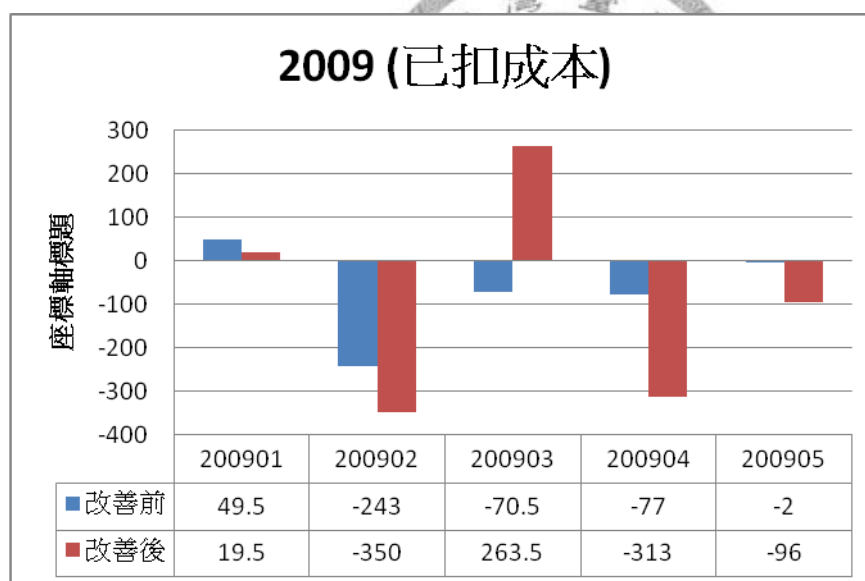


(圖六) 方法一 2008 年已扣成本後，改善前計-1296.5 點，改善後計-7.5 點

2008 年，改善前總計進場 2085 次，勝 1394 次，敗 691 次，勝率約 66.86%。
改善後總計進場 2085 次，勝 1440 次，敗 645 次，勝率約 69.06%。



(圖七) 方法一 2009 年未扣成本前，改善前計 539 點，改善後計 406 點

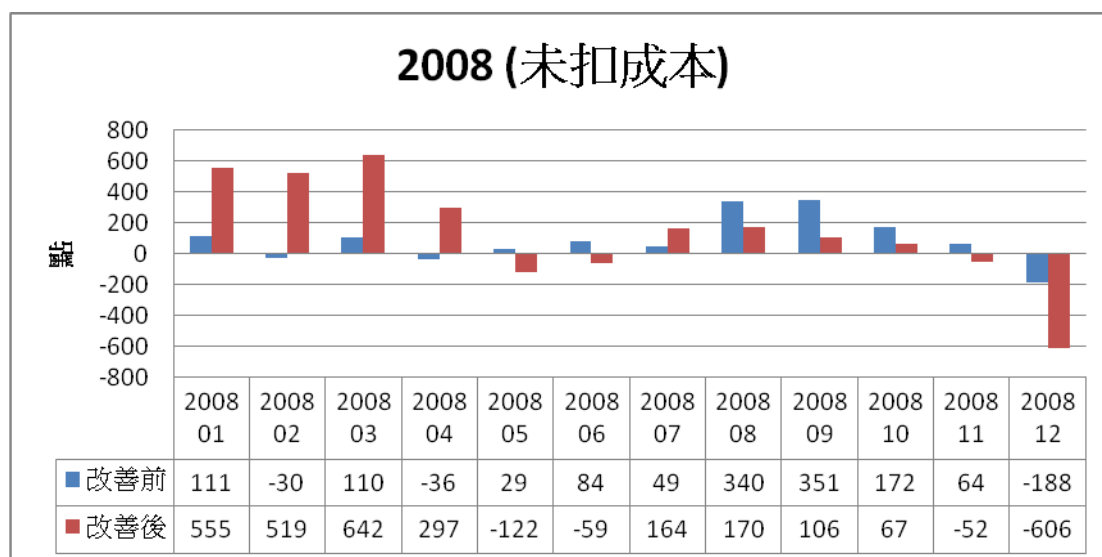


(圖八) 方法一 2009 年已扣成本後，改善前計-343 點，改善後計-476 點

2009 年 1 到 5 月，改善前總計進場 588 次，勝 380 次，敗 208 次，勝率約 64.63%。改善後總計進場 588 次，勝 374 次，敗 214 次，勝率約 63.61%。

4-2 方法二

方法二依然為與訓練方式相同的停利停損設定，但交易日出現敗績後則當日不再進場。

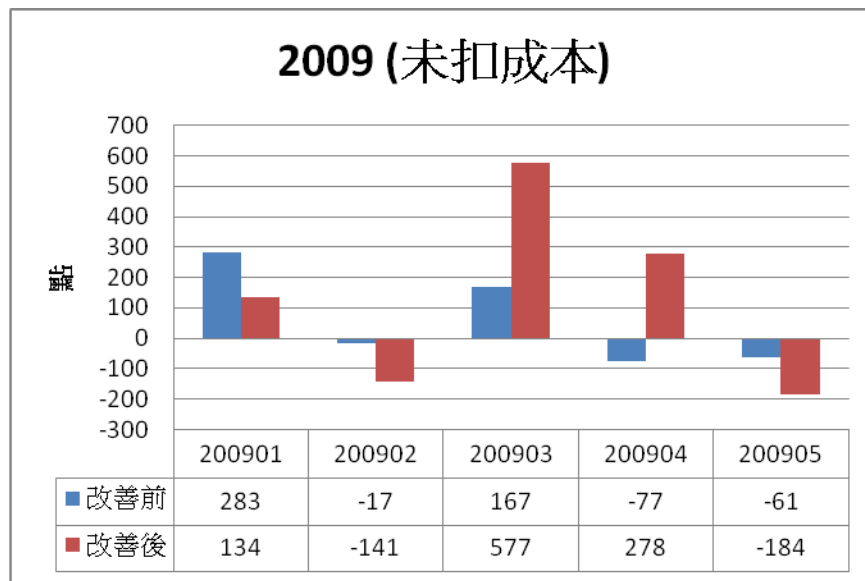


(圖九) 方法二 2008 年未扣成本前，改善前計 1056 點，改善後計 1681 點

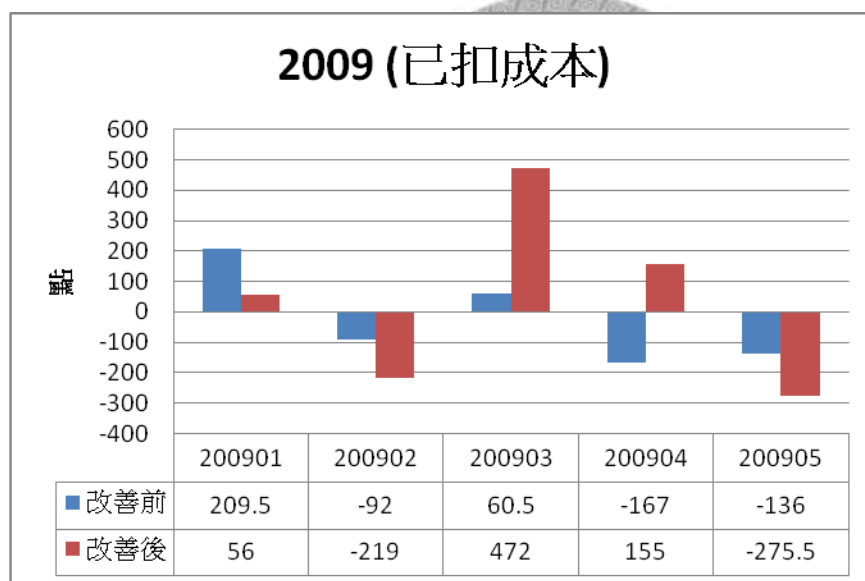


(圖十) 方法二 2008 年已扣成本後，改善前計-118 點，改善後計 284 點

2008 年。改善前總計進場 782 次，勝 529 次，敗 253 次，勝率約 67.65%。
改善後總計進場 926 次，勝 648 次，敗 278 次，勝率約 69.98%。



(圖十一) 方法二 2009 年未扣成本前，改善前計 295 點，改善後計 664 點

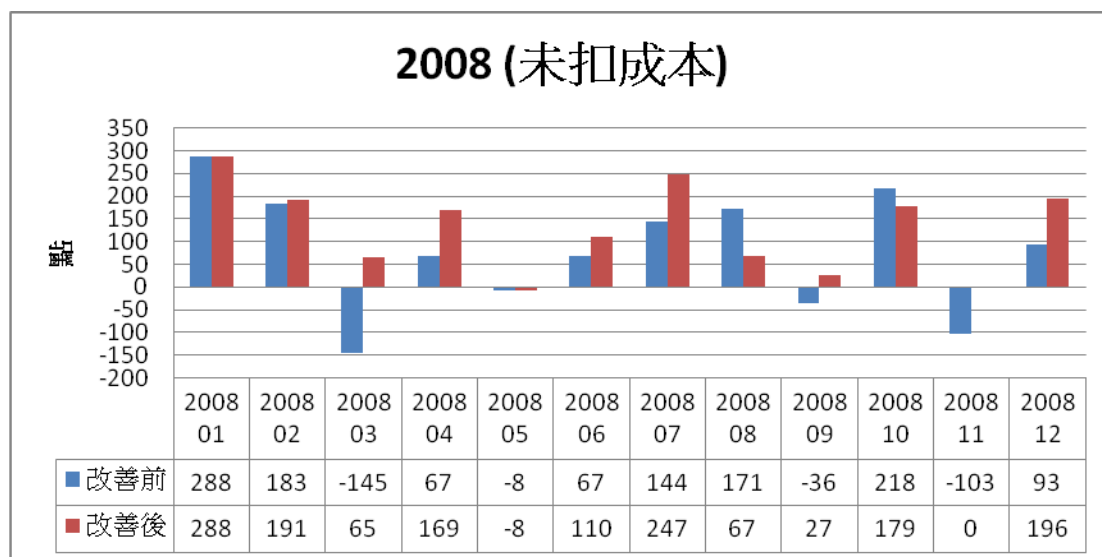


(圖十二) 方法二 2009 年已扣成本後，改善前計-125 點，改善後計 188.5 點

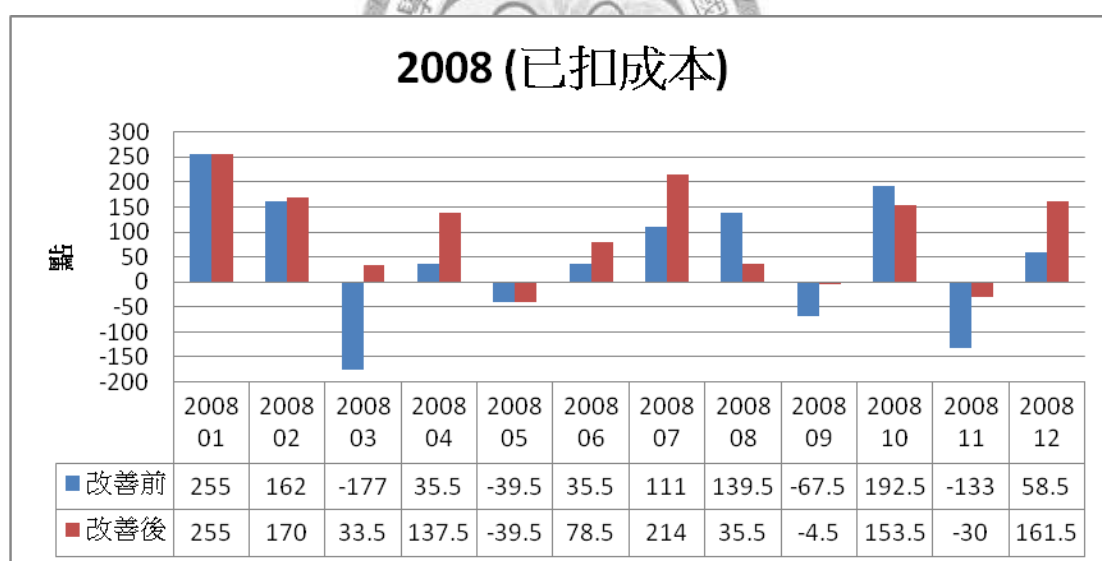
2009 年 1 到 5 月。改善前總計進場 278 次，勝 187 次，敗 91 次，勝率約 67.27%。
改善後總計進場 317 次，勝 214 次，敗 103 次，勝率約 67.51%。

4-3 方法三

方法三採取固定數值的停損停利，且每個交易日只交易第一次的訊號。

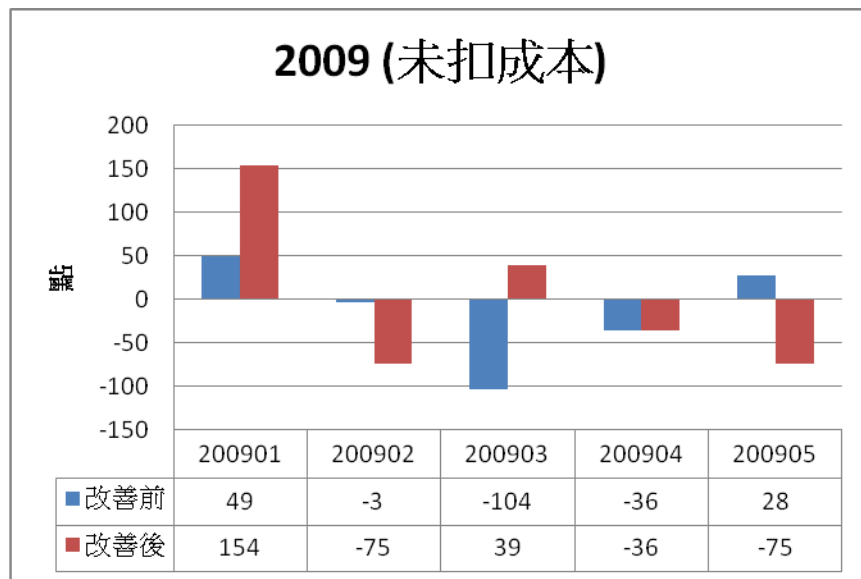


(圖十三) 方法三 2008 年未扣成本前，改善前計 939 點，改善後計 1531 點

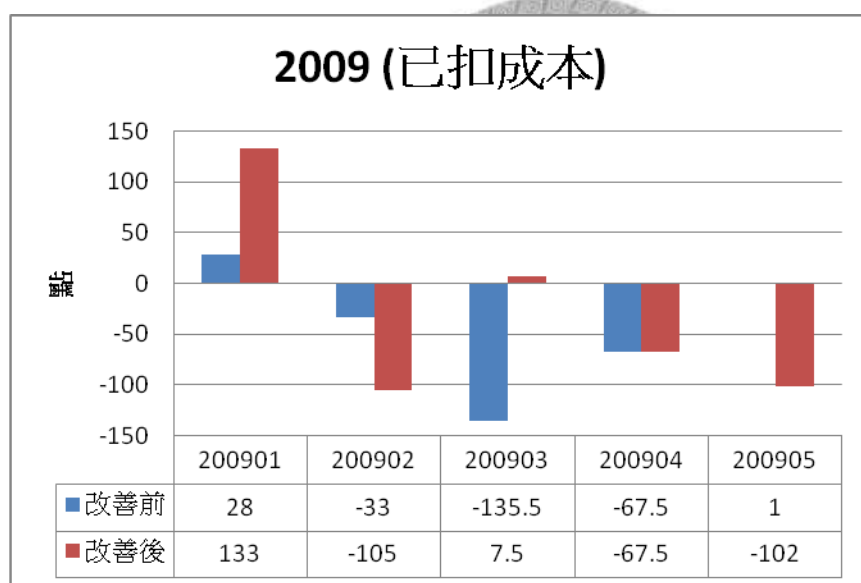


(圖十四) 方法三 2008 年已扣成本後，改善前計 572.5 點，改善後計 1165 點

2008 年。改善前總計進場 244 次，勝 221 次，敗 23 次，勝率約 90.57%。改善後總計進場 244 次，勝 225 次，敗 19 次，勝率約 92.21%。



(圖十五) 方法三 2009 年未扣成本前，改善前計-66 點，改善後計 7 點



(圖十六) 方法三 2009 年已扣成本後，改善前計-207 點，改善後計-134 點

2009 年 1 到 5 月。改善前總計進場 94 次，勝 77 次，敗 17 次，勝率約 81.91%。
改善後總計進場 94 次，勝 79 次，敗 15 次，勝率約 84.04%。

第五章 數據分析及結論

5-1 數據分析

檢測期間從 2008 年 1 月到 2009 年 5 月，總共 342 個交易日。

5-1-1 方法一

| 方法一 | 改善前 | 改善後 | 後 - 前 |
|------|--------|--------|-------|
| 進場次數 | 2673 | 2673 | 0 |
| 獲利次數 | 1774 | 1814 | 40 |
| 虧損次數 | 899 | 859 | -40 |
| 勝率 | 66.37% | 67.86% | 1.49% |
| 未扣成本 | 2372 | 3522 | 1150 |
| 已扣成本 | -1304 | -819 | 485 |

(表一) 方法一的改善前後比較表

方法一在 2673 次進場中，經由改善後，將其中 40 次的交易轉虧為盈，避免掉損失並獲得利益，勝率提高約 1.49%。惟其交易的次數相當頻繁，在扣掉交易成本後，其總損益依然皆為負值，不過仍比改善前增加了 485 點。

5-1-2 方法二

| 方法二 | 改善前 | 改善後 | 後 - 前 |
|------|--------|--------|-------|
| 進場次數 | 1060 | 1243 | 183 |
| 獲利次數 | 716 | 862 | 146 |
| 虧損次數 | 344 | 381 | 37 |
| 勝率 | 67.55% | 69.35% | 1.80% |
| 未扣成本 | 1351 | 2345 | 994 |
| 已扣成本 | -243 | 472.5 | 715.5 |

(表二) 方法二的改善前後比較表

方法二在當日有敗績後便不再進場。在改善後增加了 183 次進場，足見其獲勝次數的提高，勝率提高約 1.8%。由於減少了交易次數，扣除成本在改善後其損益已轉為正值，並比改善前增加 715.5 點。

5-1-3 方法三

| 方法三 | 改善前 | 改善後 | 後 - 前 |
|------|--------|--------|-------|
| 進場次數 | 338 | 338 | 0 |
| 獲利次數 | 298 | 304 | 6 |
| 虧損次數 | 40 | 34 | -6 |
| 勝率 | 88.17% | 89.94% | 1.77% |
| 未扣成本 | 873 | 1538 | 665 |
| 已扣成本 | 365.5 | 1031 | 665.5 |

(表三) 方法三的改善前後比較表

方法三只做每日的第一次進場，在 342 交易日中進場 338 次，有 4 天不符進場條件。經由改善後，有 6 天的進場轉虧為盈，勝率提高約 1.77%。由於交易次數少，因此扣除成本後的損益依然為正，改善後增加 665.5 點。

5-2 結論

本論文提出以增強式學習法做為改善手段，提供已有策略的程式交易者，在對其核心策略的了解之下，經過適當的設計，可利用增強式學習法扮演較複雜的篩選器角色，來改善其策略中的不足之處。

本研究實作出的交易系統，可透過 DDE 接收即時盤中資訊，並利用券商提供的 API 來下單，達到程式交易的目的。

5-3 未來展望

1. 增強式學習法的狀態訂定或是演算法實作設定的修改。
2. 使用其他的機器學習方法，例如監督式學習法或類神經網路。
3. 核心策略種類及用以改善的條件變更。
4. 應用在波段交易上。
5. 使用在其他期貨商品或金融商品上。

參考文獻

- [1] Curtis M. Faith, *Way of the Turtle: The Secret Methods that Turned Ordinary People into Legendary Traders*. McGraw-Hill Education, United States, 2007.
- [2] Richard S. Sutton and Andrew G. Barto, *Reinforcement Learning: An Introduction*. MIT Press, Cambridge, MA, 1998.
- [3] 吳承康、簡美雲、陳嫻瑩，“臺灣期貨市場交易人之行為研究”，臺灣期貨交易所 95 年度專案研究報告提要表，2006。
- [4] 陳律宇，“以自我組織特徵映射圖為基礎之模糊系統實作連續性 Q-learning”，國立中央大學資訊工程研究所碩士論文，2006。
- [5] 林典南，“使用 AdaBoost 之台股指數期貨當沖交易系統”，國立臺灣大學資訊網路及多媒體研究所碩士論文，2008。