

東吳大學巨量資料管理學院
資料科學系碩士在職專班碩士論文

指導教授：邱嘉豪 博士

多因子技術指標和機器學習的
市場活絡度預測與當沖研究

Multi-Factor Technical Indicators and Machine
Learning for Market Liquidity Forecasting and
Intraday Trading Research

研究生：林昌諒 撰

中華民國 114 年 4 月

目錄

第一章 緒論.....	4
第一節 背景	4
第二節 研究動機	6
第三節 研究目的	6
第四節 研究問題	7
第二章 文獻探討	9
第一節 道氏理論、效率市場假說與行為財務學	9
第二節 技術分析	11
第三節 當沖交易	19
第四節 異常偵測	21
第三章 研究方法	25
第一節 研究架構	25
第二節 資料處理及資料集劃分	26
第三節 交易成本設定	27
第四節 交易策略設計	28
第五節 交易模型設計	31
第六節 交易績效衡量	34

第一章 緒論

第一節 背景

獲取超額報酬一直是金融市場中的投資人關注的焦點，近年來台灣股票、期貨市場受惠於電子業蓬勃發展，呈現正成長，並逐年突破新高，然而金融市場趨勢瞬息萬變，Mandelbrot(1963)、Fama(1965)認為金融市場非常態分配，而是具有高峰(leptokurtic)與厚尾(fat tail)的特性，即發生極端價格的機率高於常態分配的假設。以台灣的歷史事件為例，1999年921大地震、2000年的網路泡沫破裂、2008年美國次級房貸危機引發金融海嘯、2020年則因COVID-19疫情爆發，導致台灣金融市場市場恐慌性拋售使得股票、期貨價格劇烈震盪，因此直至現今風險控制亦已是金融界普遍的觀點和共識，這種理論上發生機率極低，但實際上卻又不斷發生的事件，即稱為金融市場的「厚尾風險」。

從風險控的角度來看，當沖交易提供了一個有效的途徑來迴避隔夜風險。當沖交易指的是在同一交易日內完成股票或期貨的買入和賣出，以避免持有股票或期貨過夜，從而減少市場收盤後的突發事件或新聞所引發的波動風險。這種方式不僅能減少持有期間的不確定性，也能夠在市場未來走勢不明的情況下有效控制風險。

行為財務學(Behavioral Finance)的研究指出投資行為經常受到情緒與認知偏誤的影響，進而導致非理性的判斷與錯誤的決策，這些偏誤廣泛存在於市場參與者中，是導致大多數投資人無法穩定獲利的根本原因之一。例如，「錯置效果(Disposition Effect)」反映了投資人傾向於提早賣出獲利部位以鎖定小利，卻對虧損部位猶豫不決、遲遲不願停損，導致長期而言「賺小賠大」的情況(Shefrin & Statman, 1985)；而「損失厭惡(Loss Aversion)」則顯示出人們對損失的敏感度遠高於獲利所帶來的快樂，使得許多投資人即使面對明顯的不利情勢，仍選擇死守虧損部位(Kahneman &

Tversky, 1979)。此外，投資人也常因過度自信（Overconfidence）而高估自己的判斷能力，進而進行頻繁交易或高風險集中投資，而「從眾行為（Herding Behavior）」與「後見之明偏誤（Hindsight Bias）」則進一步削弱了獨立判斷與風險意識，使得投資決策更易受到外在資訊與市場情緒影響（DeBondt & Thaler, 1985; Barberis & Thaler, 2003）。

在此背景下，程式交易(Algorithmic Trading)提供了一種有效對抗人性弱點的解方。與人為操作不同，程式化的交易系統不會因市場波動而恐慌，也不會因短期損益而貪婪或猶豫，其決策完全建立在統計模型與數據分析的基礎上（Lo, 2004）。此外，透過歷史資料的回測與模擬，程式交易能夠事先驗證策略在各種市場條件下的表現，有助於提升策略的穩健性與可靠性，並且可納入完整的風險控制設計，如停損點、停利點與部位控管等，降低單筆交易失誤所造成的資金損失（Chan, 2009）。程式交易能夠透過事先明確定義的邏輯自動化執行買賣決策，排除情緒干擾，確保交易過程始終遵循一致的策略與風控規則（Aldridge, 2013），為系統化、紀律化與數據化提升投資績效與風險管理效能的普遍操作工具。然而根據Park與Irwin（2007）對近百篇與技術分析相關的論文進行回顧後發現，其中近6成支持技術分析具有獲利能力，但亦有4成認為技術分析無法帶來獲利或呈現不確定結果，顯示程式交易雖然避開了人性的弱點，但在獲利能力的穩定性仍有提升空間。

隨著計算能力的提升與大數據技術的成熟，人工智慧（Artificial Intelligence, AI）與機器學習（Machine Learning, ML）技術已被廣泛應用於各個產業領域。在金融領域中，機器學習模型被廣泛運用於量化交易、風險評估與詐欺偵測，以提升決策效率與準確性。本研究即希望可以藉由機器學習輔助的交易方式替代人工交易，並推升程式交易之獲利能力。

第二節 研究動機

獲取超額報酬並降低風險，一直是金融市場中的重要研究課題與投資人關注的焦點。透過程式交易建構交易策略，由程式買賣台股，結果顯示自訂交易策略在期貨市場上可獲得報酬(朱健榮，2008)。回溯測試和自動化交易是程式交易的2大特點，透過程式與資料，以歷史資料進行策略回溯測試，驗證策略是否能夠在過去的市場上獲得利潤，能夠推論策略的可行性，而程式依照所設計之策略規則進行買賣，不受外來因素所干預以紀律進出場，但同時缺點則是不具備投資人的直覺反應和應變能力，無法應對突如其來的事件(紀錦城，2021)。

傳統技術指標大多建立在固定公式與線性假設上，並仰賴人工觀察來判斷市場規律。然而金融市場行為具有高度複雜的特徵，難以藉由傳統方式完全掌握，使得技術指標在特定情境下有效，而在其他情境中則可能失靈。隨著資料科學的發展，機器學習成為捕捉潛藏於大量數據中的模式與異常行為的重要工具，林典南(2008)以 AdaBoost 為辨識方法，提出一套低風險且穩定報酬的交易系統，提供期貨交易者參考；林敬斌(2009)則運用增強式學習技術改善既有交易策略；林洋慶(2022)運用 XGBoost(extreme gradient boosting)機器學習演算法，提出「三次交易順勢策略」，報酬率優於買進持有及其他常見的技術指標交易策略。鍾嘉元(2024)運用 Times Net 模型能預測異常交易點，搭配交易策略為投資者提供更加穩健的當沖方式。以上研究成果顯示了機器學習技術優化金融市場交易策略的潛力，有鑑於此，本研究提出下研究動機：

1. 藉由長期歷史資料回測，驗證量化交易策略績效的有效性及穩定性。
2. 比較單純交易策略以及加入機器學習輔助交易後策略績效的改變。

第三節 研究目的

市場價短期格波動主要由即時的市場情緒、供需變動以及短期技術指標所驅動，而非基本面因素，所以當沖交易主要依賴技術面進行操作。常見的技術指標包括移動平均線、相對強弱指數、布林帶及CDP指標等，能幫助交易者判斷買入或賣出的時機。

近年來人工智慧與機器學習技術達成了人們過去難以達成的目標，而異常檢測 (Anomaly Detection) 也是該領域的其中一個研究分支，其核心目的是識別資料中偏離正常模式的異常樣本，相較於傳統時間序列模型具有處理多維度及非線性資料的優勢，被廣泛應用於多元產業領域。在金融科技方面，異常檢測常用於信用卡詐欺偵測、洗錢行為以及異常交易點的識別，有助於金融機構即時掌握風險與異常波動。Rudkiewicz與Hernes(2023)運用深度學習自動編碼器(Auto Encoder)異常檢測模型於建構加密貨幣日內交易策略並指出能幫助獲利，Kim等(2023)使用生成式對抗網路(Generative Adversarial Network)於標準普爾500指數中的產業類股指數進行異常檢測，研究結果顯示能增進投資風險管理效能。黃俊朋(2025)以基於預測的LSTM異常檢測模型提供股價偏離程度的異常分數發展交易策略，實證結果顯示，以固定決策門檻產生交易訊號的方式，常有過早出場而錯失後續上漲機會的情形，以異常分數的移動平均線交叉產生交易訊號的方式，應用在上證指數震盪或空頭走勢標的時，能有效獲利並降低風險。

本研究希望結合程式交易與機器學習技術，驗證機器學習方法在金融市場中的獲利潛力作為交易人降低交易風險的工具，並預期達成以下研究目的：

1. 建立能夠自動判斷買入或賣出時機的交易策略。
2. 以量化的方式評價交易策略的獲利能力及風險，比較多種策略的績效。
3. 運用機器學習技術提升策略獲利能力或降風險。

第四節 研究問題

本研究及於上述動機與目的，提出以下研究問題，以探討傳統量化交易策略與機器學習模型於期貨當沖交易中之獲利的有效性：

1. 建構獲利能力優於大盤或買進持有的當沖(隔日沖)量化交易策略。

以技術分析的CDP指標設計可以模擬當沖交易情境的自動交易策略，並透過回溯測試大規模歷史資料，檢驗CDP指標策略的長期獲利能力。

2. 建構機器學習模型，並探索機器學習模型用於當沖交易決策可行性與效能。

以機器學習替代技術指標模擬當沖交易情境的自動交易策略，並透過回溯測試大規模歷史資料，檢驗機器學習策略的長期獲利能力是否能贏過技術指標策略。

3. 比較以機器學習輔助相較於單純量化交易的策略表現。

運用機器學習結合技術指標策略共同判斷，找出技術指標策略未捕捉到的市場特性，建立篩選機制，並評估其對交易績效的影響。

第二章 文獻探討

第一節 道氏理論、效率市場假說與行為財務學

一、道氏理論

道氏理論 (Dow Theory) 是技術分析的奠基理論，最初由 Charles H. Dow 於 1900 至 1902 年間在《華爾街日報》(The Wall Street Journal) 的連載評論中提出。其核心觀點認為，透過觀察價格的運行模式可判斷市場趨勢，進而制定交易決策。隨後，Hamilton (1921) 在其著作《The Stock Market Barometer》中進一步發展並系統化道氏理論，而 Rhea (1932) 則進一步理論化之，並在《The Dow Theory》一書中正式整理道氏理論的六大基本原則，使其成為技術分析的核心理論架構之一。

道氏理論強調，市場價格反映了所有已知資訊，包括財務數據、經濟指標、政治事件與市場情緒等 (Rhea, 1932)。因此，市場價格的變動已經充分涵蓋了影響供需的各種因素。基於此觀點，技術分析者無需深入探究基本面資訊，而應專注於價格趨勢與圖形變化的識別與解讀。

二、效率市場假說

交易策略是否能夠長期超越市場績效，一直是金融學界關注與探討的核心議題。對此問題，學術界普遍持保留態度。其中 Fama (1970) 所提出的效率市場假說 (Efficient Market Hypothesis, EMH) 即為最具影響力的理論之一。該理論主張，市場中的投資者是理性的，且資訊能夠迅速且無障礙地傳播至所有市場參與者，因此市場價格已充分反映所有可得資訊。基於此假說，無論是透過歷史價格、公開資

訊，甚至是內部資訊，投資者皆無法持續獲得超額報酬。因此，相較於費時且成本高昂的主動型投資策略，效率市場假說建議投資人應選擇低成本的被動型投資方式，以達到長期穩健的報酬表現（Fama， 1970）。

三、行為財務學

然而在真實市場，仍有效率市場假說（Efficient Market Hypothesis，EMH）無法解釋的現象，這些現象違背了市場總是反映所有公開信息並且投資者理性行為的假設，稱為市場異象。行為財務學則試圖解釋這些異象，強調心理因素和非理性行為如何影響市場價格，使其偏離效率市場假說的預測：

1. 投資人有非理性行為：

Kahneman與Riepe（1998）指出，投資人無法做出理性決策的原因可歸納為三個主要因素。首先是對風險的態度，投資人常在面對風險時表現出非理性偏好。其次，非貝氏法則（Non-Bayesian）所引發的預期形成過程使得投資人根據有限的資訊進行不完全或偏誤的預測。第三，問題框架（Framing Effect）會影響投資人的決策方式，特別是當問題的呈現方式改變時，投資人會作出不同的反應。這些心理學因素使得投資人在面對風險與不確定性時，經常偏離理性行為模式，從而挑戰了效率市場假說中關於投資人理性行為的假設。

2. 投資人的非理性行為並非隨機發生：

Kahneman與Tversky（1979）強調，非理性行為並非隨機發生，而是呈現某種模式或趨勢。這些行為往往朝著相同的方向發展，未必會相互抵銷。Shiller

(1984) 進一步指出，當這些非理性行為被社會化，或當群體中大多數投資者聽信相同的謠言時，市場中的非理性行為現象會更加明顯，並對市場價格產生重大影響。此外，Mullainathan與Thaler (2000) 則指出，由於學習正確評價的機會成本較高，且學習過程需要長時間的經驗積累，因此投資人難以及時修正錯誤的決策，這使得市場中存在結構性非理性行為，且此行為會長期存在。

3. 套利存在條件限制：

Shleifer與Vishny (1997) 以及Thaler (1999) 指出，套利活動本身受到多重條件的限制，使其無法完全發揮其理論上的效能。Mullainathan與Thaler (2000) 進一步闡述，儘管套利理論預測市場價格應該趨於平衡，實際操作中卻存在風險與不確定性。特別是在面對不完美的替代品和有限的投資期限時，套利者可能無法在預定的時間內獲得預期回報。這些因素使得套利活動在現實市場中未必能夠實現其預期效能，進一步證明了市場並非總是完全有效的。

換言之，行為財務學認為投資人仍然有機會設計出能夠超越市場績效的交易策略，從而獲得持續的超額回報。本研究的主要目標，即探討設計出有效的當沖交易策略，並檢視其在期貨市場中的表現。

第二節 技術分析

一、技術分析理論

除了行為財務學認為市場價格非隨機發生之外，Murphy(1999)在《金融市場技術分析》一書中指出，技術分析法建立在三個基本前提假設上。首先，市場行為已經反

映了一切信息，無論是基本面、政治還是經濟因素，價格走勢已包含所有已知與未知的信息。其次，價格會沿著趨勢運行，趨勢一旦形成，將持續運行，直到有明確的反轉信號出現。最後，歷史會重演，市場參與者的行為模式具有可預測性，因為人類的心理行為使歷史模式一再重複。這些基本前提構成了技術分析的核心基礎，被廣泛應用於金融市場分析與交易決策。因此透過對歷史資料進行回測或資料探勘，投資者可以從大量的技術指標中擷取有用資訊，以判斷市場趨勢，進而獲取超額報酬。彭煥淇（2014）透過市場輪廓理論與分群分類分析，以 1分鐘K線圖進行歸納，證明市場價格並非隨機漫步，而是會受到心理因素影響，亦支持技術分析在市場行為研究中的有效性。

技術分析主要透過歷史資料來分析股價的未來趨勢，有趨勢訊號明確，以及可供短期操作之特性。技術分析指標的種類繁多，不同指標適用於不同的市場趨勢預測，因而各種指標都有其偏好使用的投資人。儘管技術分析是否能穩定創造超額報酬在學術界尚存爭議，然而針對臺灣市場的多項實證研究結果顯示，運用技術指標建構之交易策略在特定期間內確實具備優於大盤的潛力。安芷誼（2005）針對1989年至2003年間的台股資料進行分析，發現採用移動平均線法可有效提升投資績效。黃冠華（2008）以1991年至2007年的台指期資料為樣本，指出在特定的均線週期下，技術指標策略能創造顯著的超額報酬。施岳宏（2013）則針對2009年至2012年間台灣50指數成分股進行研究，發現結合移動平均線與箱型策略（採停損不停利原則）的技術分析法，在報酬表現上優於單純長期持有策略。許嘉文（2018）分析2003年至2018年台股指數期貨，結果顯示無論是簡單移動平均（SMA）、加權移動平均（WMA）或是移動平均收斂背離指標（MACD），在多頭市場條件下均可達到正向報酬。

綜合以上對不同時期台股或台指期研究結果，運用技術指標建構交易策略有利於提高較報酬甚至勝過大盤，並有時間穩定性。

二、技術指標

本研究採用CDP技術指標為建構當沖交易策略的主要指標，以及其他技術指標如SMA、Bollinger Bands、RSI、BIAS、CCI、KD、MACD、DMI等為機器學習之特徵變數，以下就本研究採用技術指標的運算方式及常見用途進行說明。

1. CDP指標

CDP 指標是一種基於前一交易日價格計算的指標，適合日內交易者用於快速判斷當日關鍵價位，用於預測當日價格的支撐與阻力位，常用於短線當沖交易。應用前一營業日的最高價、最低價、及收盤價的計算與分析，將當日的股價變動範圍為五個等級，做為超短線進出的研判標準。計算公式如下：

(1) 首先求出昨日行情的CDP值：

$$CDP = (\text{昨日最高價} + \text{昨日最低價} + 2 \times \text{昨日收盤價}) / 4$$

(2) 再分別計算AH、NH、NL及AL：

$$AH(\text{最高值}) = CDP + (\text{昨日最高價} - \text{昨日最低價})$$

$$NH(\text{近高值}) = 2 \times CDP - \text{昨日最低價}$$

$$NL(\text{近低值}) = 2 \times CDP - \text{昨日最高價}$$

$$AL(\text{最低值}) = CDP - (\text{昨日最高價} - \text{昨日最低價})$$

當價格接近但未突破AH或NH時，視為阻力位，可能賣出或觀望；接近但未突破AL或NL時，視為支撐位，可能買入或止損。若價格突破AH，可能看漲；跌破AL，可能看跌。

2. SMA (Simple Moving Average, 簡單移動平均線)

SMA 是價格的簡單平均值，用於平滑價格、確認趨勢方向或作為支撐/阻力參考，是許多指標的基礎。計算公式如下：

$$SMA(n) = (\text{過去 } n \text{ 根K的收盤價總和}) / n$$

其中 n 為設定的周期數。

當價格上穿 SMA 表示看漲訊號，可能買入；下穿 SMA 表示看跌訊號，可能賣出。短期 SMA（如 10 根K）與長期 SMA（如 50 根K）交叉時，黃金交叉（短期上穿長期）看漲，死叉（短期下穿長期）看跌。

3. Bollinger Bands (布林帶)

布林帶由中軌、上軌和下軌組成，其概念是運用平均值、標準差的概念，用於衡量市場價格的相對高低位置，進而找出市場的價格突破點、反轉點。計算公式如下：

(1) 首先求出Middle Band：

$$\text{Middle Band(中軌)} = SMA(n)$$

(2) 再分別計算Upper Band和Lower Band：

$$\text{Upper Band(上軌)} = \text{Middle Band} + k \times \sigma$$

$$\text{Lower Band(下軌)} = \text{Middle Band} - k \times \sigma$$

n 通常設20、 K 通常設2、 σ = 過去 n 天的收盤價標準差

當價格觸碰上軌表示可能超買，考慮賣出；觸碰下軌表示可能超賣，考慮買入。價格沿上軌上升表示強勢看漲，沿下軌下降表示強勢看跌。

4. RSI(Relative Strength Index， 相對強弱指數)

RSI指標用來衡量價格變動的強度，用於判斷超買或超賣狀態。計算公式如下：

(1) 首先求出RS（相對強度）：

$$RS = \text{平均上漲幅度} / \text{平均下跌幅度}$$

$$\text{平均上漲幅度} = \text{近}n\text{根}K\text{上漲價格的總和} / n$$

$$\text{平均下跌幅度} = \text{近}n\text{根}K\text{下跌價格的總和} / n$$

通常 $n = 14$ 。

(2) 再分別計算RSI：

$$RSI = 100 - [100 / (1 + RS)]$$

當 $RSI > 70$ 表示超買，可能下跌，考慮賣出； $RSI < 30$ 表示超賣，可能反彈，考慮買入。RSI 與價格背離（如價格創新高但 RSI 未創新高）可能預示趨勢反轉。RSI 在 40-60 區間波動時，市場可能處於無明顯趨勢狀態。

5. BIAS（乖離率）

BIAS來量化當前價格與移動平均線的偏離程度，用於判斷價格是否過度偏離趨勢。計算公式如下：

$$\text{BIAS}(n) = [(\text{當根K收盤價} - \text{近n根K移動平均線}) / \text{近n根K移動平均線}] \times 100\%$$

通常 $n = 6、12$ 或 24 。

BIAS > 正值閾值（如 +5%）表示價格過高，可能回落；< 負值閾值（如 -5%）表示價格過低，可能反彈。極端 BIAS 值（如 $\pm 10\%$ ）可能預示反轉。

6. CCI (Commodity Channel Index, 商品通道指數)

CCI亦用來衡量價格偏離其平均值的程度，用於識別超買或超賣區域。計算公式如下：

(1) 首先求出TP（典型價格）：

$$\text{TP} = (\text{當K最高價} + \text{當K最低價} + \text{當K收盤價}) / 3$$

(2) 再計算TP的簡單移動平均、MD（平均偏差）及CCI：

$$\text{SMA}(\text{TP}, n) = \text{近n根K典型價格的簡單移動平均}$$

$$\text{MD} = \sum |\text{TP} - \text{SMA}(\text{TP}, n)| / n$$

$$\text{CCI} = (\text{TP} - \text{SMA}(\text{TP}, n)) / (0.015 \times \text{MD})$$

通常 $n = 14$ 。

當CCI > 100 表示超買，可能回落，考慮賣出；CCI < -100 表示超賣，可

能反彈，考慮買入。CCI 穿越零線（正轉負或負轉正）可作為趨勢轉換訊號。與價格背離時，可能預示反轉。

7. KD (Stochastic Oscillator, 隨機指標)

KD指標用於衡量價格在一定周期內的相對位置，判斷超買或超賣狀態，並通過快線（%K）與慢線（%D）的交叉提供買賣訊號。計算公式如下：

(1) 首先求出RSV(未成熟隨機值)：

$$RSV = [(\text{當根K收盤價} - \text{近n根K最低價}) / (\text{近n根K最高價} - \text{近n根K最低價})] \times 100$$

通常 $n = 14$ 。

(2) 再計算%K及%D：

$$\%K = \text{RSV 的 } i \text{ 簡單移動平均值}$$

$$\%D = \%K \text{ 的 } j \text{ 簡單移動平均值}$$

通常 $i = 3$ 、 $j = 3$ 。

當%K上穿%D且兩者低於 20，表示買入訊號；而當%K下穿%D且兩者高於80，表示賣出訊號。%K或%D > 80表示超買，可能下跌；< 20 表示超賣，可能上漲。KD 與價格背離（如價格創新高或低但KD未同步新高或低）可能預示反轉。

8. MACD(Moving Average Convergence/ Divergence, 指數平滑異同移動平均線指標)

MACD指標用於判斷趨勢方向與動量，通過快線與慢線的交叉提供買賣訊號。

計算公式如下：

(1) 首先求出MACD：

$$\text{MACD} = \text{EMA}(\text{收盤價}, n1) - \text{EMA}(\text{收盤價}, n2)$$

$$\text{EMA}(n) = n \text{ 根K的指數移動平均線，通常 } n1 = 12、n2 = 26。$$

(2) 再計算DIFF(訊號線)及柱狀體：

$$\text{DIFF} = \text{EMA}(\text{MACD}, n3)$$

$$\text{柱狀體} = \text{MACD} - \text{DIFF}$$

$$\text{通常 } n3 = 9。$$

當MACD 線上穿訊號線為買入訊號，下穿為賣出訊號。柱狀圖由負轉正表示看漲，由正轉負表示看跌。

9. DMI (Directional Movement Index，動向指數)

DMI是一個技術分析系統，用於評估價格趨勢的方向與強度，包含兩個主要組成部分：PDI（正向動向指數）和 MDI（負向動向指數）。DMI通過比較多空力量的相對強弱，幫助交易者判斷市場趨勢的方向。計算公式如下：

(3) 首先求出+DM（正向動向）、-DM（負向動向）及TR（真實波幅）：

$$+DM = \text{當K最高價} - \text{前K最高價} \quad (\text{若小於0或}-DM, \text{則為 } 0)。$$

$$-DM = \text{前K最低價} - \text{當K最低價} \quad (\text{若小於0或}+DM, \text{則為 } 0)。$$

$$TR = \text{Max}(\text{當K最高價} - \text{當K最低價}, |\text{當K最高價} - \text{前K收盤價}|, |\text{當K最低價} - \text{前K收盤價}|)。$$

(4) 再計算PDI（正向動向指數）及MDI（負向動向指數）：

$$PDI = (+DM \text{ 的 } n \text{ 根K平滑值} / TR \text{ 的 } n \text{ 根K平滑值}) \times 100$$

$$MDI = (-DM \text{ 的 } n \text{ 根K平滑值} / TR \text{ 的 } n \text{ 根K平滑值}) \times 100$$

通常 $n = 14$ 。

當PDI上穿MDI表示上升趨勢開始，可考慮買入；MDI上穿PDI表示下降趨勢開始，可考慮賣出或放空。PDI和MDI的差距擴大表示趨勢強度增加；差距縮小表示趨勢減弱。

第三節 當沖交易

一、台指期貨

期貨（Futures）是一種金融合約，買賣雙方約定在未來的某一指定時點，以事先確定的數量和價格對特定標的商品進行買進或賣出的交易方式。與即期交易的現貨商品（Spot Market）相比，期貨的主要區別在於其履約時點為未來既定的時間。此外，期貨交易在期貨交易所的規範下進行，期貨合約為標準化的交易契約，明確載明買賣雙方的權利與義務。期貨交易還借助保證金交易制度，促使交易的確定完成，這使得期貨市場能夠有效運作。

台灣期貨交易所（Taiwan Futures Exchange，TAIFEX）於1996年12月經政府批准設立，並於1997年3月通過《期貨交易法》立法，隨後在1998年7月21日正式開業。台灣期貨交易所自成立以來推出了多項期貨商品，其中首項期貨商品便是「台股期

貨」，該商品自推出後便成為台灣期貨市場的核心交易工具。本研究選擇台灣股價指數期貨（TAIEX Futures）作為實證對象，其期貨合約規格如表2-1所示。

二、當沖

「當沖」即「當日沖銷」（Day Trading）的簡稱，是指在同一交易日內，針對同一投資標的進行買進和賣出的操作，從而實現平倉並結算交易的行為。這種交易方式的特點是投資人在當日內完成所有交易，避免持倉過夜，因此能夠有效減少由於市場消息、新聞報導或其他外部因素所帶來的風險，然而當沖交易因為在較短的時間內作出決策，若交易次數過於頻繁，也有因交易成本過高而侵蝕利潤的風險。

三、交易策略

為避免非預期風險及過度交易的問題，許多研究者加入濾網或停損停利機制對試圖改善策略績效。停損的主要目的是於進場後控制虧損，避免在錯誤判斷市場趨勢時承擔過大的損失。當價格達到預設的停損價位時，交易者即刻出場，減少虧損的擴大。相對地，停利的目的是提高勝率，當持有的部位獲利時，通過提前平倉來確保收益，雖然可能因此錯過更多的潛在利潤，但也能避免獲利回吐成虧損。濾網（filter）則是一種過濾條件或過濾機制，旨在避免因為市場噪音或非預期的風險造成的過度交易或錯誤交易，用來在進場前篩選出更具價值的交易機會，過濾掉可能帶來虧損或不確定性的交易。吳百正（2004）發現，儘管KD指標、DMI指標及MACD指標在交易成本考慮下無顯著獲利，但加入停損機制後，策略的操作績效顯著提高。高明駿（2013）指出，透過調整K棒時間週期的濾網設計，並結合不同的進出場時間，部分當沖策略的績效可顯著改善。柯麗美（2014）發現停損機制能有效減少虧損，而停利機制對整體績效影響較小，並且以5分鐘K棒最適用於當沖交易。許文華（2014）發現CDP指標與Pivot Point指標可於台灣指數期貨當沖交易獲得正報酬，並且濾網設計有助於減

少虧損。莊文澤（2018）則以不同的K棒時間週期行當沖策略的回測，結果亦發現5分鐘K棒最適用於應對突發行情，能有效避免每筆交易的虧損過高。這些研究表明，濾網設計和停損機制對提高當沖策略的績效有顯著影響。。

歸納以上，透過濾網、停損停利機制甚至K線的時間刻度，都能有效避免過度交易引發的利潤侵蝕，是提升交易程式績效穩定性的關鍵。本文將沿用過去文獻研究成果，採用5分鐘K線為資料，並在交易策略中納入停損停利及濾網機制。

第四節 異常偵測

一、機器學習在交易策略的應用

近年來，機器學習技術也被廣泛應用於期貨當沖策略。林典南（2008）使用AdaBoost演算法進行台股期貨的當沖研究，結果顯示，增設交易濾網及定期重新訓練模型能增加獲利的可能性。賴怡玲（2009）運用增強式學習進行台指期的當沖策略研究，並強調最大化獎勵函數的學習方式，達到高達93%的策略報酬率。施孝承（2011）則利用決策樹模型進行資料探勘，有效辨識當日價格漲跌對後市發展的影響。楊迪敦（2015）指出，加入倒傳遞類神經網路(Back-Propagation Networks, BPN)的濾網後，當沖策略的整體績效可提高20%。吳政晏（2019）研究開盤價突破策略並結合LSM模型，發現這些策略在使用最適參數組合後能有效改善績效。鍾嘉元（2024）將交易策略搭配TimesNet模型預測異常交易點，提供更加穩健的當沖方式，這些研究展示了以機器學習技術作為濾網對於提升交易策略的獲利績效具有相當潛力。

二、異常偵測演算法

異常偵測（Anomaly Detection）是機器學習演算法的一種應用，透過資料分析技術來識別資料集中與大多數資料不同或不符合預期模式的異常值。這些異常點可能

代表錯誤資料、異常行為，甚至是潛在的風險或攻擊，因此異常偵測在資料挖掘、金融監控、網路安全、醫療診斷、工業設備監測等領域具有重要應用價值，常見的異常偵測模式包括：

1. 監督式異常偵測：

監督式異常偵測 (Supervised Anomaly Detection) 依賴標記過的資料集，所有資料被分別標記為正常資料或異常資料，並使用這些標籤化資料訓練模型。在訓練過程中，模型學習正常資料的特徵與分佈，並根據這些學習到的模式來判斷新資料是否為異常。監督式異常偵測的主要優勢在於其基於標籤資料進行學習，因此通常能提供較高的準確性。然而此方法需要大量的標籤資料，而現實中異常資料往往十分稀少，蒐集足夠的異常樣本往往困難，此外，如果異常資料的特徵與正常資料過於相似，則模型的判別能力可能受限，影響偵測效果。

2. 半監督式異常偵測：

半監督式異常偵測 (Semi-Supervised Anomaly Detection) 在訓練時僅需部分資料具備標籤 (通常是正常資料)，而大部分資料則未被標記。模型先使用標記資料學習正常與異常的差異，接著利用未標記資料進行訓練，試圖將未知資料分類為正常或異常。此方法的優勢在於無需大量標記資料即可訓練模型，並能有效利用未標記資料，提高訓練效率，然而準確性通常低於監督式異常偵測，因為未標記資料的數量與分佈可能影響模型學習的效果。

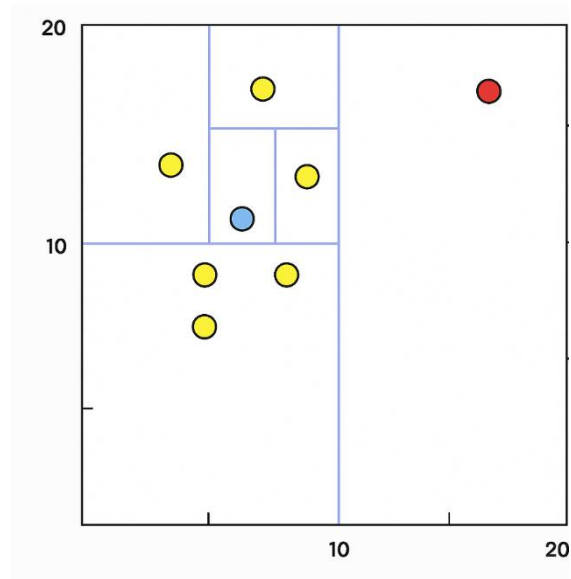
3. 非監督式異常偵測：

非監督式異常偵測 (Unsupervised Anomaly Detection) 模型不依賴任何標記過的正常或異常數據，而是透過聚類、密度估計、距離計算等方法，將數據分為多個子群，並從中識別異常數據點。由於無需標註數據，此方法特別適用於無法獲取足夠標籤數據的應用場景，如網路入侵偵測、金融詐欺偵測和工業設備故障診斷等。此外，非監督式方法能夠發掘新的、未知的異常模式，而不僅限於識別過去已知的異常類型，因而在動態環境中具有更強的適應性。

三、孤立森林演算法

孤立森林演算法 Isolation Forest (IForest) 演算法由 Fei Tony Liu、Zhi-Hua Zhou 等人提出，屬於非監督式異常偵測演算法，採用了非參數化(Non-parametric)的設計不依賴於資料的具體分佈形態，以隨機選擇的特徵及切分值為基礎，構建一系列二元樹(稱為 iTree)來達到辨認異常點的目的。IForest 假設異常資料點在特徵空間中會呈現孤立狀態，即與其他資料點的距離較遠，或在特徵值上表現出顯著差異。這一假設與傳統的基於密度或距離的異常檢測方法有所不同，因為這些方法依賴於資料點之間的相似性或密度來進行識別。IForest 主要通過資料分割的方式來識別異常，其程序是以一個隨機超平面對資料空間進行切割，每次切割將資料空間一分為二，形成兩個子空間，接著再對每個子空間再次施以隨機超平面切割，不斷重複進行直到每個子空間中僅剩下一個資料點為止(圖2-1)。正常值通常屬於高密度區域，需要經過多次切割才會被分離到單獨劃分的狀態，故其與根節點之間的距離較長；而異常值通

常屬於高密度區域，會較早在切割過程中被單獨劃分，因此其與根節點之間的距離相對較短(圖2-2)。而因為單棵iTree的結果有一定的局限性，為了提高檢測精度，IForest通過集成多棵iTree的結果，計算每個資料點的平均分割距離，進而得出其異常分數，異常分數的值介於0至1之間，越接近1表示為異常點的可能性越高。



常分數，異常分數的值介於0至1之間，越接近1表示為異常點的可能性越高。

圖 2-1 隨機超平面切割

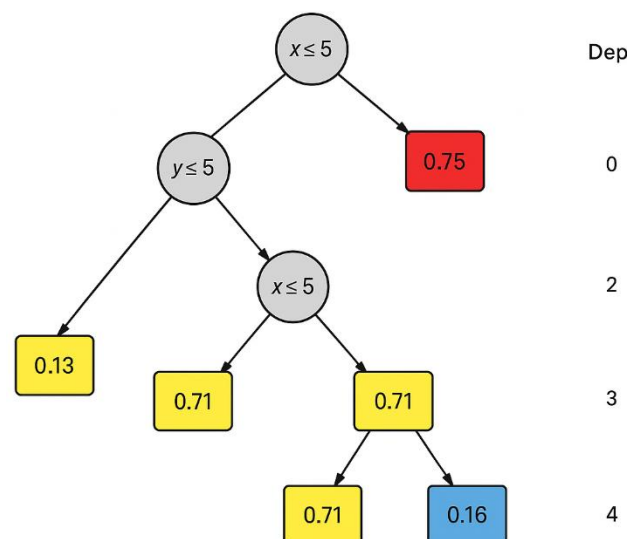


圖 2-2 資料與根節點之間的距離

第三章 研究方法

第一節 研究架構

本研究之整體架構如圖3-1所示。首先進行K線資料的蒐集與前處理，並計算包括KD、RSI、MACD及CDP等多項技術指標，作為後續策略建構的依據。接著，以CDP指標搭配不同的停損與停利條件，設計多種交易策略，並透過長期回測篩選出能夠穩定獲利的最佳策略。再以該最佳策略於訓練集期間的每一次進場時點為基準，擷取進場前固定時間步長內的各項技術指標作為特徵，輸入孤立森林（Isolation Forest， IForest）模型進行訓練，藉以模擬策略判斷進場的邏輯，並以驗證集資料調整模型參數，以獲得最佳模型。最後，透過回測比較僅使用傳統交易策略、僅使用機器學習模型進行交易、以及機器學習模型輔助的策略交易三種交易方式，藉以探討傳統量化交易與機器學習策略之間的差異，並探討模型輔助策略交易對績效的提升效果。

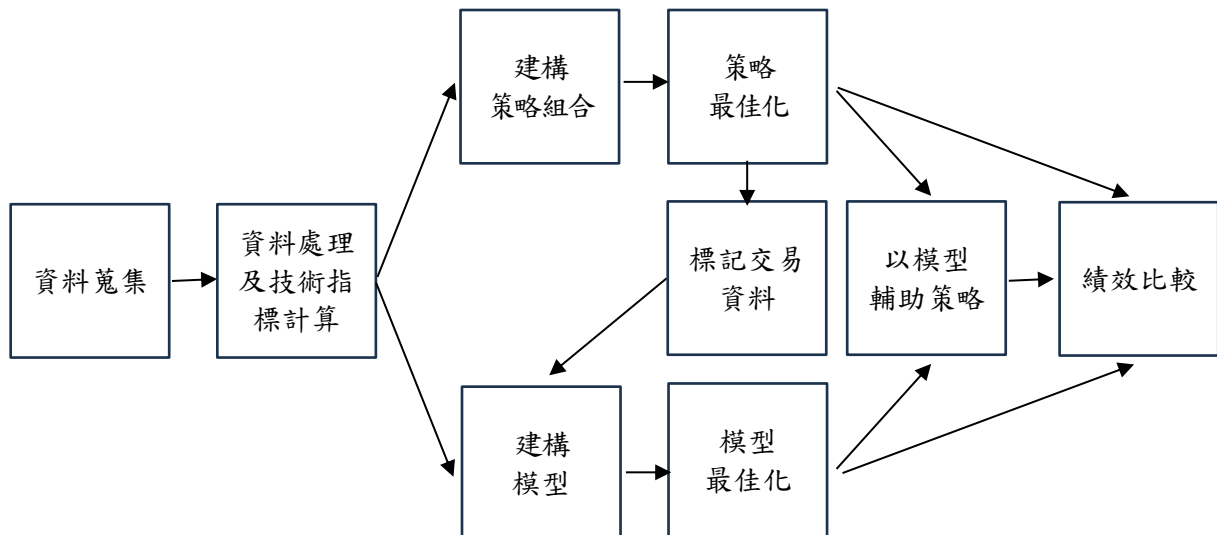


圖 3-1 研究架構圖

第二節 資料處理及資料集劃分

一、資料處理

本研究以台灣期貨指數市場為研究對象，運用台灣期貨交易所台灣指數期貨自2000年1月2日至2024年12月31日止的一般盤(即日盤，早上8:45至下午13:45)及盤後交易(即夜盤，下午15:00至早上05:00)的逐筆歷史交易資料轉換為約67萬筆5分鐘K線的開、高、低、收盤價格以及成交量資料，並計算逆勢操作系統(CDP)、能量潮(OBV)、VMA、RSI、MACD、超買超賣指標(OBOS)、乖離率(BIAS)、隨機指標(KD)、指數平均移動平均線(MACD)、相對強弱指標(RSI)、威廉指標(W%R)、移動平均線(MA)、動量指標(MTM)、趨向指標(DMI)、商品通道指標(CCI)等技術指標作為進、出場策略及機器學習模型特徵。

二、資料集劃分

本研究將資料分為訓練、驗證及測試3個部分，以2001年1月2日至2020年3月10日資料為訓練集(占約60%)，進行交易策略設計並回朔測試，確立最佳交易策略及機器模型訓練，接著再以2020年3月11日至2022年7月26日資料為驗證集(占約20%)，檢驗交易策略績效並決定最佳模型參數，最後以2022年7月27日至2024年12月31日資料為測試集(占約20%)，確認最佳參數的績效穩定性(表3-2)。

表 3-1 資料集劃分

資料集名稱	涵蓋區間	交易日數(日)	K棒數量	占比(%)
訓練集	2001.01.02-2020.03.10	4,739	398,848	59.8
驗證集	2020.03.11-2022.07.26	583	132,834	19.9
測試集	2022.07.27-2024.12.31	592	135,017	20.3
全部資料	2001.01.02-2024.12.31	6,234	666,699	100.0

註：期交所2017年5月15日起推動夜盤交易。

第三節 交易成本設定

期貨交易成本主要來自期貨交易所徵收的交易稅、期貨公司收取的交易手續費以及市場滑價所帶來的額外成本。這些費用會隨著市場條件及政策的變動而有所調整：

一、交易稅

現行台股期貨的交易稅為十萬分之二，交易稅的計算公式為：

交易稅 = 每口契約價值 × 交易稅率，四捨五入至整數位

每口契約價值 = 期貨成交價格(或最後結算價) × \$200

以指數在20,000點為例，一次買或賣交易稅=20000 × 200 × 0.0002，約為80元，一對買賣交易稅約為160元。

二、交易手續費

期貨手續費依各家不同證券商或不同營業員也會有不同的收費方式，各家券商和期貨商的手續費大約在新台幣50至100元之間不等。在實際交易中可能因價格在小範圍波動而導致短時間內買進賣出的訊號不斷被錯誤觸發，進而導致交易成本過高，從而削弱交易策略報酬。

三、交易滑價

滑價（Slippage）是指在進行期貨或其他金融商品交易時，實際成交價格與預期成交價格之間的差異。這種差異可能發生在任何市場中，尤其是在市場波動性較大或流動性較低的情況下。滑價通常會在執行交易時出現，尤其是對於快速進出市場的交易者或大宗交易。在此設定交易一次可能損失滑價2點，即400元。

綜合上述，本研究估計交易稅、交易手續費及滑價成本為新台幣560元，以整數600元為每一對買賣交易總成本進行策略的績效評估。

第四節 交易策略設計

一、最大持有部位及交易時間設定

臺灣期貨交易所規定的交易時間為早上08:45至下午1:45，下午15:00至隔日早上5:00，本研究設計不同進、出場條件組合為多種交易策略，每個策略皆逐筆判斷自2000年1月2日至2024年12月31日全部交易時段的5分鐘K資料，模擬進、出場行為並計算獲利績效。為避免同時持有多個倉位造成互相影響，不易於評估交易決策的獲利能力，所以限制同時最大持有部位為1口台指期。

二、交易流程

本研究回測策略交易期間為2000年1月1日至2024年12月31日，含括訓練、驗證及測試集資料，自每日8:45及15:00起判斷是否符合進場條件，進場後從進場時點起逐筆資料判斷是否符合出場條件，若符合停損條件則優先執行停損出場，若未符合停損條件且符合停利條件，則進行停利出場。出場後自出場時點起重新判斷是否符合進場條件，若符合條件可以再進行進場，並繼續判斷出場條件，如此反覆進行進場出場操作，直到強制平倉時間，若仍持有部位則以收盤價強制平倉。

三、進場條件

本研究以CDP指標為主要進場條件，用前1交易日的最高價、最低價、及收盤價透過公式計算出當日價格的AH、NH、CDP、NL、AL等5個支撐與壓力價位，作為判斷

進場的依據，具體交易策略如下。

1. 當價格觸及NL時，買進一口多單。
2. 當價格觸及NH時，買進一口空單。
3. 當價格觸及AH時，買進一口多單。
4. 當價格觸及AL時，買進一口空單。

四、出場停損條件

1. 點數停損

(1) 若進場為多單，當價格觸及買進價格以下固定點數時，以賣出價格以下固定點價格停損。

(2) 若進場為空單，當價格觸及賣出價格以上固定點數時，以買進價格以上固定點價格停損。

2. CDP指標壓力支撐停損

(1) 若在NL買進一口多單，當價格觸及AL時，以AL價格停損。

(2) 若在NH買進一口空單，當價格觸及AH時，以AH價格停損。

五、出場停利條件

1. 點數停利

(1) 若進場為多單，當價格觸及買進價格以上固定點數時，以賣出價格以上固定點價格停利。

(2) 若進場為空單，當價格觸及賣出價格以下固定點數時，以買進價格以下固定點價格停利。

2. 移動平均線停利

- (1) 因為在當根K收盤前，移動平均的價位仍會變動，所以以前一K的移動平均價為觸發出場的價格。
- (2) 若進場為多單，當價格自移動平均線以上向下觸及前一根K移動平均價時，賣出前一根K移動平均價格。
- (3) 若進場為空單，當價格自移動平均線以下向上觸及前一根K移動平均價時，買入前一根K移動平均價格。

3. 百分比停利

- (1) 若進場為多單，當價格上漲超過固定點數後又向下回跌固定百分比時，以固定百分比價格賣出。
- (2) 若進場為空單，當價格下跌超過固定點數後又向上回漲固定百分比時，以固定百分比價格買入。

4. 分階段百分比停利

- (1) 若進場為多單，當價格上漲超過固定點數後又向下回跌固定百分比時，以固定百分比價格賣出，並依據上漲幅度調整向下回跌時觸發停利的的百分比。
- (2) 若進場為空單，當價格下跌超過固定點數後又向上回漲固定百分比時，以固定百分比價格買入，並依據下跌幅度調整向上回漲時觸發停利的的百分比。

六、強制平倉條件

本研究為當沖操作，設定完全避免留倉風險，所以時間為下午1:45或隔日早上5:00時，若仍持有部位則以下午1:45或隔日早上5:00的收盤價強制平倉出場。

七、策略組合

本研究以進場條件與出場條件搭配為多種策略組合，並逐一依據交易流程執行後比較獲利績效以選擇最佳策略，所有策略組合如表3-3：

表 3-3 交易策略組合

進場條件	停損條件	停利條件	交易策略組合
CDP指標	點數停損	點數停利	策略A1
		移動平均線停利	策略A2
		百分比停利	策略A3
		分階段百分比停利	策略A4
	CDP指標壓力支撐停損	點數停利	策略B1
		移動平均線停利	策略B2
		百分比停利	策略B3
		分階段百分比停利	策略B4

第五節 交易模型設計

本研究在建構一個有長期正報酬的交易策略後，藉由異常偵測進行交易訊號篩選，試圖進一步推升交易策略的報酬。為評估個績效推升的程度，將比較最佳策略與最佳策略經異常

偵測模型優化後的回測績效，異常偵測模型建立方法如下：

一、輸入特徵選取

在特徵變數選取方面，本文參考台股期當沖策略研究，選擇能量潮(OBV)、VMA、RSI、MACD、超買超賣指標(OBOS)、乖離率(BIAS)、隨機指標(KD)、指數平均移動平均線(MACD)、相對強弱指標(RSI)、威廉指標(W%R)、移動平均線(MA)、動量指標(MTM)、趨向指標(DMI)、商品通道指標(CCI)等常用的技術指標作為特徵變數，並以最佳策略進場時點的前2根5分K的各技術指標作為模型輸入特徵。

二、模型參數設定

為探究不同模型參數對交易策略優化的潛力，本研究以2種主要參數建立多個模型，探究參數對模型效能的影響，並用以挑選出最佳模型：

1. 森林中樹的數量：

森林中樹的數量($n_{estimators}$)增加可提升模型穩定性，但回報逐漸遞減，且會增加運算成本。一般設定為100即可取得穩定結果。

2. 認定異常的比例：

認定異常的比例(contamination)介於0到0.5，是孤立森林模型的一個超參數，表示數據集中預期的異常樣本比例（範圍通常為 0 到 0.5）。它在訓練階段用來指導模型假設多少百分比的數據點是異常的。

3. 門檻值：

門檻值(threshold)是指將孤立森林輸出的異常分數轉換為二元分類（異常或正常）的決策邊界。孤立森林會為每個樣本生成一個異常分數（通常範圍在0到1，值越高越異常），而門檻值決定何時將樣本標記為異常

三、模型建立流程

本文以最佳策略的進場交易紀錄分為多方與空方分別進行模型訓練，為挑選較佳模型，將森林中樹的數量設定為較少、標準、較多3個層級，各層級樹的數量分別為100、256、512個，而認定異常的比例(contamination)則自0至0.5分為100等分，以網格搜尋法(Grid Search)將最佳策略在訓練集期間(2001年1月2日至2020年3月10日)進場交易紀錄的輸入特徵投入不同參數的模型進行訓練，訓練完成後再投入驗證期間(2020年3月11日至2022年7月26日)進場交易紀錄的特徵變數進行預測，並比較對績效表現的改善程度，最後再以測試集(2022年7月27日至2024年12月31日)確認模型的穩定性。

四、模型評估方式

雖然孤立森林本質上為非監督式機器學習方法，而本研究目的為藉由機器學習方法提高交易獲利能力，淘汰較低獲利或負獲利的交易，所以設定將扣除成本後獲利小於1點（相當於進場價格為20,000點時扣除成本獲利0.005%）的交易標籤設為1，其餘設為0，與孤立森林判斷的異常與正常進行比較。本研究中「異常樣本」被視

為正類，正常樣本則歸類為負類。透過模型預測驗證集的結果與真實標記進行交叉比對，建立混淆矩陣（Confusion Matrix），並計算正確率（Accuracy）、精確率（Precision）、召回率（Recall）等分類效能指標，分別用來衡量模型判定為異常之樣本中，有多少比例為真正的異常、所有實際異常樣本中，模型成功偵測的比例。另外從獲利能力提升的角度考量，以驗證集獲利能力提升能力進行評估，包含總獲利、每次交易平均獲利、交易次數進行評估。

第六節 交易績效衡量

在完成基礎交易策略與機器學習模型之建構後，將分別以及結合應用於測試集期間（自2022年7月27日至2024年12月31日）之資料進行回測，以檢驗其於未見資料上的獲利表現。因本研究旨在建構一套具備長期正報酬潛力之量化交易策略，為客觀評估各策略於測試期間之績效，評估面向涵蓋扣除交易成本後之獲利能力、績效穩定性以及風險程度，將採用量化交易領域常用之績效進行比較與分析：

一、總淨利潤：

總淨利潤(Total Net Profit)是指在回測期間內，交易策略扣除所有成本（包括交易費用、滑價等）後的總利潤，直接反映策略的獲利能力。

二、年報酬率

年報酬率是指策略在回測期間內的平均年化收益率，以百分比表示。透過將總報酬標準化為年度收益，用來衡量策略長期獲利潛力。

三、平均每一單損益

平均每一單損益是指策略在回測期間內，每筆交易的平均盈虧金額，計算方式為總淨利潤除以總下單次數，反映單筆交易的平均獲利能力。

四、獲利因子

獲利因子(Profit Factor)：總盈利與總虧損的比率，計算方式為所有獲利交易總額除以所有虧損交易總額。若數值大於1，表示策略具備盈利能力；數值越高，代表獲利相對於虧損的優勢越顯著。

五、勝率

勝率(Percent of Win)是指獲利交易次數占總交易次數的比例，以百分比表示。反映策略在交易中收益為正的機率，勝率越高代表策略執行結果的穩定性較佳。

六、總下單次數

總交易次數(Total Number of Trades)是指在回測期間內執行的總交易次數，反映策略的容易觸發的程度與交易頻率，高下單次數可能提升數據的統計可靠性，但也可能因過度交易增加成本。

七、最大回撤

最大回撤(Maximum Drawdown)是指交易期間內，從歷史最高點（累積獲利的峰值）下跌到最低點時所經歷的最大損失幅度，通常以百分比或絕對值表示，反映在最糟糕情況下可能承擔的最大資金回落，是評估交易策略或投資組合風險的重要指標之一。

八、最大連續虧損天數

最大連續虧損天數(Recovery Time)是指資產淨值從進入虧損狀態起到重新回到原始投資淨值（轉虧為盈）所需的最長天數，用來衡量投資者在策略執行過程中，可能需要忍受的最長連續虧損時間，是評估策略耐受性和投資者心理承受能力的重要參考。

九、夏普比率

夏普比率(Sharpe Ratio)是用來衡量策略相對於無風險資產的獲利表現，由於本研究主要目的在比較不同策略獲利能力，無風險利率不影響策略間比較基準，此處將無風險資產的年報酬率假設為零。夏普比率越高，代表在承擔相同風險的情況下，策略能提供更高的收益。計算方式如下，其中 R_p 為策略年報酬率， R_f 為無風險資產的年報酬率， σ_p 為策略年投資報酬率波動（標準差）：

$$\text{夏普比率} = (R_p - R_f) / \sigma_p$$

十、索提諾比率

索提諾比率(Sortino Ratio)同樣是衡量「承擔單位風險能得到多少報酬」的指標，差別在於在計算風險時，只考慮虧損的波動大小，因為Sortino的理論認為正報酬的波動程度(上行波動率)是交易的加分項，因此不應包括在風險計算中，所以將正報酬的波動程度從公式中剔除，計算波動程度時只計入負報酬的波動程度 σ_d （下行波動率）：

$$\text{索提諾比率} = (R_p - R_f) / \sigma_d$$

十一、卡瑪比率

卡瑪比率是用來評估策略總收益相對於最大回撤風險的指標，計算方式為年化報酬率除以最大策略虧損的絕對值，反映策略在承受最大損失情況下的報酬效率，數值越高代表策略在風險管理與收益間的平衡越優。