

國立虎尾科技大學
資訊管理系
專題製作報告

應用深度學習於金融市場模型

中華民國 113 年 5 月

國立虎尾科技大學
資訊管理系
專題製作報告

< 應用深度學習於金融市場模型 >

成果發表：
評審委員：_____

指導老師：楊達立老師

學 生： 41041209 張婷穎
41041213 蔡昀真
41041223 黃柏浚
41041226 賴益廷

中華民國 113 年 5 月

摘要

現今有許多用於預測金融市場的深度學習模型，常用的有長短期記憶模型（Long Short-Term Memory, LSTM）[1]、循環神經網路（Recurrent Neural Network, RNN）、門控循環單元（Gated Recurrent Unit, GRU）等，根據公開的研究及論文中，我們發現深度學習模型在預測價格上，並無法應用於即時的金融市場當中。此外，在交易的過程中，人們容易產生錯失恐懼症（Fear of missing out, FOMO），導致執行錯誤的決策，無法像機器一樣的正确執行策略，造成不可彌補的虧損，而金融市場學習的門檻很高，需要花費大量的時間成本，因此本研究提出一種方法，以交易為核心，價格預測為輔，利用深度強化學習中的Q-Learning與深度Q網路（Deep Q Network, DQN）作為我們的交易代理人，進行短線投機交易，本研究從台灣股票市場、美國股票市場及加密貨幣市場中，選取了這些市場中市值較大的股票和加密貨幣作為研究對象，以每小時價格資料作為我們的資料集，結合時間序列模型預測之價格，將其作為新的技術指標，當作特徵傳入DQN模型中進行訓練與學習，使DQN模型學會短線投機行為，以獲得更高的報酬率為目標，最後將模型應用於量化交易上，實現自動化投機交易系統。

關鍵字：深度學習、深度強化學習、短線交易、投機、策略交易

誌謝

本研究項目達成的每一步，都凝聚了團隊夥伴的共同奮鬥與努力。特別地，我們要向我們的指導導師—楊達立教授，表達深深的謝意。在從選擇研究主題、進行細緻的案例分析到整個論文撰寫的過程中，楊教授不斷地提供專業的見解與激勵，使我們能夠順暢地完成了這項研究。

目錄

頁數

摘要	I
誌謝	II
目錄	III
圖目錄	V
表目錄	VI
第一章 研究動機與研究問題	- 1 -
第二章 文獻回顧與探討	- 2 -
2.1 深度學習	- 2 -
2.2 深度強化學習	- 3 -
2.3 金融市場中的技術指標	- 4 -
2.4 迴歸策略	- 4 -
2.5 趨勢策略	- 5 -
2.6 短線交易結果評估	- 6 -
第三章 研究方法及步驟	- 8 -
3.1 取得資料	- 8 -

3.2 資料處理.....	- 9 -
3.3 訓練價格預測模型.....	- 10 -
3.4 定義交易決策模型.....	- 13 -
3.5 訓練交易決策模型.....	- 15 -
第四章 研究結果與討論.....	- 16 -
4.1 模型在各市場學習趨勢策略的結果	- 16 -
4.2 模型在各市場學習迴歸策略的結果	- 20 -
4.3 不同模型交易能力比較結果.....	- 22 -
第五章 結論.....	- 24 -
第六章 參考文獻.....	- 25 -
附錄	- 27 -
作者簡介.....	- 35 -

圖目錄

	頁數
圖 2.1 LSTM 長短期記憶架構圖.....	2 -
圖 2.2 深度強化學習工作原理.....	3 -
圖 4.1 2382 廣達趨勢策略收益回撤圖	17 -
圖 4.2 MSFT 微軟趨勢策略收益回撤圖.....	18 -
圖 4.3 BTC 比特幣趨勢策略收益回撤圖	19 -
圖 4.4 2382 廣達迴歸策略收益回撤圖	20 -
圖 4.5 MSFT 微軟迴歸策略收益回撤圖.....	21 -
圖 4.6 BTC 比特幣迴歸策略收益回撤圖	22 -
圖 4.7 不同模型交易能力比較圖	23 -

表目錄

	頁數
表 2.1 SQN 評分標準	- 7 -
表 3.1 時間序列模型輸入格式的參數表	- 10 -
表 3.2 模型參數表	- 11 -
表 3.3 模型比較實驗結果	- 12 -
表 4.1 指標代號對應表	- 16 -
表 4.2 2382 廣達趨勢策略統計表	- 17 -
表 4.3 MSFT 微軟趨勢策略統計表	- 18 -
表 4.4 BTC 比特幣趨勢策略統計表	- 19 -
表 4.5 2382 廣達迴歸策略統計表	- 20 -
表 4.6 MSFT 微軟迴歸策略統計表	- 21 -
表 4.7 BTC 比特幣迴歸策略統計表	- 22 -

第一章 研究動機與研究問題

近年來，以深度學習或機器學習預測金融市場一直是金融界及計算機領域中非常熱門的話題，並且有許多的研究資料。由於延遲性等問題，實際上無法單純透過深度學習來預測金融市場的未來，因此也有了許多結合與改良的相關研究，例如：混合注意力機制（Hybrid Attention Networks, HAN）與長短期記憶（LSTM）、加速梯度長短期記憶（Accelerated Gradient Long-Short Term Memory, AG-LSTM）[2]、結合卷積神經網路與長短期記憶[3]等，此外也有結合技術指標的相關研究，例如：利用情緒指標及長短期記憶模型預測台灣加權股價指數[4]，而研究中也有提到結合與改良深度學習的方式，確實可以提高預測的準確率，而達到更好的報酬率，從中可以發現相關研究都是以時間序列模型預測價格為主，深度學習結合交易策略卻鮮少人探討，事實上在行為經濟學上已經有許多的研究表明，不管是投機客或是投資客，在交易過程中都存在著情感，終究無法避免人性的考驗，容易產生錯失恐懼症（FOMO），FOMO的產生容易影響交易決策，導致交易過程中犯錯，無法像機器一樣準確的執行交易[5]，因此本研究提出以策略交易為核心，價格預測為輔的量化交易模型，除了預測市場短期走勢外，透過深度強化學習中的深度Q網路（DQN）結合強化學習中的Q-Learning算法，並以最大投機報酬率為目標，進行策略交易模型的訓練，使模型作為我們的交易代理人，代替我們執行交易決策。

第二章 文獻回顧與探討

2.1 深度學習

深度學習是人工智慧中的一種方法，以神經網路（Neural Network, NN）為基礎框架延伸的一種機器學習結構，現今已有許多延伸框架，例如：卷積神經網路（Convolutional Neural Network, CNN）、循環神經網路（Recurrent Neural Network, RNN）等，由於本計畫為時間序列問題，因此採用循環神經網路作為我們的深度學習架構，而循環神經網路會隨著長時間的訓練中，其產生的誤差於反向傳遞時會導致梯度消失或梯度爆炸的問題。為了解決上述問題，於1997年由Hochreiter和Schmidhuber提出了長短期記憶模型[6]，該模型基於循環神經網路的架構進行改良，與循環神經網路相比，加入了閘門的機制，分別是遺忘閘門、輸入閘門及輸出閘門，並利用訊號控制閘門的開關，使長短期記憶能夠處理長時間的問題，發展至今已經是循環神經網路中最常用的框架之一，其架構圖如圖2.1所示。

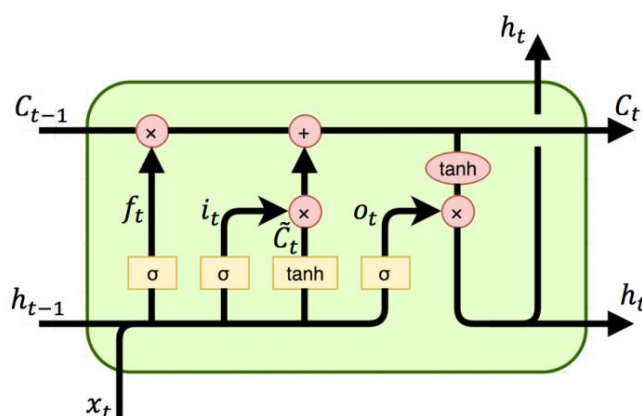


圖 2.1 LSTM 長短期記憶架構圖（來源：<http://dprogrammer.org/rnn-lstm-gru>）

根據Zexin Hu、Yiqi Zhao和Matloob Khushi所做的調查[7]，使用深度學習進行外匯和股票的價格預測，回報率最高是由J. WU等人所提出的基於深度強化學習的股市量化交易[8]，該篇研究中於前期選擇模型時，根據不同的模型框架進行價格預測的測試，發現長短期記憶模型的預測效果是最好的，於是本研究採用長短期記憶相關的架構作為預

測價格的模型。

2.2 深度強化學習

深度強化學習（Deep Reinforcement Learning, DRL）是一種結合了深度學習（Deep Learning, DL）和強化學習（Reinforcement Learning, RL）的技術，用於解決具有高度複雜性的決策過程問題。其中強化學習是在探討如何在嘗試錯誤的過程中讓智慧型代理人（Agent）學習做出更好的決策，代理人通過與環境（Environment）交互來學習如何達到某個目標。在這一過程中，代理人根據其所採取的行動獲得獎勵（Reward）或懲罰（Punishment），並利用這些獎勵來指導其未來的行動，以獲得最大化的累積獎勵，其工作原理如圖2.2所示。

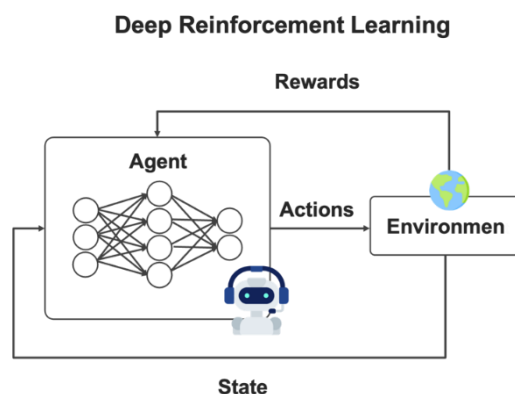


圖 2.2 深度強化學習工作原理

本研究在深度強化學習中使用深度Q網路方法，以下簡稱為DQN模型，並結合強化學習中的Q-Learning算法，進行決策交易模型的訓練。DQN模型是由DeepMind的研究團隊開發，在2015年的研究中成功地將深度學習與強化學習結合起來，實現了在多個Atari 2600遊戲上達到了人類專家水平的成就[9]。這項研究開創了深度強化學習領域的一個重要里程碑，為後續的研究奠定了基礎。DQN模型引入了經驗回放機制，使學習過程中多次利用之前的經驗，降低資料之間的相關性及提高學習效率，在不斷變化的金融市場中，可以通過持續學習以適應市場的變化。此外，DQN模型可以通過與環境的交互學習交易策略，使其能夠在市場上做出理性的決策，使用DQN模型既節省了金融市場的

學習成本，又可以減少人為的錯誤及風險，提高交易的效率，並且在學習過程中，避免重大的損失，以最大化獎勵為目標，選擇最適合的決策來進行交易。

2.3 金融市場中的技術指標

技術指標是分析股票、外匯、商品或其他可交易資產價格行為的數學計算工具。通常交易者和分析師使用這些指標來預測市場趨勢、確定入場和退出點，以及識別潛在的交易機會。技術指標可以分為幾個主要類型，包括趨勢指標、動量指標、波動性指標和成交量指標。在本研究中，使用常見的九種技術指標，用於當作時間序列模型的特徵，使模型在預測價格時，根據技術指標在不同價格的特徵進行學習，從而使預測價格更為貼近真實價格。

2.4 迴歸策略

迴歸策略在金融領域是一種基於統計方法的交易策略，策略的基本假設是價格或收益率會圍繞其歷史平均值波動，即所謂的均值迴歸。這種策略假設在經過一段時間的極端變動後，價格或收益會返回到其歷史之平均水平。在多種的交易工具和方法中，布林通道（Bollinger Bands）因其廣泛的認可度和流行性，已經成為實施迴歸策略中不可或缺的一個重要工具。

布林通道，由John Bollinger在1980年代提出[10]，是一種圍繞價格移動平均線及其波動性的技術分析工具。該指標包含三條線：中央線通常是20日的簡單移動平均線（SMA），而上軌和下軌則分別表示價格的上下波動範圍，通常設置為移動平均線兩側的兩個標準差。布林通道的核心在於評估資產價格的相對高低以及波動性，提供了一個視覺化的方式來觀察價格是否達到異常波動的區域，從而判斷可能的迴歸機會。

本研究引入了布林通道的概念，利用時間序列模型預測最高價、

最低價及收盤價，將最高價作為布林通道之上軌，最低價作為布林通道之下軌，而收盤價則做為均值，當預測的收盤價高於預測的最高價時，這表示市場價格可能過熱，存在迴歸的可能性，即預示著價格可能會出現下跌，反之，當預測的收盤價低於預測的最低價時，即預示著價格可能會出現上漲。在訓練DQN模型時，我們將該策略加入到模型當中，通過外部影響DQN模型的決策，使DQN模型學會應用此策略並進行改進，實現短線投機的目的。

2.5 趨勢策略

趨勢策略是一種基於識別並跟隨市場趨勢來進行交易的方法，是金融市場中最常見和廣泛應用的交易策略之一，該策略的核心假設是一旦市場形成了某個方向的移動趨勢，該趨勢有可能會在一段時間內持續。因此，通過識別這些趨勢並相應地調整交易決策，交易者可以利用市場的動向來獲取利潤。趨勢可以分為上升趨勢、下降趨勢和橫向趨勢，交易者通常使用技術指標，如：移動平均線（Moving Average, MA）、相對強弱指數（Relative Strength Index, RSI）、指數平滑異同移動平均線（Moving Average Convergence Divergence, MACD）等來識別市場趨勢，在眾多的趨勢策略中，移動平均線策略無疑是最為經典且廣泛應用於市場上的一種。該策略的核心在於使用移動平均線來平滑價格資料，從而清晰地揭示出市場的趨勢方向。透過觀察不同周期的移動平均線之間的交叉情況，例如黃金交叉（Bullish Cross）和死亡交叉（Bearish Cross），交易者可以捕捉到市場趨勢變化的早期信號。本研究在DQN模型的訓練中，加入移動平均線策略的概念，以真實價格及預測價格的交叉，從而判斷出市場趨勢的方向，當真實價格由下往上穿過預測價格時，則表示趨勢從下降趨勢轉變為上升趨勢，因此在做交易決策時，應該選擇買進，反之，當真實價格由上往下穿過預測價格時，則該選擇賣出。

2.6 短線交易結果評估

為了全面評估短線交易的效能與風險，本研究採用了幾個關鍵的性能指標：最大回撤（Maximum Drawdown, MDD）、系統品質數（System Quality Number, SQN）及管理資金比率（MAR Ratio）。這些指標能夠從不同角度提供策略的表現評估，幫助交易者理解各策略的風險與回報。

1. 最大回撤（MDD）：

最大回撤是衡量投資組合潛在損失的指標，它表示從一個高點到下一個低點的最大跌幅，通常用來評估策略在熊市期間的表現。公式如下：

$$MDD = \frac{Peak - Trough}{Peak}$$

其中，Peak 是策略達到的最高價值，Trough 是從這個高點後達到的最低價值。

2. 系統品質數（SQN）：

SQN 是由交易系統設計專家 Van K. Tharp 博士提出[11]，用以評估交易系統性能的指標。SQN 計算公式為：

$$SQN = \sqrt{N} \times \frac{Mean(R)}{SD(R)}$$

其中，N 是交易次數，R 代表每次交易的收益率，Mean(R) 是收益率的平均值，SD(R) 是收益率的標準差。SQN 較高表示系統性能穩定且可靠，如果 $SQN > 2$ ，則表示交易策略相當不錯。針對 SQN 的值，我們給定了範圍作為交易系統的評分標準，其表格如表 2.1 所示。

表 2.1 SQN 評分標準

SQN	交易系統能力
<1.00	差
1.01~2.00	普通
2.01~3.00	好
3.01~5.00	優秀
5.01~7.00	專業
>7.00	完美

3. 管理資金比率（MAR Ratio）

MAR 比率是比較投資回報與最大回撤的指標，常用於評估交易策略的風險調整回報。計算公式為：

$$MAR = \frac{TP}{MDD}$$

其中，TP（Total Profit）表示總收益，MDD 是最大回撤。MAR 比率較高意味著相對於其承受的風險，策略獲得了較高的回報。

通過這些指標，我們可以對交易結果進行全面的量化分析，不僅衡量短線交易的盈利能力，還能評估其在面對市場波動時的風險承受能力。這種多角度的評估方法為短線交易的選擇提供了科學依據，有助於我們衡量策略與模型的好壞。

第三章 研究方法及步驟

3.1 取得資料

首先，我們必須取得市場的價格走勢，來作為我們訓練預測價格模型的資料集，為此，我們針對不同的金融市場，使用不同的 API 工具來收集資料。在台灣股票市場方面，我們使用永豐金證券 API[12] 來取得每分鐘 K 線資料；美國股票市場方面，則是選擇使用 yfinance 套件來取得 Yahoo Finance 提供的美股資料；加密貨幣市場則是使用當前全球最大的加密貨幣交易所幣安[13]，作為資料的來源，透過幣安提供的第三方 Python 套件，來取得加密貨幣的價格走勢。本研究所採用的資料集涵蓋了從 2020 年 3 月 2 日至 2022 年 3 月 31 日的時期，這個時間範圍包括牛市和熊市的行情，以便全面評估所開發模型在不同市場狀態下的表現和穩健性。在此期間，全球金融市場經歷了顯著的波動，其中包括由 COVID-19 大流行引發的市場動盪和隨後的經濟復甦，這使得該時期的資料特別適合用於測試金融模型對於極端市場條件的適應能力。而在時間週期方面，考慮到本研究的重點在於短線交易，我們選擇每小時作為資料的時間間隔。這種選擇反映了短線交易者對市場動態快速反應的需求，並使模型捕捉到這些短期內的價格波動，從而更精確地模擬和預測市場行為，並做出交易決策。每小時資料提供了足夠的細節，使得模型能夠識別每小時的價格趨勢和潛在的交易機會，這對於深度強化學習模型做出短線交易的決策至關重要。

此外，在標的的選擇中，本研究選擇了市值較大的股票和加密貨幣作為分析對象，在台股方面，我們選擇了市值排名前面的股票，包括 2330 台積電、2317 鴻海、2412 中華電及 2382 廣達。在美股方面，由於資料集的搜集較為困難，因此選擇的標的有所限制，標的包括金融領域的 JP Morgan Chase (JPM)、科技領域的 Microsoft (MSFT) 及影視串流服務的 Netflix (NFLX)。至於加密貨幣市場，我們選擇了比特幣 (BTC)、以太幣 (ETH) 及幣安幣 (BNB)。這些標的的選擇有

助於分析市場的趨勢，並在多變的市場條件中尋找投機機會，其中的共通點在於有著較高流動性和顯著的價格波動，市值大的股票和加密貨幣具有較高的交易量，這意味著這些資產的買賣更加容易，能夠迅速吸引和集中大量資金，從而為投資者提供豐富的交易機會。此外，這些市場的價格波動性較大，這種快速的價格變動提供了利用短期價格差異獲利的機會，特別是對於那些尋求快速獲利的短線操作者來說，顯著的價格波動不僅增加了交易的風險，同時也提高了潛在的獲利空間，這樣的市場為後續的短線交易提供了理想的環境。

3.2 資料處理

在資料處理的過程中，我們取得到的資料集格式都不一致，為了解決格式問題，本研究制定了一套統一資料格式的標準，我們將來自不同市場的資料集標準化，統一資料欄位為日期（datetime）、開盤價（open）、最高價（high）、最低價（low）、收盤價（close）、成交量（volume）。

接下來，為了確保模型在訓練過程中的學習效率，我們需要將資料進行正規化（Normalization），把所有資料特徵調整至相同的範圍，這裡我們把資料縮放到 0-1 的區間範圍內。資料縮放是資料前處理的一個很重要的關鍵，確保模型在訓練過程中不會因為特徵間的範圍差異過大而影響學習效率。此外，本研究設定預測目標為收盤價、最高價、和最低價。這些預測目標將直接影響到後續深度強化學習模型的訓練，因為這些目標是策略執行的關鍵，為了讓深度強化學習模型學會策略的應用，是進行短線交易策略的進出場依據。隨後，將資料集分割成訓練集和測試集，分配比例為 60% 和 40%。這樣的分配確保了模型有充足的資料進行學習，同時留出足夠的測試資料來驗證模型的預測能力。在資料處理的最終階段，我們將資料轉換成時間序列模型的輸入格式，詳細的參數設定如表 3.1 所示。

表 3.1 時間序列模型輸入格式的參數表

參數 市場	批次大小 (Batch Size)	序列長度 (Sequence Length)	特徵數 (Feature Size)
台股	32	25	26
美股	32	40	26
加密貨幣	32	120	26

在我們的實驗中發現，當批次大小 (Batch Size) 過大時，雖然可以顯著減少訓練時間，但模型的預測性能卻有所下降。經過反覆測試，我們確定最佳的Batch Size為32，這一設定能夠在不犧牲預測性能下，最大程度地提高價格預測模型的效能，序列長度 (Sequence Length) 我們設定為五天的價格走勢資料。考慮到不同市場的交易時間差異，對Sequence Length的數值做了不同的設定，台灣股票市場的每日交易時間為4.5小時，美國股票市場為7.5小時，而加密貨幣市場則為24小時不間斷交易，我們將各市場的交易時間無條件進位，並將這個數值乘以五天。這樣的計算方法決定了表3.1中所示的序列長度。至於特徵數，包括了開盤價、最高價、最低價、收盤價、成交量及技術指標等關鍵特徵，使模型能夠更容易捕捉到關鍵。

3.3 訓練價格預測模型

在時間序列模型的選擇方面，本研究比較了多種模型，包括長短期記憶模型 (LSTM)、門控循環單元模型 (GRU)、雙向長短期記憶模型 (Bi-LSTM)、雙向門控循環單元模型 (Bi-GRU)、卷積長短期記憶模型 (CNN-LSTM)、卷積門控循環單元模型 (CNN-GRU)，並從中找出在相同參數及相同資料集的情況下的最佳模型來當作我們的價格預測模型，詳細參數如表3.2所示。

表 3.2 模型參數表

模型 超參數	LSTM	GRU	Bi- LSTM	Bi- GRU	CNN- LSTM	CNN- GRU
隱藏層單元數量 (hidden_size)	128	128	128	128	128	128
全連接層單元數量 (fc_size)	128	128	128	128	128	128
模型層數 (num_layers)	2	2	2	2	2	2
隨機丟棄率 (dropout_prob)	0.1	0.1	0.1	0.1	0.1	0.1
輸入通道數 (input_channels)	x	x	x	x	1	1
卷積核大小 (kernel_size)	x	x	x	x	3	3
卷積輸出大小 (cnn_out_size)	x	x	x	x	32	32
池化核大小 (pool_kernel_size)	x	x	x	x	2	2
LSTM/GRU 輸出大小 (lstm/gru_out_size)	x	x	x	x	64	64
卷積次數 (num_convolutions)	x	x	x	x	3	3

實驗結果經過十次測試並取平均值後表明，Bi-GRU在所有比較模型中表現最佳，損失值最小，並且決定係數最接近1。這表明Bi-GRU在預測變數的變異性方面具有卓越的表現，模型的預測值與實際觀察值之間的差異非常小，如表3.3所示。因此，在價格預測價格上面，我們選擇使用Bi-GRU來作為價格預測模型。Bi-GRU在相同資料集與相同參數下，於不同的時間序列模型中有著較低的損失和更高的預測準確度，在預測結果上面，較貼近真實價格。經過我們的實測，價格預測模型的精確度將會影響深度強化學習模型的決策，當準確度越高時，交易決策的結果越好，因此Bi-GRU被選為本研究的價格預測模型。

表 3.3 模型比較實驗結果（上方數字為 MSE，下方數字為 R^2 ）

模型 標的	Bi-GRU	LSTM	GRU	Bi-LSTM	CNN- LSTM	CNN- GRU
2317 鴻海	0.069 0.995	0.07 0.994	0.071 0.994	0.071 0.994	4.587 0.654	4.105 0.691
NFLX.US	47.279 0.995	363.364 0.962	84.531 0.991	268.649 0.972	565.237 0.941	576.586 0.94
ETHUSDT	1794.108 0.997	3470.844 0.994	2519.37 0.996	3198.584 0.994	70487.108 0.877	45183.762 0.921

在進行模型訓練時，我們特別注意避免過度擬合（Overfitting）和低度擬合（Underfitting）的問題，過度擬合通常發生於模型過於徹底地學習訓練資料中的每一細節，包括那些不具代表性的噪聲和特徵，而這些特徵並不適用於未見過的資料中，這會導致模型在訓練資料上表現出色，但在新的或驗證資料上表現不佳。相反地，低度擬合發生在模型未能在訓練資料上獲得足夠的學習，通常是因為模型過於簡單，無法捕捉資料中的關鍵變化，導致模型在訓練集和測試集上都表現不佳。為了應對這些問題，我們在訓練中加入了提前停止法（Early Stopping），該方法涉及在訓練過程中監控驗證集上的損失，若在連續數個訓練週期內驗證損失沒有顯著改善，則自動停止訓練。這不僅節省了計算資源，也有助於維持模型對新資料的泛化能力。此外，由於 Bi-GRU 模型在訓練中已經展現出與真實價格非常接近的預測結果，微調超參數對預測性能的提升作用有限，因此我們選擇保持原有的參數設定，具體如表 3.2 所示。

由於不同的股票或加密貨幣在價格、交易量及技術指標上具有顯著差異，這些差異反映了各自市場的特定特性和交易者行為。因此，本研究針對每種商品，單獨訓練對應的模型，以確保我們的模型能夠更精確地理解和預測各個市場的獨特動態。

3.4 定義交易決策模型

在本研究中，交易決策模型指的是深度強化學習模型，我們選擇深度強化學習中的 DQN 模型，結合強化學習中的 Q-Learning 算法，進行決策交易模型的訓練，以有效地學習並執行交易策略。DQN 模型在模擬的交易環境中進行訓練，搭配我們設定的交易策略，從而學習如何在實際市場中做出利潤最大化的交易決策。

首先，我們必須建立一個交易環境，使 DQN 模型能夠在環境內進行學習，透過與環境的交互，從中學習交易決策，該環境就像是真實的交易系統，有著完整的交易功能，包括交易命令的發出與執行、交易方向的選擇、資金管理和收益率計算等，這樣的設計可以使 DQN 模型能夠在一個接近實際市場的條件下進行學習交易策略。此外，環境還需要提供反饋及獎懲給模型，使模型對於決策進行修正，學習如何使獎勵最大化。

在初始化交易環境時，我們需要先進行環境參數的設定，參數包括了輸入的資料、資金設定、時間窗口大小、動作的種類及交易策略。在輸入資料方面，我們整合了來自預測模型的輸出資料及實際的市場價格資料作為環境的輸入，將預測模型的輸出資料當作新的技術指標，搭配真實市場價格，進行策略的設計。資金的設定則是包含了初始資金和每筆交易所能動用的最大資金比例，這是控制風險的一種策略，確保模型在保持收益最大化的同時，也能避免過度風險。時間窗口大小（Time Window Size）的設定則是基於資料的時間範圍選定，這決定了模型評估市場狀況所依賴的資料窗口大小。而動作的種類，指的是模型能進行的操作有哪些，在本研究中，可操作的動作分別是開單買進（long）、關單賣出（close_long）及不做任何操作（hold）。最後是交易策略的制定，透過我們設計的交易策略，指導模型在適當時機執行買入或賣出操作。

在交易策略的設計上，我們只進行做多的交易，不考慮做空的情況。我們提出了基於迴歸策略和趨勢策略的兩種交易方法。首先，我們的迴歸策略利用預測價格中的最高價及最低價作為價格的上限及下限，並以收盤價作為價格的平均水平，當真實價格低於下限時，進行買入的操作，當價格回到平均水平時，進行賣出的操作，此策略的核心基於一個假設，當價格經歷劇烈波動後，往往會迴歸到它的平均水準，利用價格的暫時偏離所出現的交易機會，從中進行短線操作，期望通過市場的自我調節而獲得收益。而我們的趨勢策略則是利用預測價格中的最高價、最低價及收盤價，作為趨勢判斷的依據，當真實價格超過最高價時，代表上漲的力量非常強大，則表示當前為上漲的趨勢，應進行買入的操作，當真實價格回到預測價格的收盤價，即代表趨勢可能準備轉換了，因此進行賣出的操作，當真實價格低於最低價時，則表示當前為下跌趨勢，不應該進行任何操作，此策略的原理是，一旦市場形成了某個方向的移動趨勢，這一趨勢有可能會在一段時間內持續，因此交易的方向應與當前市場的趨勢保持一致。

在環境反饋的設計上，我們將每天的價格漲跌幅作為環境的狀態，並反饋給模型，使模型能夠根據市場的即時動態做出反應，並針對這些變化進行學習，從而使模型選擇更為精確的交易決策。在獎懲方面，我們設計了一套獎懲機制，當模型有訂單並決定關閉訂單時，若下一個小時的價格繼續上漲，則視為模型未能把握住持續獲利的機會，因此給予等同於漲跌幅的懲罰；反之，如果下一個小時的價格未能繼續上漲，模型則能夠獲得漲跌幅的獎勵加上交易所獲得的實際收益。此外，若模型選擇不進行任何交易操作，則會受到輕微的懲罰，以激勵模型探索並利用可利用的市場機會。這樣的獎懲機制，既鼓勵模型學習如何最大化收益，同時減少不必要的交易風險，使模型可以更有效地學習何時進行交易及何時保持觀望，從而提高整體的交易效率和報酬率。

3.5 訓練交易決策模型

在本研究中，我們使用了深度 Q 網絡（DQN）來訓練交易決策模型，模型的參數設定在訓練過程中起著關鍵作用，因為它們直接影響到模型的學習效率和最終性能。DQN 模型與價格預測模型一樣，均採用了提前停止法（Early Stopping），這是為了避免過度擬合，並在模型學習足夠的特徵後即時停止訓練，以維持模型對新資料的泛化能力。此外，我們將批次大小（Batch Size）設定為 2048，這是經過一系列試驗後確定的最佳值。在增加批次大小的過程中，訓練速度有著顯著得提升，我們發現與價格預測模型不同的是，所得到的報酬結果並沒有下降，但我們也觀察到，當批次大小達到一定程度後，模型的報酬率會達到峰值並有所下降。根據我們的測試，設定批次大小為 2048 能夠在加速訓練的同時，維持模型的高效表現和報酬率。在環境參數方面，我們將初始資金設定為 100,000 單位。為了簡化分析過程，本研究並未進行不同市場貨幣單位之間的換算，而是專注於探討模型對於市場變化的預測能力和對交易策略的適應性。這樣的設定允許我們更直觀地評估模型在各種市場條件下的表現，而不受貨幣價值差異的影響。每筆交易的資金比例為 0.5，即每筆交易使用的資金不得超過總資金的 50%，透過限制每筆交易可使用的資金比例，我們能夠有效控制風險，確保模型在面對不利市場條件時，仍有足夠的資金進行其他潛在的盈利交易。時間窗口大小則是與價格預測模型的序列長度相同，都是以五天為基準，交易策略方面，則是將上述兩種交易策略都單獨進行了訓練，並且與價格預測模型一樣，對於每個交易商品，分別訓練專屬的模型，以便更準確地把握並預測每個市場的特定變化。

第四章 研究結果與討論

本章節將詳細展示並分析本研究所得到的結果。我們使用了Bi-GRU結合DQN的模型，在不同策略、不同交易市場下的表現進行全面評估。為了準確評估模型的交易能力及風險管理能力，我們使用了多個指標來衡量模型的交易能力，每個指標都有對應的代號，針對每個代號所代表的意義，我們將其整理成表4.1。

表 4.1 指標代號對應表

指標名稱	代號
總收益	TP
總交易次數	TT
總報酬率	RR
勝率	WR
盈虧比	O
最大損失	ML
最大回撤	MDD
管理資金比例	MAR
系統品質數	SQN

此外，本章節還包括了對本研究的模型與其他研究者所提出的模型進行比較。透過比較交易的結果，我們不僅能夠評估本研究模型的能力，還可以確定模型的優勢所在。

4.1 模型在各市場學習趨勢策略的結果

在台灣股票市場中，我們選擇2382廣達作為代表，其統計結果如表4.2所示，收益回撤圖如圖4.1所示。從兩年的回測結果上來看，經過DQN模型的學習，此策略能夠有效地從市場上獲利，總報酬率達到12.39%。在各項指標中，可以看到勝率方面的表現特別突出，高達72.22%，反映了其選擇交易時機的精確性。在風險管理上，相對於總資金，最大損失控制在了1.6%，7.38的MAR指標進一步證實了在風險

管理方面的卓越表現。最後的SQN指標為2.86，這表明該交易系統不僅盈利性好，而且具有良好的穩定性。然而，該模型在交易機會的次數上仍有提升空間，在兩年的實驗期間內，交易次數相對較少，這可能限制了其總體報酬率的進一步提升。

表 4.2 2382 廣達趨勢策略統計表

TP	TT	RR	WR	O	ML	MDD	MAR	SQN
12,394	36	12.39%	72.22%	1.35	-1,679	-1,679	7.38	2.86

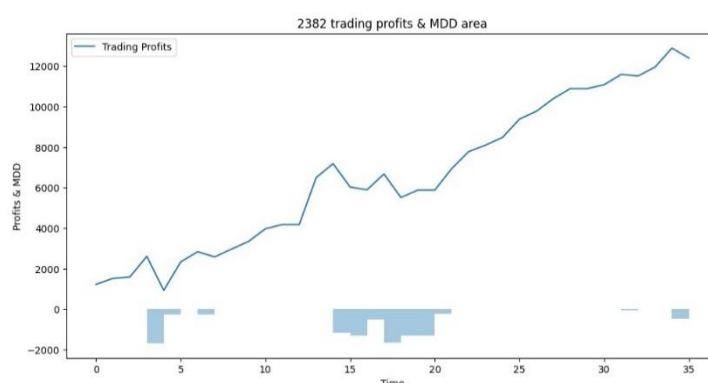


圖 4.1 2382 廣達趨勢策略收益回撤圖

在美國股票市場中，我們選擇MSFT微軟作為代表，其統計結果如表4.3所示，收益回撤圖如圖4.2所示。從回測的結果來看，同樣的交易策略在美國市場同樣展現出了顯著的盈利潛力。在兩年的實驗期間內，MSFT微軟實現了18.24%的總報酬率，進一步確證了模型掌握了獲利的能力。從各項指標來分析，MSFT微軟的勝率達到了68.63%，雖略低於2382廣達的表現，但依然表明了模型在選擇進出市場時機的高度精確性。相較於2383廣達的結果，MSFT微軟的交易次數在兩年的實驗期間有51次的交易，雖然仍不算高，但相對於2382廣達的情況有所改善。

表 4.3 MSFT 微軟趨勢策略統計表

TP	TT	RR	WR	O	ML	MDD	MAR	SQN
18,236	51	18.24%	68.63%	1.36	-1,771	-3,758	4.85	2.56

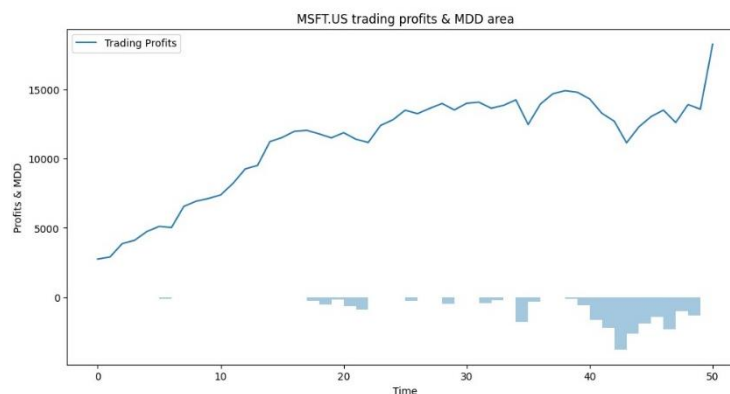


圖 4.2 MSFT 微軟趨勢策略收益回撤圖

緊接著，我們進一步探討了加密貨幣市場，以比特幣作為加密貨幣市場的代表，根據兩年的回測結果，模型在比特幣上顯示出了極高的盈利潛力和卓越的市場表現，總報酬率達到了32.10%，這一結果明顯高於先前的傳統股票市場，顯示出加密貨幣市場的投機性和高回報潛力。從各項指標來看，模型在比特幣上的勝率高達79.17%，這不僅反映了模型選擇進出時機的極高精確性，也突顯了加密貨幣市場波動性大的特點下，模型依舊能保持著精確的進出場時機。此外，管理資金比例高達11.67，這一結果不僅反映了在面對高風險的情況下仍能實現高額回報，也顯示了模型在極端市場條件下依然能保持良好的資金管理，並突顯了加密貨幣市場相對於傳統股票市場的高報酬潛力。系統品質數（SQN）為3.71，則進一步證明了該模型在比特幣市場中的穩定性和盈利性。

表 4.4 BTC 比特幣趨勢策略統計表

TP	TT	RR	WR	O	ML	MDD	MAR	SQN
32,095	48	32.10%	79.17%	1.25	-998	-2,751	11.67	3.71

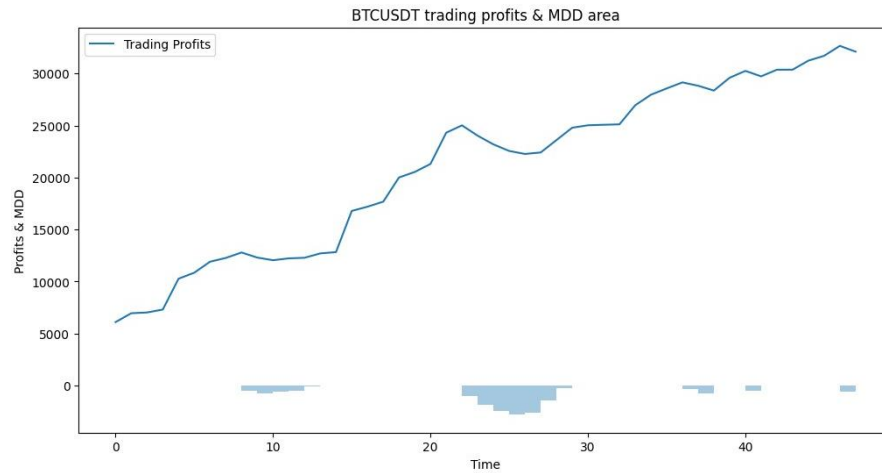


圖 4.3 BTC 比特幣趨勢策略收益回撤圖

透過上述呈現的結果，我們觀察到模型在不同市場之間表現的差異，我們的模型不僅在高報酬率的實現上表現卓越，也在波動性較高的市場環境中保持了高勝率和優秀的資金管理能力。另外，針對本研究涉及的每一個商品的詳細結果，我們已將之整理於附錄中供讀者進一步查閱。

4.2 模型在各市場學習迴歸策略的結果

在台灣股票市場的研究中，我們選定 2382 廣達作為案例，以評估迴歸策略的效能。具體統計結果已在表 4.5 中詳細列出，並在圖 4.4 中提供了收益與回撤圖。根據結果表示，DQN 模型學習了迴歸策略後，能夠於市場上獲利，實現了 31.56%的總報酬率。在勝率方面，與趨勢策略的結果無太大差異，依然維持著高勝率的水準，在風險管理上，可以看到該模型的最大損失控制在-2,214，相比於趨勢策略，風險稍微增加，但獲利能力卻是大幅上升，其中 MAR 指標為 9.98，說明了風險管理依然有著卓越的表現，而 SQN 指標為 3.26，表示該交易系統有著良好的獲利性及穩定性，並且交易能力遠大於使用趨勢策略的模型。

表 4.5 2382 廣達迴歸策略統計表

TP	TT	RR	WR	O	ML	MDD	MAR	SQN
31,557	62	31.56%	69.35%	1.48	-2,214	-3,162	9.98	3.26

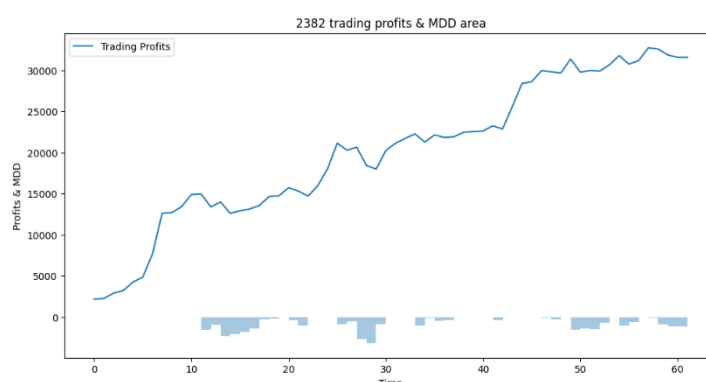


圖 4.4 2382 廣達迴歸策略收益回撤圖

在美國股票市場的研究中，與趨勢策略的研究對象相同，選擇 MSFT 微軟作為案例，以評估迴歸策略的效能。具體統計結果已在表 4.6 中詳細列出，並在圖 4.5 中提供了收益與回撤圖。從結果來看，同樣的交易策略在美國市場上有著更好的績效，在兩年的回測期間中，模型在 MSFT 微軟實現了 52.12%的報酬率，並且依然維持著高勝率的交易能力，勝率高達 70.11%。在高勝率的情況下，風險管理也保持著高水平，MAR 指標為 17.54。從 SQN 指標的結果來看，交易系統已達到了優秀的水準，達到了 4.87。

表 4.6 MSFT 微軟迴歸策略統計表

TP	TT	RR	WR	O	ML	MDD	MAR	SQN
52,118	87	52.12%	70.11%	1.90	-2,970	-2,970	17.54	4.87

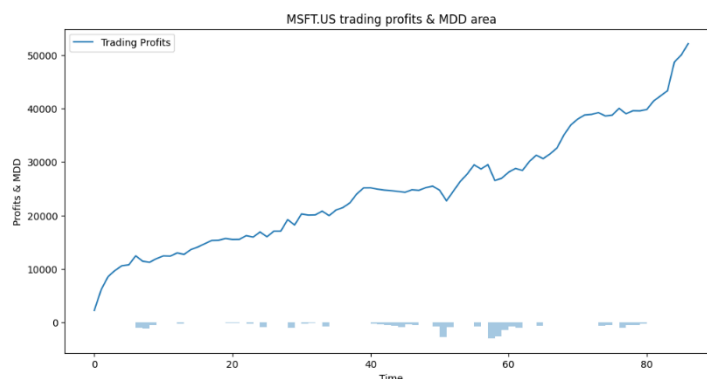


圖 4.5 MSFT 微軟迴歸策略收益回撤圖

在加密貨幣領域上，在趨勢策略的結果上，我們可以得知加密貨幣市場相對於傳統股票市場的高報酬潛力，但高報酬的情況下，與此同時價格的高波動增加了巨大的風險，因此在加密貨幣上，資金管理能力是非常重要的。經過模型重新學習了迴歸策略後，我們發現模型在比特幣上有著更顯著的表現，其統計結果如表 4.7 所示，收益與回撤圖如圖 4.6 所示。其統計結果表示在總報酬率上獲得了 430.31%，使得本金翻了 4 倍，這一結果相較於趨勢策略有著更巨大的成長，顯示出了模型的潛力。此外，在勝率方面，模型依然保持在 60% 以上的水準，而代表風險管理能力的 MAR 指標，也可以看到有著極高的數值，達到了 23.25，最後的 SQN 指標，則是表示該交易系統是專業等級。在這些回測結果中，我們可以發現不管是趨勢策略還是迴歸策略，我們的模型在短線交易上都有獲利的能力，並且還有風險管理的能力，能夠利用短進短出來實現收益，其中，學習迴歸策略的模型，相較於趨勢策略，有著更卓越的交易能力。

表 4.7 BTC 比特幣迴歸策略統計表

TP	TT	RR	WR	O	ML	MDD	MAR	SQN
430,315	362	430.31%	65.47%	1.73	-16,585	-18,508	23.25	6.89

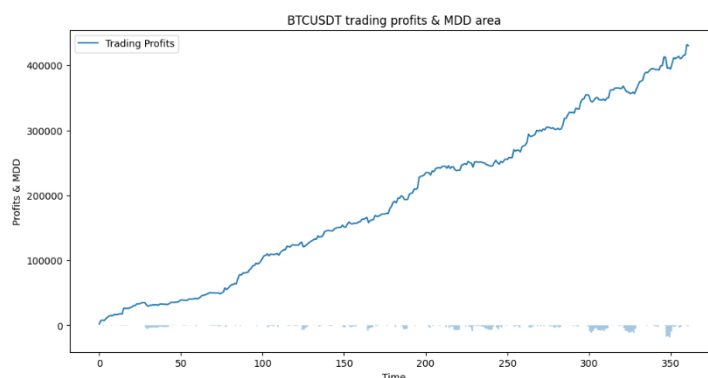
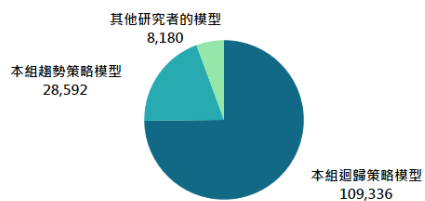


圖 4.6 BTC 比特幣迴歸策略收益回撤圖

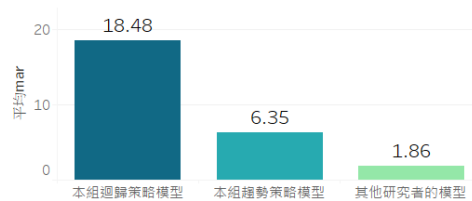
4.3 不同模型交易能力比較結果

除了上述兩個模型外，我們還引入了由量化交易研究員wbbhcb開發的深度強化學習模型[14]。在相同的輸入和參數設定下，我們對比了不同模型之間的交易能力，相關的結果展示於圖4.7中。從結果圖可以明顯看出，在獲利能力及風險管理方面，我們的模型顯示出了明顯的優勢。特別是在市場極端波動的情況下，我們的模型在確定進出市場的關鍵時刻表現相對精準。基於這些發現，我們相信我們的模型具有作為交易代理人的潛力，有望實現本研究的目標，開發一個全自動化的短線交易系統。這樣的系統將有助於提高交易的效率與效果，並減少人為操作錯誤，保證策略的一致性和持續性執行。

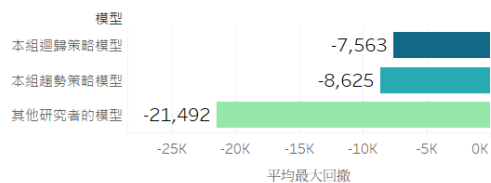
平均收益比較



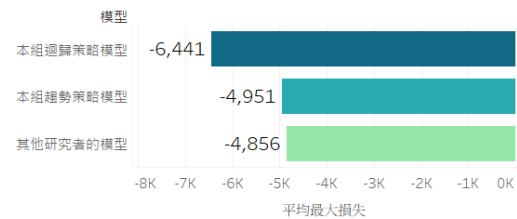
平均MAR



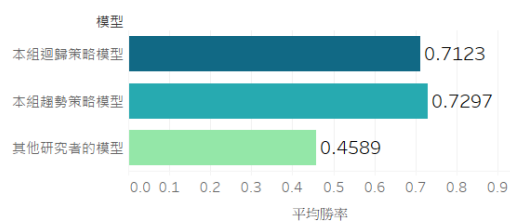
平均最大回撤



平均最大損失



平均最大損失



平均總報酬率

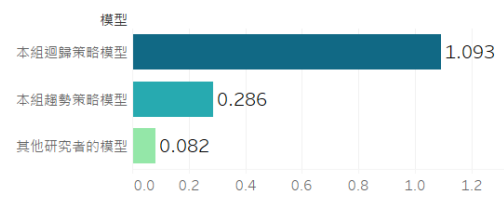


圖 4.7 不同模型交易能力比較圖

第五章 結論

本研究採用了深度學習技術及深度強化學習技術，先使用時間序列模型，進行價格的預測，模型訓練過程中，搭配九種技術指標作為特徵，使預測價格更為貼近真實的價格，提高了價格預測的準確性，進一步幫助了後續的交易決策模型在實際操作中的可靠性和穩定性，交易決策模型中，我們利用深度Q網路（DQN）模型，結合我們設計的策略，來進行短線投機交易的決策。在進行短線交易方面，我們所訓練的模型不僅提高了交易的勝率，也顯著改善了最大回撤（MDD）與系統品質數（SQN）等重要的性能指標。透過交易決策模型的學習能力，模型能夠在大量的歷史資料中學習和提取有效的交易訊號與特徵，並且透過我們所制定的趨勢策略及迴歸策略，進行策略的學習，這證明了深度強化學習在市場交易的潛力。研究結果顯示，這些模型在獲利能力和風險管理方面，相比於其他研究者的交易模型，有顯著的提升，驗證了我們的模型在金融市場上的應用價值。此外，我們實驗的對象包含台股、美股及加密貨幣市場，展示了我們使用相同的模型下，對不同市場動態的適應性及穩定性。透過觀察研究結果，可以發現在使用趨勢策略及迴歸策略的模型當中，同樣的商品及同樣的時間區段，使用迴歸策略的模型所展示的結果都優於使用趨勢策略的模型，迴歸策略模型展現出了良好的獲利性及穩定性。

第六章 參考文獻

- [1] Liu, H.-W., & Wang, J.-H. (2020) . Combining hybrid attention networks and LSTM for stock trend prediction. In *Proceedings of the 32nd Conference on Computational Linguistics and Speech Processing (ROCLING 2020)* (pp. 304–318) . Taipei, Taiwan: The Association for Computational Linguistics and Chinese Language Processing (ACLCLP) .
- [2] Deepika, & Bhat, N. M. (2021) . An efficient stock market prediction method based on Kalman filter. *Journal of The Institution of Engineers (India) : Series B*. <https://doi.org/10.1007/s40031-021-00583-9>
- [3] 吳宜謙. (2021) . 使用卷積-長短期記憶神經網路進行股票交易 [Master's thesis, 國立清華大學計算與建模科學研究所].
<https://hdl.handle.net/11296/53r486>
- [4] 張士琳. (2018) . 利用情緒指標及長短期記憶模型預測台灣加權股價指數 [Master's thesis, 國立臺北商業大學財務金融研究所].
<https://hdl.handle.net/11296/gnxq3q>
- [5] Güngör, S., Küçün, N. T., & Erol, K. (2022) . Fear of missing out reality in financial investments. *International Journal of Business & Management Studies*, 3 (10) . <https://doi.org/10.56734/ijbms.v3n10a4>
- [6] Hochreiter, S., & Schmidhuber, J. (1997) . Long short-term memory. *Neural Computation*, 9 (8) , 1735–1780.
<https://doi.org/10.1162/neco.1997.9.8.1735>
- [7] Chatterjee, A., Ayadi, O. F., & Maniam, B. (2002) . The applications of the Fibonacci sequence and Elliott wave theory in predicting the security price movements: A survey.
- [8] Wu, J., et al. (2019) . Quantitative trading on stock market based on deep reinforcement learning. In *2019 International Joint Conference on Neural Networks (IJCNN)* (pp. 1-8) .

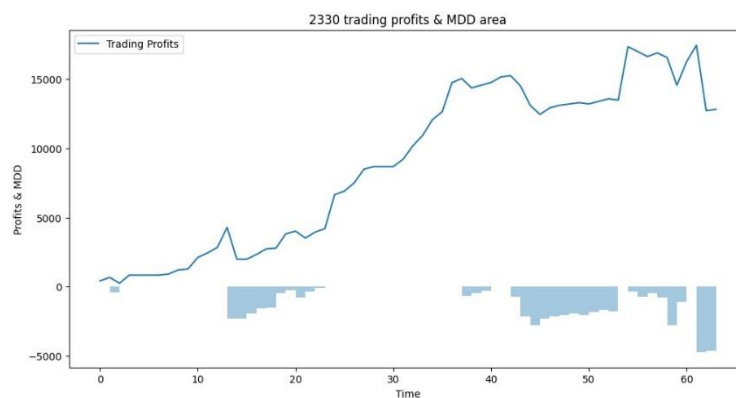
- [9] Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A. A., Veness, J., Bellemare, M. G., Graves, A., Riedmiller, M., Fidjeland, A. K., Ostrovski, G., Petersen, S., Beattie, C., Sadik, A., Antonoglou, I., King, H., Kumaran, D., Wierstra, D., Legg, S., & Hassabis, D. (2015) . Human-level control through deep reinforcement learning. *Nature*, 518 (7540) , 529–533. <https://doi.org/10.1038/nature14236>
- [10]Wikipedia contributors. Bollinger Bands. Wikipedia, The Free Encyclopedia. Retrieved from https://en.wikipedia.org/wiki/Bollinger_Bands
- [11]Tharp, V. K. (2008) . *Van Tharp's Definite Guide to Position Sizing SM: How to Evaluate Your System and Use Position Sizing*.
- [12]SinoTrade. Retrieved from <https://sinotrade.github.io/>
- [13]CoinMarketCap. Retrieved from <https://coinmarketcap.com/zh-tw/rankings/exchanges/>
- [14]Wbbhcb. Retrieved from https://github.com/wbbhcb/stock_market/

附錄

專題製作預定進度甘特圖（Gantt Chart）：整學年度之進度



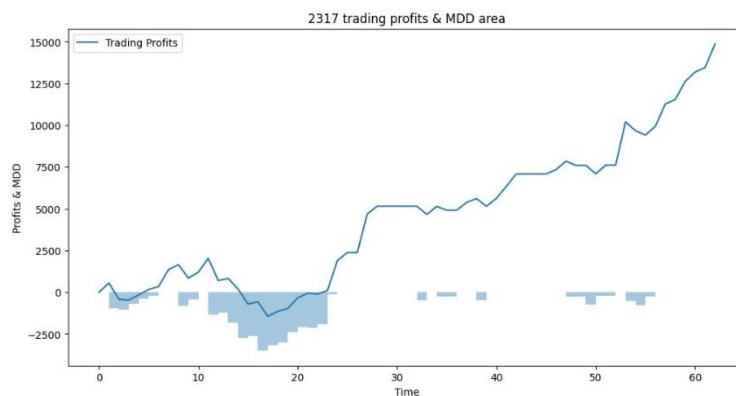
2330 台積電 趨勢策略收益回撤結果圖與統計表



2330 台積電 (趨勢策略)

TP	TT	RR	WR	O	ML	MDD	MAR	SQN
12,823	64	12.82%	68.75%	0.85	-4,732	-4,732	2.70	1.49

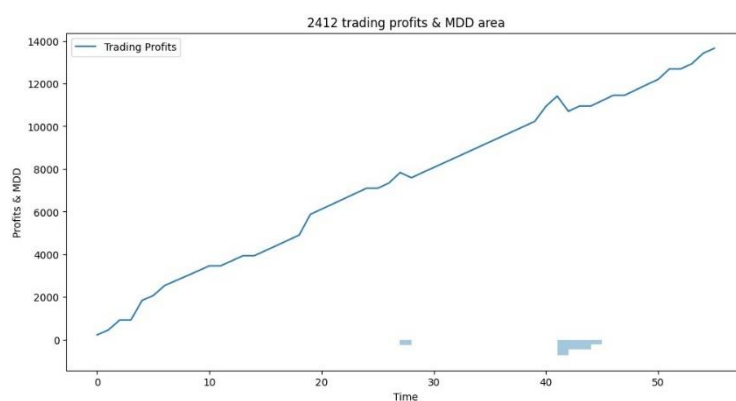
2317 鴻海 趨勢策略收益回撤結果圖與統計表



2317 鴻海 (趨勢策略)

TP	TT	RR	WR	O	ML	MDD	MAR	SQN
14,854	63	14.85%	57.14%	2.09	-1,333	-3,473.3	4.28	2.72

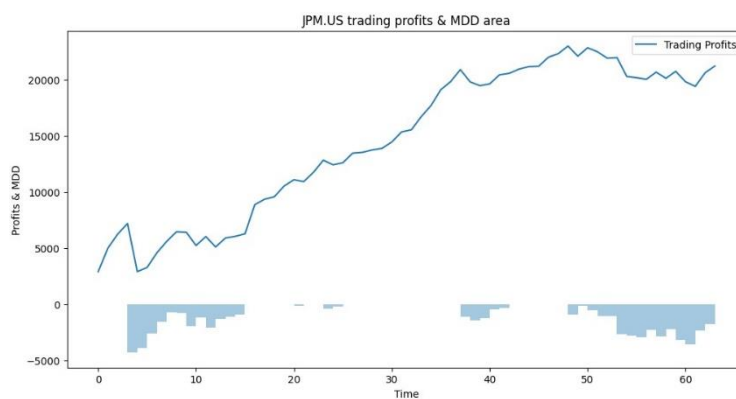
2412 中華電 趨勢策略收益回撤結果圖與統計表



2412 中華電 (趨勢策略)

TP	TT	RR	WR	O	ML	MDD	MAR	SQN
13,650	56	13.65%	83.93%	2.92	-714	-714	19.12	7.74

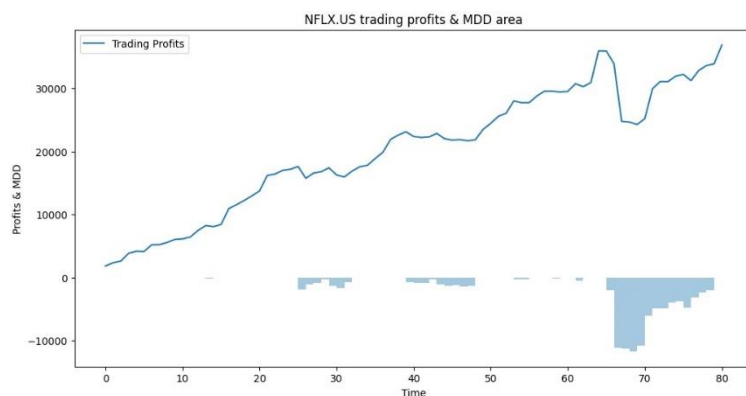
JPM 摩根大通 趨勢策略收益回撤結果圖與統計表



JPM 摩根大通 (趨勢策略)

TP	TT	RR	WR	O	ML	MDD	MAR	SQN
21,209	64	21.20%	73.44%	0.91	-4,286	-4,286	4.95	2.65

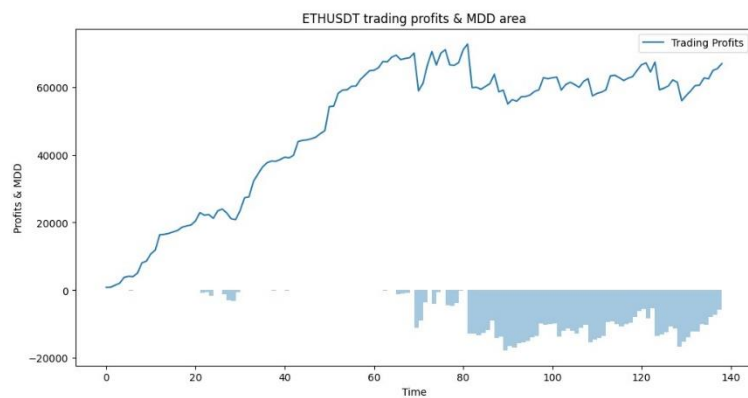
NFLX 網飛 趨勢策略收益回撤結果圖與統計表



NFLX 網飛 (趨勢策略)

TP	TT	RR	WR	O	ML	MDD	MAR	SQN
36,901	81	36.90%	74.07%	1.02	-9,152	-11,687	3.16	2.69

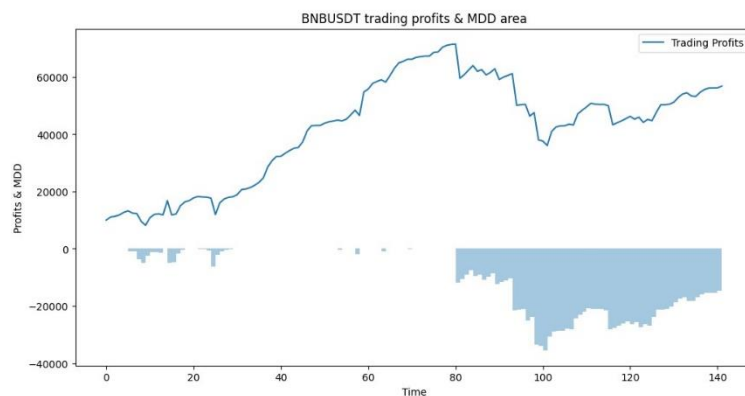
ETH 以太幣 趨勢策略收益回撤結果圖與統計表



ETH 以太幣 (趨勢策略)

TP	TT	RR	WR	O	ML	MDD	MAR	SQN
66,909	139	66.90%	77.70%	0.53	-12,878	-17,706	3.78	2.26

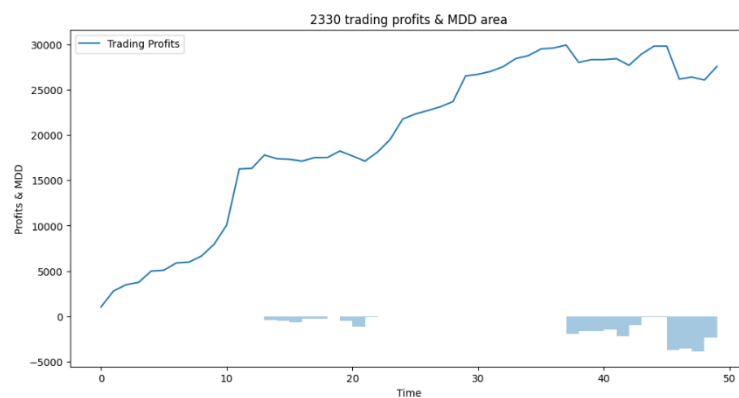
BNB 幣安幣 趨勢策略收益回撤結果圖與統計表



BNB 幣安幣 (趨勢策略)

TP	TT	RR	WR	O	ML	MDD	MAR	SQN
56,847	142	56.84%	74.65%	0.59	-11,960	-35,458	1.60	1.89

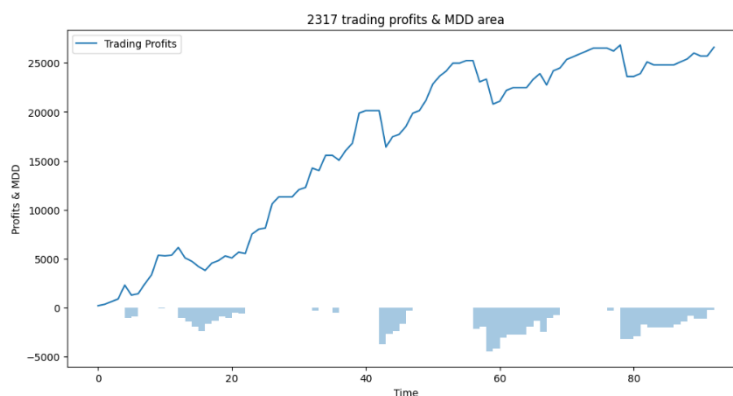
2330 台積電 迴歸策略收益回撤結果圖與統計表



2330 台積電 (迴歸策略)

TP	TT	RR	WR	O	ML	MDD	MAR	SQN
27,584	50	27.59%	76.00%	1.35	-3,636	-3,861	7.14	3.06

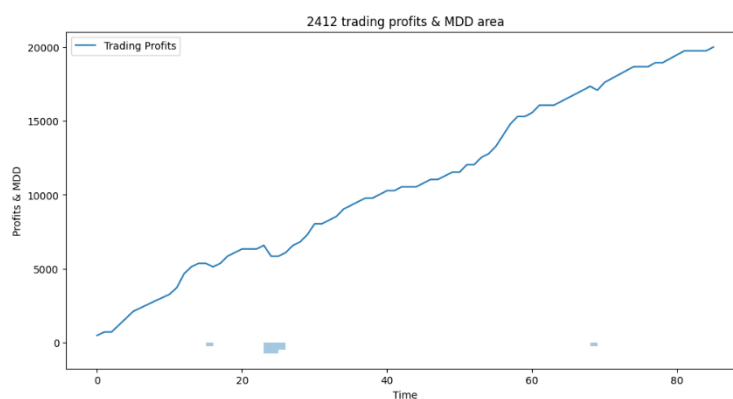
2317 鴻海 迴歸策略收益回撤結果圖與統計表



2317 鴻海 (迴歸策略)

TP	TT	RR	WR	O	ML	MDD	MAR	SQN
26,606	93	26.61%	63.44%	1.42	-3,717	-4,435	6.00	2.82

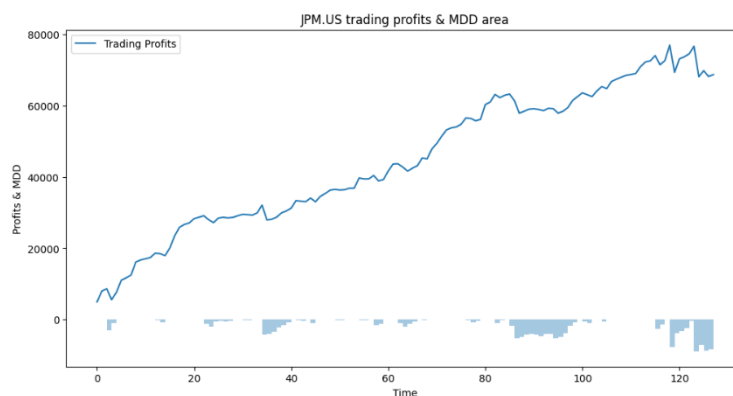
2412 中華電 迴歸策略收益回撤結果圖與統計表



2412 中華電 (迴歸策略)

TP	TT	RR	WR	O	ML	MDD	MAR	SQN
20,019	86	20.02%	70.93%	7.02	-736	-736	27.18	8.98

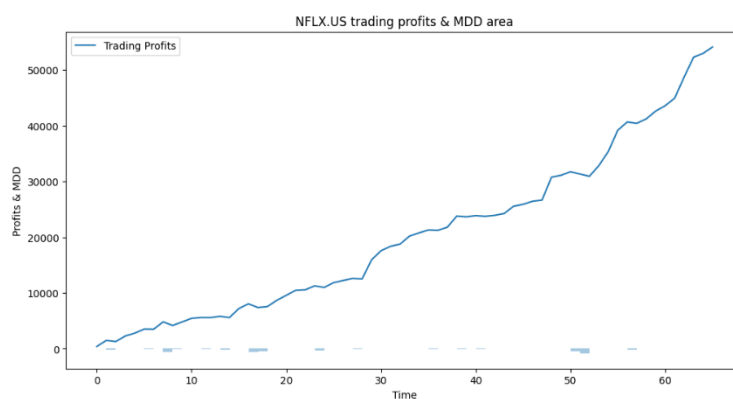
JPM 摩根大通 迴歸策略收益回撤結果圖與統計表



JPM 摩根大通（迴歸策略）

TP	TT	RR	WR	O	ML	MDD	MAR	SQN
68,703	128	68.70%	73.44%	0.89	-8,564	-8,899	7.72	3.40

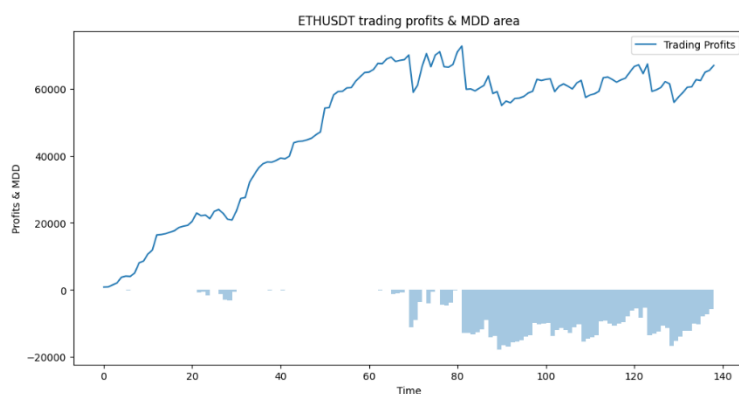
NLFX 網飛 迴歸策略收益回撤結果圖與統計表



NLFX 網飛（迴歸策略）

TP	TT	RR	WR	O	ML	MDD	MAR	SQN
54,122	66	54.12%	78.79%	4.42	-657	-817	66.23	6.30

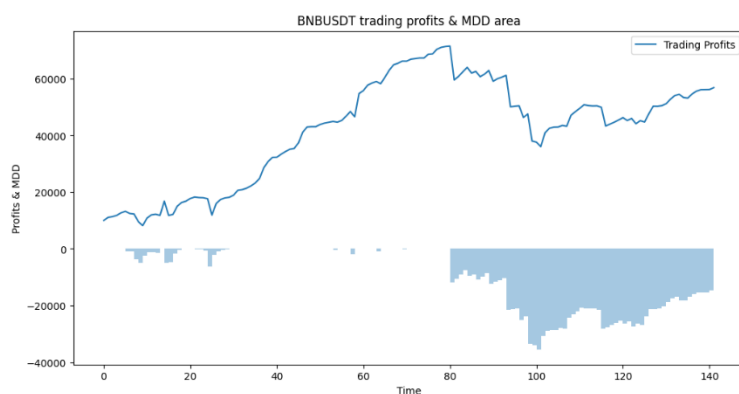
ETH 以太幣 迴歸策略收益回撤結果圖與統計表



ETH 以太幣 (迴歸策略)

TP	TT	RR	WR	O	ML	MDD	MAR	SQN
35,227	26	35.23%	73.08%	1.90	-4,371	-5,019	7.02	2.07

BNB 幣安幣 迴歸策略收益回撤結果圖與統計表



BNB 幣安幣 (迴歸策略)

TP	TT	RR	WR	O	ML	MDD	MAR	SQN
347,100	226	347.10%	71.68%	1.27	-20,958	-27,214	12.75	6.07

作者簡介



姓 名：張婷穎
學 號：41041209
畢業學校：國立虎尾科技大學



姓 名：蔡昀真
學 號：41041213
畢業學校：國立虎尾科技大學



姓 名：黃柏浚
學 號：41041223
畢業學校：國立虎尾科技大學



姓 名：賴益廷
學 號：41041226
畢業學校：國立虎尾科技大學