

國立陽明交通大學
資訊管理與財務金融學系財務金融碩士班
碩士論文

Graduate Program of Finance

National Yang Ming Chiao Tung University
Master Thesis

圖關注網路之股票選取與報酬預測

Stock Selection and Return Prediction via Graph Attention
Networks

研 究 生： 羅頤 (Lo, Yi)

指導教授： 黃宜侯 (Huang, Alex YiHou)

中華民國 一一五年六月

June 2026

圖關注網路之股票選取與報酬預測

Stock Selection and Return Prediction via Graph Attention Networks

研究 生： 羅頤

Student : Yi Lo

指導 教授： 黃宜侯 博士

Advisor : Dr. Alex YiHou Huang



June 2026

Taiwan, Republic of China

中華民國 一一五年六月

誌謝

謝天謝地，感謝國立陽明交通大學提供良好的研究環境與資源，使我能順利完成學業。



圖關注網路之股票選取與報酬預測

學生：羅頤

指導教授：黃宜侯 博士

國立陽明交通大學
資訊管理與財務金融學系財務金融碩士班

摘要

中文摘要就從這邊開始寫。

本研究採用圖注意力網絡 (GAT) 和動態多因子模型 (DMFM)，利用產業圖和全市場圖的結構特性進行個股收益率預測，並通過動態投資組合回測評估模型的實際投資績效。

關鍵字：圖注意力網絡、股票收益率預測、量化投資策略

Stock Selection and Return Prediction via Graph Attention Networks

Student : Yi Lo

Advisor: Dr. Alex YiHou Huang

Graduate Program of Finance
National Yang Ming Chiao Tung University

Abstract

Write your English abstract here. This research employs Graph Attention Networks (GAT) and Dynamic Multi-Factor Models (DMFM) to predict individual stock returns using structural characteristics of industry graphs and market-wide graphs, and evaluates the practical investment performance through dynamic portfolio backtesting.

Keywords: Graph Attention Network, Stock Return Prediction, Quantitative Investment Strategy.

目錄

中文摘要.....	i
英文摘要.....	ii
目錄.....	iii
圖目錄.....	vi
表目錄.....	viii
第 1 章 緒論.....	1
1.1 研究背景.....	1
1.2 研究問題.....	2
1.3 研究方法與貢獻.....	3
1.4 論文架構.....	4
第 2 章 文獻回顧.....	5
2.1 線性迴歸在股票報酬預測中的角色與限制 (Linear Regression)	5
2.2 非線性機器學習模型與特徵工程導向方法 (Tree-based Models)	6
2.3 技術指標如何轉化為深度學習可用之特徵表示.....	8
2.4 LSTM 等時間序列模型在金融預測中的典型做法與延伸	9
2.5 圖神經網路的基本概念與代表性架構.....	11
2.6 GAT 及其變形在股票預測／排序任務上的應用與限制.....	12

2.7 因子訊號評估指標：IC 與 ICIR 的理論意涵與實務角色.....	16
第 3 章 實驗設計.....	17
3.1 資料來源與預處理.....	18
3.1.1 資料來源.....	18
3.1.2 TEJ 資料欄位與研究變數表格	19
3.1.3 資料預處理.....	19
3.1.4 資料集劃分.....	20
3.2 特徵工程與圖結構.....	21
3.2.1 特徵選取與分類.....	21
3.2.2 圖結構設計與建構流程.....	22
3.3 模型設計.....	25
3.3.1 基準模型：GATRegressor.....	25
3.3.2 DMFM 模型架構 (Wei et al. 2022)	25
3.4 損失函數.....	26
3.4.1 總損失函數.....	26
3.5 評估指標.....	27
3.5.1 預測品質指標.....	27
3.5.2 投資組合績效指標.....	28
3.5.3 基準比較與流程摘要.....	28

第 4 章 實驗結果.....	29
4.1 模型對比.....	29
4.1.1 評估指標說明.....	29
4.1.2 模型比較結果.....	30
4.2 投組績效驗證.....	31
4.3 特徵重要性排名.....	33
第 5 章 結論.....	34
5.1 未來研究.....	34
參考文獻.....	35
第 A 章 Appendix 附錄.....	40
A.1 模型輸入特徵完整清單.....	40
A.2 完整實驗輸出補充.....	45
A.2.1 各模型於不同時間窗之完整預測輸出補充（測試集）.....	45
A.2.2 完整累積報酬圖.....	46
A.2.3 Daily IC 補充結果.....	49
A.2.4 IC 分佈補充結果.....	52

圖 目 錄

圖 1 XGBoost 訓練與測試流程示意：多輪 CART 樹建模、權重調整與預測聚 合。	7
圖 2 LSTM 單元的資訊流 (Forget / Update / Output)。其中 x_t 為當期輸入， h_{t-1} 為前一期隱狀態 (hidden state)， c_{t-1} 為前一期記憶狀態 (cell state)。	9
圖 3 GNN 架構示意：由輸入層 X 經過多層圖卷積/訊息傳遞後得到輸出層表 示 Z 。	11
圖 4 GAT 的鄰居注意力 (neighbor attention) 機制示意。左圖呈現注意力係數 α_{ij} 的計算流程；右圖示意節點 i 對其鄰居集合 $\mathcal{N}(i)$ 分配不同權重後進 行聚合，並以 concat/avg 的方式得到更新後的節點表示。	13
圖 5 本文實驗設計流程總覽（對應第三章各小節之處理步驟與順序）	17
圖 6 三種樣本視窗之資料劃分（短期/中期/長期）	21
圖 7 產業圖示意 (Industry Graph)：同產業股票完全連結並加入自環 (self- loop)。圖中每個子圖對應一個產業群集，群集內節點兩兩相連以表示產 業內共同訊號傳遞。	23
圖 8 全市場圖示意 (Universe Graph)：完全連結並包含自環 (self-loop)。在 此圖結構下，每個節點可與所有其他節點交換訊息，以表徵跨產業的市 場共通因子影響。	24

圖 9 edge_index 表示方式示意圖：左側為小型圖結構（節點 A、B、C），右側為其對應之 edge_index。edge_index 以 $[2, E]$ 形式紀錄所有邊，每一欄代表一條由起點（start）指向終點（end）的連結關係。 25

圖 10 各模型在不同時間窗之關鍵預測指標比較(ICIR、IC、RMSE、DirAcc)。

31

圖 11 GAT 與 DMFM 在 Short/Medium/Long 三種時間窗之累積報酬曲線比較。 32

圖 12 DMFM 在不同時間窗的注意力前五特徵權重。 33

圖 13 short 時間窗各模型累積報酬曲線。 46

圖 14 medium 時間窗各模型累積報酬曲線。 47

圖 15 long 時間窗各模型累積報酬曲線。 48

圖 16 short 時間窗各模型 Daily IC。 49

圖 17 medium 時間窗各模型 Daily IC。 50

圖 18 long 時間窗各模型 Daily IC。 51

圖 19 short 時間窗各模型 IC 分佈。 52

圖 20 medium 時間窗各模型 IC 分佈。 53

圖 21 long 時間窗各模型 IC 分佈。 54

表目錄

表 1 TEJ 匯出之原始欄位（資料來源欄位）	18
表 2 模型輸入特徵摘要（完整 54 項清單見附錄表 6）	22
表 3 各模型於不同時間窗之預測指標比較（測試集）	30
表 4 GAT、DMFM 與 0050 於回測區間之績效比較	32
表 5 DMFM 注意力權重前五特徵（依時間窗）	33
表 6 本研究實際使用之模型輸入特徵 (feature_cols，共 54 項；全部由 TEJ 原始欄位衍生)	40
表 7 各模型於不同時間窗之完整預測輸出補充（測試集）	45

第一章、緒論

1.1 研究背景

股票市場的價格形成同時受到公司基本面、產業景氣、整體市場風險偏好與短期交易行為等因素影響，使得報酬序列呈現高雜訊、非線性與時變（time-varying）特性。在量化投資（quantitative investment）實務中，常以可計算的市場訊號作為決策依據，其中技術指標（Technical Indicators, TIs）透過歷史價格與成交量序列，將趨勢、動能與波動等資訊轉換為量化特徵，長期被廣泛用於特徵工程（feature engineering）與交易策略建構。

近年深度學習（deep learning）在金融預測任務上的應用逐漸普及，主要優勢在於能以資料驅動方式學習非線性映射，並在高維特徵下自動萃取有效表示。然而，當技術指標與衍生特徵數量擴張時，特徵集合往往伴隨高度相關、冗餘與尺度不一致，導致模型訓練成本提高、收斂不穩定，並可能削弱泛化能力。換言之，研究焦點逐步由「是否使用技術指標」轉向「如何將技術指標組織為深度模型可有效吸收的表示」。

另一方面，金融市場並非由彼此獨立的資產所構成。台灣股票之間普遍存在產業層級與市場層級的共同因子影響（common factors），使個股報酬在截面（cross-section）上呈現結構性相依。傳統時間序列模型（如 LSTM）多以單一資產的歷史序列為主要訊號來源，較難直接刻畫股票之間的關係拓撲；因此近年研究開始引入圖神經網路（Graph Neural Networks, GNNs），以「節點（股票）—邊（關係）」的方式建模產業或市場連結，並透過訊息傳遞（message passing）與注意力機制（attention mechanism）聚合鄰居資訊，以提升截面預測與排序任務的表現與可解釋性。

在上述脈絡下，結合多因子表示 (multi-factor representation) 與圖注意力 (graph attention) 的模型架構，特別是能分離產業與全市場共同影響的設計，提供了一條兼顧「特徵可學習性」與「關係結構建模」的路徑。本文後續將採用深度多因子模型結合圖注意力之框架，並以台灣日頻資料進行實證檢驗。

1.2 研究問題

綜合既有研究與實務需求，本文關注的核心問題可整理如下。

第一，技術指標所形成的高維特徵集合雖能涵蓋多面向市場訊號，但亦容易引入冗餘與尺度差異，使模型難以穩定學習；因此需要一套可重複、可檢核的特徵建置與整理流程，將價格與量能序列轉換為適合深度模型輸入的特徵表示，並降低不必要的噪音干擾。

第二，股票之間存在產業共通訊號與市場共同波動，僅依賴單一股票的時間序列特徵，可能無法充分利用截面關係資訊；因此需要能顯式建模股票關係的學習架構，使模型同時吸收「個股特徵」與「關係拓撲」兩類訊號。

第三，即使引入圖結構，模型表現仍高度仰賴圖的構建方式與關係品質；若未能有效分離產業與全市場層級的共同影響，模型可能將「產業押注」或「市場曝險」誤判為選股能力。因而本文特別關注：是否能透過分層中性化 (hierarchical neutralization) 的方式，將產業與市場影響自特徵表示中拆解，進而提升預測訊號的穩健性。

第四，在金融預測情境下，模型品質不宜僅以單一誤差指標衡量；更關鍵的是模型在截面排序上的有效性與穩定性。因此本文採用資訊係數 (Information Coefficient, IC) 與其穩健性指標 ICIR (Information Coefficient Information Ratio, ICIR) 作為主要評估基礎，並輔以投資組合績效指標進行整體驗證。

1.3 研究方法與貢獻

本文採用台灣股票日頻資料進行實證。資料主要來源為 TEJ 台灣經濟新報 (Taiwan Economic Journal, TEJ)，並匯出為 CSV 作為後續處理之輸入。以「交易日 × 股票」為觀測單位，保留日期、股票代碼、OHLC、成交量／成交值、市值與產業分類等欄位；技術特徵則由價格與量能序列計算而得，其中 RSI、MACD、KD 等指標透過 TA-Lib 套件產生，最終形成固定的 54 維特徵集合供模型輸入（詳見第三章的變數表格與附錄清單）。

在關係建模方面，本文依第三章之設計建構兩類圖結構：(1) 產業圖 (Industry Graph)，以同產業股票完全連結並加入自環；(2) 全市場圖 (Universe Graph)，以完全圖表示整體市場共同因子影響。圖結構以 PyTorch Geometric (PyG) 的 edge_index 形式保存，並與特徵張量共同形成可重複使用的訓練產物 (artifacts)。

模型設計以 Deep Multi-Factor Model (DMFM) 為主體，透過「特徵編碼 → 產業中性化 → 全市場中性化 → 階層式拼接 → 深度因子學習 → 因子注意力」的流程，在每個交易日的截面上輸出各股票之預測訊號；同時以較簡化的圖模型 (GATRegressor) 作為基準模型，進行可比性的對照實驗。評估上以 IC、Daily IC、ICIR 與產業中性 IC 作為核心預測指標，並以投資組合回測指標（年化報酬率、Sharpe ratio、勝率等）檢驗策略層級的有效性；資料切分採時間序列切分以避免資訊洩漏，並設計不同樣本長度（短期／中期／長期）檢驗穩健性。

基於上述方法設計，本文的主要貢獻歸納如下：

- **資料與特徵流程的可重現性：**以 TEJ 日頻資料為基礎，建立從原始欄位到 54 維技術特徵（含 TA-Lib 指標）的完整特徵建置與張量化流程，並以 artifacts 形式保存，便於重複訓練與評估。
- **關係結構的分層建模：**同時建構產業圖與全市場圖，並在模型中引入分層中性化

機制，以拆解產業與市場共同影響，提升截面預測訊號的穩健性與可解釋性。

- **以排序導向指標為核心的實證驗證：**以 IC/ICIR 與產業中性 IC 作為主要評估基礎，並搭配投資組合績效指標，提供模型在台灣市場情境下的多面向驗證結果。

1.4 論文架構

本文共分為五章，各章內容如下。

第一章為緒論，說明研究動機與背景、研究問題、研究方法與預期貢獻，並概述全文架構。

第二章為文獻回顧，依序整理技術指標在深度學習中的特徵表示、時間序列模型（如 LSTM）於金融預測的典型做法、圖神經網路的基本概念與代表性架構，以及圖注意力模型（GAT）在股票預測／排序任務上的應用與限制，作為後續方法設計之理論與實證依據。

第三章為實驗設計，說明資料來源與預處理流程、特徵工程與圖結構建構方式、模型架構（基準模型 GATRegressor 與 DMFM）、損失函數設定，以及評估指標與實驗切分規則。

第四章為實驗結果，呈現不同模型之預測表現比較、IC／ICIR 與產業中性 IC 之結果分析，並以投資組合績效驗證模型訊號在策略層面的有效性與穩健性（含不同樣本長度設定之比較）。

第五章為結論，總結本文之主要發現，討論方法限制與可能的改進方向，並提出後續在實務應用與研究延伸上的建議。

第二章、文獻回顧

本章依循金融預測方法由淺入深的研究脈絡，依序回顧：(1) 線性迴歸 (Linear Regression, LR) 在報酬預測中的典型定位與限制，(2) tree-based models 在非線性預測與特徵工程中的角色，(3) 技術指標如何轉化為深度學習可用之特徵表示，(4) LSTM 等時間序列模型在金融預測中的典型做法與延伸，(5) 圖神經網路 (GNN) 如何建模股票之間的結構性關係，(6) GAT 及其變形在股票預測／排序任務上的應用與限制，(7) 因子投資情境下常用的訊號評估指標 IC／ICIR，並據此銜接本研究後續以「多因子表示 + 圖注意力」為核心之模型設計與實驗流程。

2.1 線性迴歸在股票報酬預測中的角色與限制 (Linear Regression)

線性迴歸模型 (Linear Regression, LR) 在股票報酬預測中長期扮演基準方法 (baseline) 角色，核心優勢是結構簡潔、可解釋性高，且便於對照不同模型的增量效果。其方法論基礎可追溯至經典因子定價文獻：Fama 與 French (1993) 指出，市場、規模、帳面市值比等共同因子可解釋股票與債券報酬的平均差異，奠定了以線性因子暴露衡量風險溢酬的實證框架。^[1]

在線性框架下，動能 (momentum) 亦成為重要延伸。Jegadeesh 與 Titman (1993) 發現，買進過去贏家、放空過去輸家的策略在 3–12 個月持有期具有顯著正報酬，且並非完全由傳統系統性風險所解釋；此結果支持在線性多因子模型中納入動能相關訊號，以提升截面報酬解釋能力。^[2]

然而，近年研究顯示僅依賴線性假設仍有侷限。Gu、Kelly 與 Xiu (2020) 在大規模資產定價比較中指出，樹模型與神經網路可透過非線性與特徵交互作用捕捉額外可預測性，並在經濟績效上顯著優於傳統回歸策略；同時，其辨識出的關鍵訊號仍與動能、流動性、波動度等經典因子高度一致。^[3] 其他截面研究亦顯示，機器學習相對 OLS 可帶來額外預測增益，反映線性模型在高維與非線性情境下的擬合能力限制。^[4]

綜合而言，LR 的主要限制包含：(1) 加總線性假設難以刻畫非線性反應與高階交互作用，(2) 高維特徵下易受共線性與參數不穩定影響，(3) 對時間變動關係與 regime shift 的適應能力有限。此亦是後續文獻轉向 tree-based 與深度學習模型的重要動機。

2.2 非線性機器學習模型與特徵工程導向方法（Tree-based Models）

為克服線性模型的表達能力限制，tree-based models（如 Decision Tree、Random Forest、Gradient Boosting、XGBoost、LightGBM、CatBoost、Extra Trees 等）在金融預測文獻中逐漸成為主流方法之一。相較於單一決策樹，集成式（ensemble）模型往往在股票方向預測、波動度預測與企業風險辨識任務上展現更穩健且一致的表現，並常被視為強力的 baseline。^[5, 6, 7]

tree-based 方法的核心優勢在於能自動捕捉特徵的非線性關係與高階交互作用，並且可透過特徵重要性（feature importance）或 SHAP 類方法提供一定程度的解釋性。在金融預測應用中，相關研究亦指出 tree-based ensembles 在技術指標輸入下可支持交易規則設計，並在模擬環境中呈現相對於基準的超額報酬；此外，在波動度與風險管理任務上，樹模型亦具備一定效率優勢。^[8, 9]

在 gradient boosting 家族中，XGBoost 透過二階梯度近似、顯式正則化、縮減（shrinkage）、欄位抽樣與對稀疏特徵友善的分裂策略，顯著提升了樹集成模型在大規模結構

化資料上的訓練效率與泛化能力；因此在金融 tabular 預測任務中，常被視為兼具效能與可解釋性的強基準模型。^[10] 依圖 1 的流程，XGBoost 先由同一批訓練資料建立多輪 CART 樹 ($f(L_1), f(L_2), \dots, f(L_n)$)，每一輪透過 *weighting increase* 強化前一輪較難擬合樣本的權重，再產生對應的分輪預測 $\hat{y}_L^{(1)}, \hat{y}_L^{(2)}, \dots, \hat{y}_L^{(N)}$ ；最後將各輪預測做聚合得到最終輸出。此機制的重點在於「逐輪修正誤差 + 集成輸出」，因此通常較單一決策樹更穩健。^[10]

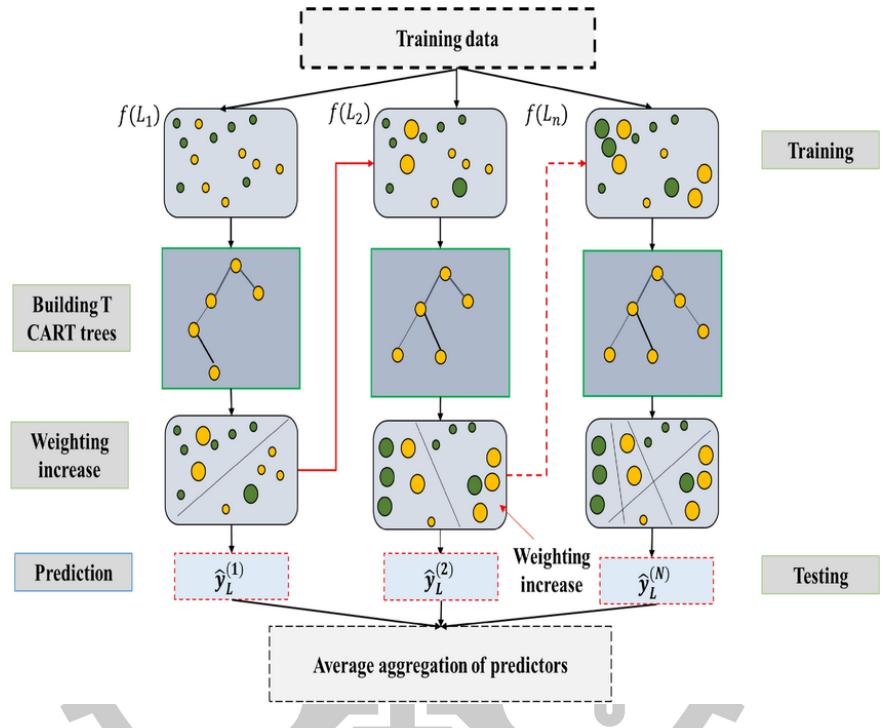


圖 1: XGBoost 訓練與測試流程示意：多輪 CART 樹建模、權重調整與預測聚合。

另一方面，較早期的決策樹交易研究已顯示，若將技術分析的候選交易點（如 filter rule）交由 C4.5 演算法進行篩選與分類，可在台灣與 NASDAQ 市場取得優於單純 filter rule 的績效，說明 decision tree 不僅可用於純預測，也可作為規則型交易策略的決策器。^[11]

然而，tree-based models 多將每個樣本視為相對獨立的觀測點，對時間相依結構的刻畫通常依賴人工設計的 lagged features 或 rolling statistics；當市場進入快速結構轉換或非平穩狀態 (non-stationarity) 時，僅靠靜態特徵映射可能不足以穩健追蹤動態行為。因此，後續研究開始引入以序列建模為核心的深度學習架構（如 LSTM），並進一步發

展 hybrid 模型，以結合 tree-based 的非線性擬合優勢與深度模型的時間依賴建模能力。

[12]

2.3 技術指標如何轉化為深度學習可用之特徵表示

技術指標 (Technical Indicators, TIs) 以歷史價格與成交量為基礎，將市場的趨勢、動能與波動等訊號轉換為可計算的量化特徵，長期被用於股票預測任務的特徵工程。當指標種類擴張時，特徵集合往往同時伴隨高度相關、冗餘與尺度不一致，使模型訓練成本上升，並可能影響泛化能力與穩健性。[13, 14, 15]

從文獻脈絡來看，技術指標的價值並非僅止於「交易規則」，而是可視為一種可重用的訊號轉換層。經典研究已顯示，移動平均與區間突破等簡單規則在長期樣本下具有統計上的可辨識性，代表價格序列中可能存在可被轉換後捕捉的結構訊號。[13] 後續研究則進一步指出，若先對多個技術指標進行對齊與降噪（例如以 partial least squares 提取共同變動成分），可提升報酬預測的樣本內與樣本外表現，說明「指標表示品質」本身就是預測力來源之一。[14] 在更高頻與型態辨識情境下，結合技術分析與模式辨識的交易規則亦可形成風險調整後可解釋的訊號，反映技術特徵在不同時間尺度下都可能提供增量資訊。[16]

因此，研究焦點逐步由「是否使用技術指標」轉向「如何將技術指標組織為深度學習可有效吸收的特徵表示」。實務上常見流程可分為四步。第一，**多視窗指標化**：由 OHLCV 建立趨勢 (MA/EMA)、動能 (RSI/MACD/KD)、波動 (ATR/std)、量能與流動性等指標族，並在不同視窗長度上展開，以保留短中長期訊號。第二，**尺度與分佈對齊**：對各指標進行 rolling z-score、rank normalization、winsorization 或中性化處理，降低厚尾分佈與異常值對梯度學習的干擾。第三，**時序化表示**：將單期指標向量堆疊為時間窗張量（例如 $X_t \in \mathbb{R}^{L \times F}$ ），並加入一階/二階差分、斜率或 regime proxy，以提升模型對轉折與狀態切換的敏感度。第四，**冗餘抑制與語義壓縮**：透過相關性過濾、

PLS/PCA、可學習 gating 或自編碼器，將高度共線的原始指標壓縮為較穩定的潛在因子表示。[14, 17, 18, 15, 19, 12]

在此基礎上，技術指標不再只是「手工規則」，而是作為深度模型的結構化輸入介面。對序列模型（LSTM/Transformer）而言，指標張量提供可被注意力或記憶機制吸收的時間上下文；對圖模型（GNN/GAT）而言，指標向量則可作為節點屬性，與資產關係邊共同學習截面與網路效應。相關研究亦指出，多指標組合能提供較完整的市場訊號，但其效益高度依賴指標集合的挑選與組織方式。[17, 18, 15]

總結而言，技術指標在深度學習框架中的關鍵角色是「將高噪音價格路徑轉換為可學習表示」：有效方法不是無限制增加指標數量，而是透過對齊、時序化與壓縮，提升訊號密度、抑制冗餘，並使表示能與模型結構共同最佳化，以提高跨期間的穩健性與可轉移性。[12, 19]

2.4 LSTM 等時間序列模型在金融預測中的典型做法與延伸

時間序列模型在金融預測中主要利用價格與特徵序列的時間依賴性，學習由過去資訊推估未來表現。長短期記憶網路（Long Short-Term Memory, LSTM）透過閘門（gates）機制控制資訊流動，能緩解長序列學習的梯度消失（vanishing gradients）問題，因此成為金融時間序列建模的代表性方法之一。

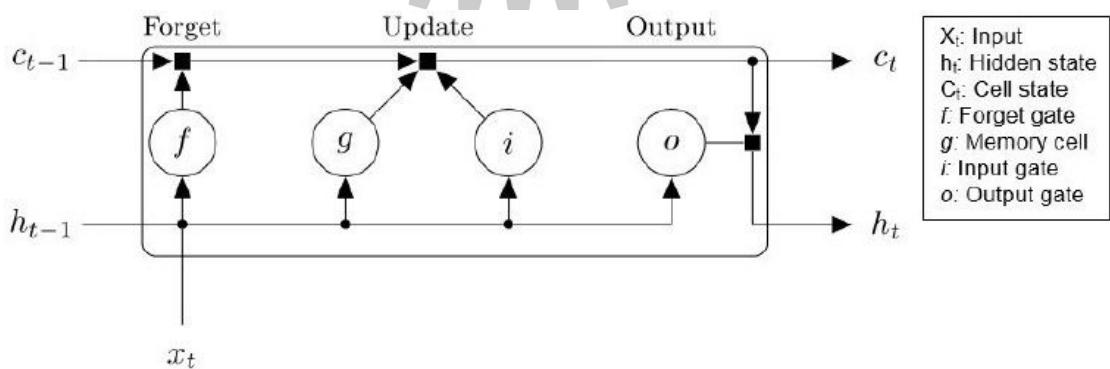


圖 2: LSTM 單元的資訊流（Forget / Update / Output）。其中 x_t 為當期輸入， h_{t-1} 為前一期隱狀態（hidden state）， c_{t-1} 為前一期記憶狀態（cell state）。

如圖 2 所示，LSTM 會沿著記憶狀態（cell state）主幹 $c_{t-1} \rightarrow c_t$ 保留長期資訊，並透過三個閘門控制「遺忘（Forget）」、「更新（Update）」與「輸出（Output）」：(1) 遺忘閘門 f_t 決定保留多少過去記憶 c_{t-1} ；(2) 更新步驟由輸入閘門 i_t 與候選記憶 g_t （圖中 memory cell）共同決定要寫入多少新資訊；(3) 輸出閘門 o_t 決定由當期記憶 c_t 產生多少輸出到隱狀態 h_t 。對應的計算可寫為：

$$\begin{aligned}
 f_t &= \sigma(W_f[h_{t-1}, x_t] + b_f), \\
 i_t &= \sigma(W_i[h_{t-1}, x_t] + b_i), \\
 g_t &= \tanh(W_g[h_{t-1}, x_t] + b_g), \\
 c_t &= f_t \odot c_{t-1} + i_t \odot g_t, \\
 o_t &= \sigma(W_o[h_{t-1}, x_t] + b_o), \\
 h_t &= o_t \odot \tanh(c_t),
 \end{aligned} \tag{2.1}$$

其中 $\sigma(\cdot)$ 為 sigmoid 函數、 $\tanh(\cdot)$ 為雙曲正切函數， \odot 表示逐元素相乘（element-wise multiplication）。圖中的方形節點可視為逐元素乘法與加總的運算節點，對應到 $f_t \odot c_{t-1}$ 與 $i_t \odot g_t$ 兩條路徑匯入 c_t 的更新。

在實務與研究中，LSTM 的延伸方向常包含多目標預測、結合注意力機制（attention mechanism）與去噪／重加權策略，以及高頻資料下的深層序列建模。Zaheer 等人設計多參數預測架構以同時預測不同價格維度，並比較不同深度模型配置在特定資料規模下的表現差異。^[20] Qiu 等人將注意力機制引入 LSTM 架構，以在時間維度上動態調整特徵權重，強化模型在金融資料高雜訊情境下的預測能力。^[21] 在高頻（例如 5 分 K）資料情境下，鄭邦廷以疊層式 LSTM（Stacked LSTM）捕捉更複雜的非線性時間依賴，用於買賣點預測。^[22] 此外，廖俊翔透過自相關分析與特徵篩選納入跨市場外部特徵，並展示外部訊號在一定程度上能降低預測誤差。^[23]

綜合金融時間序列預測的系統性綜述研究，深度序列模型（LSTM/CNN/Hybrid）在

預測誤差指標上多能優於傳統統計模型與部分經典機器學習方法，然而研究亦普遍指出：在金融市場高雜訊與非平穩環境下，單純改善 error metrics 並不必然等價於可穩健轉化為可交易的超額績效，顯示實證評估必須結合更貼近投資目標的指標與測試框架。
[12, 24]

整體而言，時間序列方法能有效吸收單一資產（或單一特徵集合）的歷史資訊，但其多數設計仍以「序列本身」為主要訊號來源，較難直接刻畫股票之間的結構性關係與共通影響；因此後續研究開始引入圖結構以建模市場關係。

2.5 圖神經網路的基本概念與代表性架構

圖神經網路(Graph Neural Networks, GNNs)以圖作為基本資料結構，透過節點(node)與邊(edge)描述實體及其關係，並以訊息傳遞(message passing)機制聚合鄰居資訊，學得具備拓撲語意的節點表示。GNN 的基本概念最初由 Scarselli 等人提出，他們引入遞迴神經網路框架以處理圖結構資料，並透過疊層訊息傳遞與節點狀態更新來學習圖表示。[25] 後來，Wu 等人對 GNN 架構進行系統性整理，並區分頻域(spectral-based)與空域(spatial-based)圖卷積之主要差異，同時歸納多類代表性 GNN 架構族群。[26]

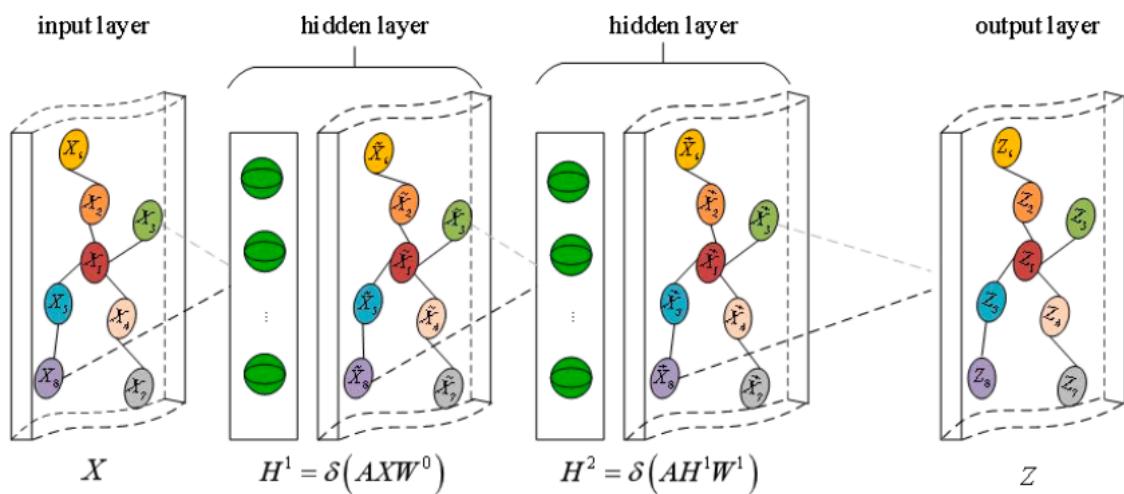


圖 3: GNN 架構示意：由輸入層 X 經過多層圖卷積/訊息傳遞後得到輸出層表示 Z 。

如圖 3 所示，GNN 先以輸入層 X 表示節點初始特徵，接著在隱藏層透過鄰接矩陣 A 與權重矩陣 W 做鄰居聚合與線性轉換，例如 $H^1 = \delta(AXW^0)$ 、 $H^2 = \delta(AH^1W^1)$ 。其中 $\delta(\cdot)$ 為非線性函數。經過多層更新後，模型在輸出層得到節點表示 Z ，可用於節點分類、排序或報酬預測等任務。

在金融市場中，股票並非獨立個體，而是受到產業鏈、供應鏈、風格因子與市場共同訊號等多來源關係影響，進而產生共動性（co-movement）與傳染效應（spillover）。近期研究指出，將股票視為「互動節點」並以圖結構建模其關係（例如產業關係、價格相關性、供應鏈或知識圖譜），可使模型在預測報酬、方向與風險指標時，相較於僅使用序列模型的 baseline 呈現更佳表現，特別在市場波動或危機期間更具優勢。[27, 28, 29]

因此，圖結構方法為金融預測提供了「同時利用個股特徵與關係拓撲」的建模途徑，也為後續在多因子框架中處理共通影響、關係層級分離與中性化（neutralization）奠定方法論基礎。

2.6 GAT 及其變形在股票預測／排序任務上的應用與限制

圖注意力網路（Graph Attention Network, GAT）在圖神經網路的鄰居聚合（neighbor aggregation）框架中引入注意力機制，使模型能在不同鄰居節點之間**自動學習差異化的權重分配**，從而提升在異質關係（heterogeneous relations）與噪音邊（noisy edges）存在時的表徵能力與可解釋性。在金融市場中，股票之間的連結可能同時來自產業關係、共同市場因子、或價格共動等多重來源，其影響強度亦可能隨時間改變；因此，相較於固定權重的鄰居平均聚合，GAT 以資料驅動方式決定「哪些關係在當期更重要」，在概念上更符合金融資料非平穩（non-stationary）與關係強度變動的特性。

如圖 4 所示，GAT 的核心是對節點 i 與其鄰居 $j \in \mathcal{N}(i)$ 計算注意力係數 α_{ij} ，並以

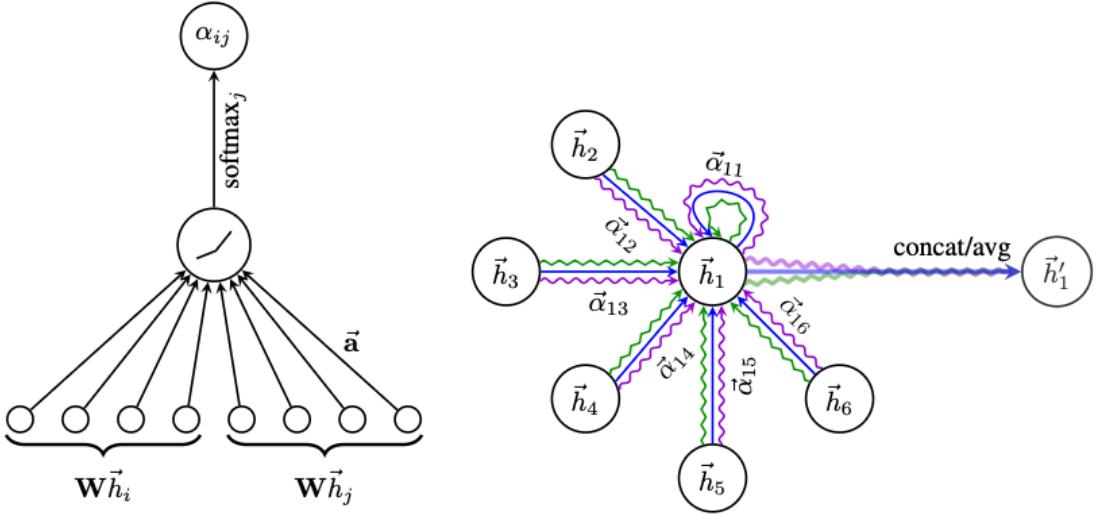


圖 4: GAT 的鄰居注意力 (neighbor attention) 機制示意。左圖呈現注意力係數 α_{ij} 的計算流程；右圖示意節點 i 對其鄰居集合 $\mathcal{N}(i)$ 分配不同權重後進行聚合，並以 concat/avg 的方式得到更新後的節點表示。

加權聚合方式更新節點表示。以注意力係數為例，其形式可寫為：

$$\alpha_{ij} = \text{softmax}_j \left(\text{LeakyReLU} \left(a^\top [Wh_i \parallel Wh_j] \right) \right), \quad (2.2)$$

其中 h_i 為節點表示， α_{ij} 代表節點 i 對鄰居 j 的注意力權重。上述表示式可視為將線性投影、注意力打分與 softmax 正規化三個步驟合併表示；為更清楚對應金融應用中的「關係加權」直覺，下文將其拆解如下。

(1) 特徵線性投影 (Linear Projection) 首先將節點特徵投影至共享表示空間：

$$z_i = Wh_i, \quad (2.3)$$

其中 W 為可學習參數矩陣， z_i 表示節點 i 的轉換後特徵。

(2) 注意力打分 (Attention Scoring) 接著對每一條邊 (i, j) 計算未正規化的注意力分數：

$$e_{ij} = \text{LeakyReLU}\left(a^\top [z_i \| z_j]\right), \quad (2.4)$$

其中 a 為可學習參數向量， $[\cdot \| \cdot]$ 表示向量串接 (concatenation)。直觀而言， e_{ij} 衡量「鄰居 j 的資訊對節點 i 的更新是否重要」。

(3) 正規化與鄰居聚合 (Normalization & Aggregation) 為使同一節點的鄰居權重可比較，GAT 對 e_{ij} 在鄰居集合內做 softmax 正規化：

$$\alpha_{ij} = \frac{\exp(e_{ij})}{\sum_{k \in \mathcal{N}(i)} \exp(e_{ik})}. \quad (2.5)$$

最後，以注意力權重對鄰居特徵加權求和，得到更新後的節點表示：

$$h'_i = \sigma \left(\sum_{j \in \mathcal{N}(i)} \alpha_{ij} z_j \right), \quad (2.6)$$

其中 $\sigma(\cdot)$ 為非線性函數。此聚合形式可理解為：節點 i 不再平均吸收所有鄰居資訊，而是以 α_{ij} 作為「關係重要性」進行選擇性吸收，從而降低弱關係或噪音邊對表徵的干擾。

多頭注意力 (Multi-head Attention) 在實作上，GAT 常使用多頭注意力 (multi-head attention) 提升穩定性與表徵能力。對於 K 個 head，可表示為：

$$h_i'^{(k)} = \sigma \left(\sum_{j \in \mathcal{N}(i)} \alpha_{ij}^{(k)} W^{(k)} h_j \right), \quad k = 1, \dots, K. \quad (2.7)$$

常見作法是在中間層以串接 (concat) 整合多頭資訊：

$$h'_i = \parallel_{k=1}^K h_i'^{(k)}, \quad (2.8)$$

而在最後一層以平均（avg）穩定輸出：

$$h'_i = \frac{1}{K} \sum_{k=1}^K h_i'^{(k)}. \quad (2.9)$$

GAT 在股票預測／排序任務中的應用。在股票預測任務中，GAT 類方法的效能高度依賴圖結構建構方式與關係訊號品質。現有研究整理指出，節點特徵常結合 OHLCV、技術指標、基本面或文字／情緒訊號，並以產業連結、價格相似性、供應鏈關係或動態學習圖等方式建立邊；在此框架下，GAT 能透過注意力機制學習「誰在影響誰」以及影響強度，進而提升截面預測或排序任務的表現。[30, 31, 32]

此外，金融領域亦出現大量 spatio-temporal GAT 變形（如 STGAT、FSTGAT 等），透過結合時間編碼器（例如 LSTM/TCN/Transformer）與圖注意力聚合，同時刻畫時間依賴與資產關係結構；相關研究報告顯示，這類模型在部分情境下可帶來預測誤差改善，並在以投資績效衡量的實證框架中展現更高的應用潛力。[27, 28]

限制與挑戰。儘管 GAT 類方法在金融預測中提供更具彈性的關係建模能力，但仍可能面臨下列限制。第一，當圖中存在大量噪音邊或關係訊號品質不足時，注意力權重可能無法穩定聚焦於有效鄰居，導致模型對邊的品質高度敏感並增加過擬合風險。第二，圖建構規則、關係來源與連邊密度會直接影響訊息傳遞範圍與聚合分佈，使模型在跨市場或跨期間的可轉移性（transferability）與穩健性（robustness）下降。第三，若採用靜態圖，模型可能難以充分反映市場結構與股票關係隨時間演化的特性，進而降低對結構變動的敏感度。因此，在金融預測中使用 GAT 類方法時，除模型設計外，更需搭配嚴謹的資料切分方式、跨期間檢驗、與穩健性測試，以確保其績效增益並非僅來自特定樣本區間或特定關係建構假設。

2.7 因子訊號評估指標：IC 與 ICIR 的理論意涵與實務角色

在因子投資（factor investing）與截面預測（cross-sectional prediction）情境中，單純以 MSE/MAE 等誤差衡量模型表現，未必能完整反映訊號在「排序（ranking）」與「可投資性（investability）」上的價值。因此，資訊係數（Information Coefficient, IC）與資訊比率（Information Ratio, IR；亦常以 ICIR 表示）成為衡量因子訊號品質的重要指標。IC 通常被定義為因子分數與下一期報酬之截面相關（常使用 rank correlation），用以衡量訊號對未來報酬的預測能力；ICIR 則進一步將訊號的平均效果與波動性整合，用以刻畫訊號的穩健性。^[33, 34]

經典的主動管理基本法則（fundamental law）指出，策略的資訊比率可近似表示為 $IR \approx IC \times \sqrt{breadth}$ ，意即在訊號品質（IC）固定下，獨立下注機會（breadth）越多，策略的風險調整績效越高。然而近期研究亦強調，真實市場中 IC 往往具有時間變動性（time-varying），且 IC 的波動（IC volatility）會直接削弱策略可達成的穩健 IR；此外，換手率（turnover）與交易成本亦會顯著壓低可實現的績效，因此在實務上必須將 turnover-adjusted IR 納入考量。^[35, 36, 37]

因此，IC / ICIR 除了作為單一因子有效性的檢驗工具外，也常被用於多因子組合與機器學習選因流程：研究通常保留具備穩定正 IC 與較高 ICIR 的訊號，以提升多因子架構在跨期間的穩健性與可投資性。^[38, 39]

第三章、實驗設計

本章旨在說明本文之整體實驗設計與實作流程，內容涵蓋資料來源與預處理、特徵工程與圖結構建構、模型架構與訓練設定、損失函數設計，以及評估指標與策略驗證方法。本文以台灣股票市場日頻資料為研究對象，重點檢驗 Deep Multi-Factor Model (DMFM) 在超額報酬率預測任務上的表現，並與基準模型 GATRegressor 進行系統性比較。為使讀者能快速掌握各小節之方法銜接與處理順序，圖 5 彙整本文完整實驗流程，後續將依據該流程逐步展開說明。

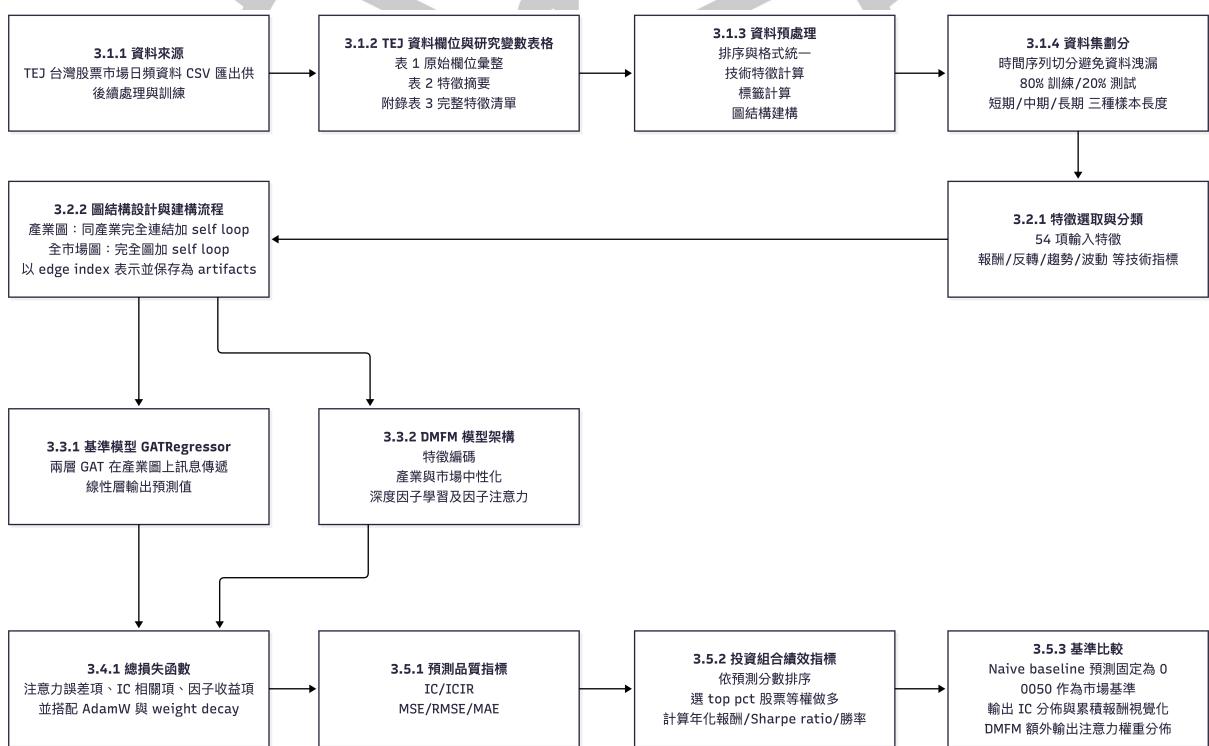


圖 5: 本文實驗設計流程總覽（對應第三章各小節之處理步驟與順序）

3.1 資料來源與預處理

3.1.1 資料來源

資料取自 TEJ 台灣經濟新報 (Taiwan Economic Journal, TEJ) 之台灣股票市場日頻資料，並以 CSV 形式匯出以供後續處理與模型訓練使用。資料以「交易日 × 股票」為觀測單位，至少包含日期 (Date)、股票代碼 (Stock ID) 與收盤價 (Close) 等必要欄位；同時保留開高低收 (OHLC)、成交量／成交值、市值、流通股數、估值指標 (如 PB、PS) 與產業分類等資訊，以支援特徵建置與圖結構構建。樣本期間由參數 start_date 與 end_date 控制，並以台股交易日序列作為時間索引基準。

除 TEJ 可直接取得之原始欄位外，模型輸入之技術特徵由價格與成交量等基礎序列計算而得；其中 RSI、MACD、KD 等指標主要透過 TA-Lib 套件生成，再彙整為固定的特徵集合供模型輸入。為避免於文字中重複列舉欄位與變數定義，TEJ 原始欄位彙整於表 1，最終模型輸入之 54 個特徵摘要見表 2 (完整清單見附錄表 6)。

表 1: TEJ 匯出之原始欄位 (資料來源欄位)

TEJ 欄位名稱	用途	備註 (單位/說明)
年月日	時間索引	交易日 (Date)
開盤價 (元)	交易價格	Open (元)
最高價 (元)	交易價格	High (元)
最低價 (元)	交易價格	Low (元)
收盤價 (元)	交易價格	Close (元)
成交量 (千股)	交易量能	Volume (千股)
成交值 (千元)	交易金額	Amount (千元)
本益比-TEJ	估值資訊	Price-to-Earnings (TEJ 口徑)
股價淨值比-TSE	估值資訊	Price-to-Book (TSE 口徑)
證券代碼 _ 純代碼	個股識別	Stock ID
證券名稱	個股識別	公司名稱
TEJ 產業 _ 名稱	產業資訊	產業分類 (TEJ 分類；用於建構產業圖)

3.1.2 TEJ 資料欄位與研究變數表格

所有輸入資料均由 TEJ 匯出為日頻 CSV。匯出檔案包含：(1) 市場交易與公司屬性之原始欄位（表 1），以及 (2) 由程式以 TEJ 原始欄位衍生之技術特徵集合 `feature_cols` ($n=54$)（表 2）。

3.1.3 資料預處理

資料預處理主要涵蓋資料格式統一、特徵與標籤計算、以及張量化整理，並以一致規則處理缺失值以降低訓練偏誤。整體流程如演算法 1 所示。

Algorithm 1 資料預處理流程

- 1: 讀取 TEJ 匯出之日頻資料（至少包含 Date、Stock ID、Close；其餘欄位供特徵與圖結構使用）
 - 2: 依 Stock ID 與 Date 排序；進行日期格式與數值型態轉換，並做基本缺漏/異常檢核
 - 3: 依價格與量能序列計算技術特徵（含 TA-Lib 指標），形成 `feature_cols`
 - 4: 計算標籤： $y_t = \frac{P_{t+k} - P_t}{P_t}$ # k 日未來報酬率 (forward return)
 - 5: 由產業欄位建構產業圖 E_{ind} ；由股票清單建構全市場圖 E_{uni}
 - 6: 建立特徵張量 $F^t \in \mathbb{R}^{T \times N \times F}$ 與標籤張量 $y^t \in \mathbb{R}^{T \times N}$
 - 7: 特徵 NaN/Inf 以 0 取代；標籤 NaN 保留並於訓練/評估以 mask 排除
 - 8: 特徵正規化由模型端 Batch Normalization (BatchNorm) 處理
-

缺失值處理 (Missing values)：不對缺失值進行插補，以避免引入主觀假設。特徵層面將 NaN/Inf 以 0 取代以維持張量完整性；標籤層面保留 NaN，並在訓練與評估時以有效樣本遮罩 (mask) 排除，使模型僅在可觀測標籤上學習。

特徵標準化 (Normalization)：資料前處理階段不額外做截面標準化，特徵保留原始尺度；模型端於輸入層使用 **Batch Normalization (BatchNorm)** 進行正規化：

$$x_{normalized} = \frac{x - \mu_{batch}}{\sqrt{\sigma_{batch}^2 + \epsilon}} \quad (3.1)$$

其中 μ_{batch} 與 σ_{batch}^2 為 mini-batch 統計量， ϵ 為數值穩定項（預設 10^{-5} ）。

標籤定義 (Label definition) : 標籤定義為未來 k 日報酬率 (forward- k return) :

$$y_t = \frac{P_{t+k} - P_t}{P_t} \quad (3.2)$$

其中 P_t 為時間 t 收盤價， k 為預測視窗（預設 $k = 5$ ）。

3.1.4 資料集劃分

為避免模型訓練過程產生未來資訊滲入 (look-ahead bias)，本文採用時間序列切分 (time-series split) 進行訓練集與測試集的劃分。具體而言，資料依日期由早至晚排序後，以前 80% 交易日作為訓練集 (Train)，後 20% 交易日作為測試集 (Test)，確保模型僅能利用過去資訊進行預測與評估。

此外，為檢驗模型在不同樣本長度下之穩健性，本文設計短期、中期與長期三種樣本視窗，並皆採用相同的 80%/20% 時間切分方式。圖 6 彙整三種視窗之資料涵蓋區間與對應的訓練／測試切分位置，以利直觀比較各樣本長度下可用資訊量與測試期間差異。

- **短期資料**：2019-09-16 ~ 2020-12-31 (共 319 個交易日；訓練 255 日、測試 64 日)
- **中期資料**：2019-09-16 ~ 2022-12-31 (共 809 個交易日；訓練 647 日、測試 162 日)
- **長期資料**：2019-09-16 ~ 2025-09-12 (共 1460 個交易日；訓練 1168 日、測試 292 日)

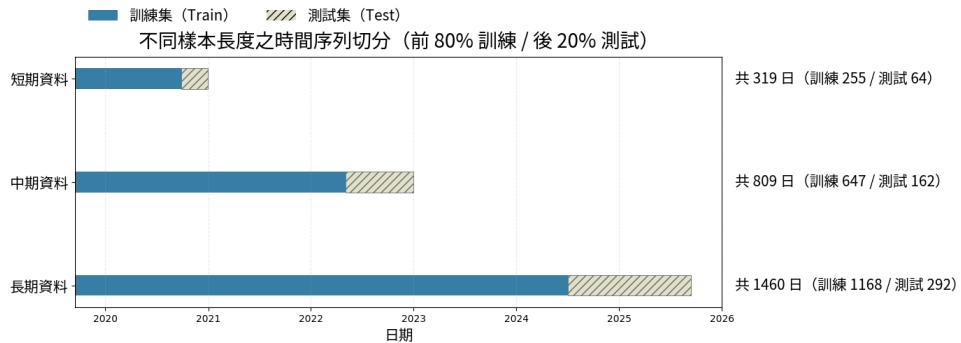


圖 6: 三種樣本視窗之資料劃分 (短期/中期/長期)

3.2 特徵工程與圖結構

3.2.1 特徵選取與分類

模型輸入特徵涵蓋報酬、趨勢、波動、量能、流動性與技術指標等訊號來源如表 2，並以多視窗 (rolling window) 構建不同時間尺度之市場資訊。各特徵之實際使用變數名稱、定義、視窗長度與對應 TEJ 原始欄位已於附錄表 6 完整列示。除 TEJ 匯出欄位外，技術指標部分由 TA-Lib 依價格與量能序列計算，不另引入外部第三方因子資料。

表 2: 模型輸入特徵摘要 (完整 54 項清單見附錄表 6)

類別	代表特徵	定義 (由 TEJ 欄位計算)	視窗
報酬	ret	$P_t/P_{t-k} - 1$ (TEJ 收盤價)	1/5/20
反轉	rev	$-ret_k$ (TEJ 收盤價)	1/5/10
趨勢	px_over_sma	$P_t/SMA_k(P)_t - 1$ (TEJ 收盤價)	5/20/60
動能差	mom_diff	$ret_{k_1} - ret_{k_2}$ (例如 $10 - 1, 20 - 1$; $-$ TEJ 收盤價)	-
波動	std_ret	$Std(r_{t-k+1}, \dots, r_t)$ (TEJ 收盤價)	5/20/60
分配	skew, kurt	20 日報酬偏度/峰度 (TEJ 收盤價)	20
量能	vol_over_ma	$V_t/SMA_k(V)_t - 1$ (TEJ 成交量)	5/20
量能	up_vol, down_vol	上/下漲日平均量 (TEJ 成交量；以日報酬正負分組)	20
流動性	amihud	$\frac{1}{k} \sum \frac{ r }{Amt + \epsilon}$ (TEJ 成交值)	5/20
技術指標	rsi, macd, stoch_k	RSI / MACD / KD (以 TEJ 序列由 TA-Lib 計算)	14/20
風險	atr, mdd	ATR (TEJ 高低收) / 最大回撤 (TEJ 收盤價)	14/20
風險	beta, idio_vol	rolling beta / 特質波動 (以個股與市場報酬 rolling 估計)	60
區間極值	roll_max, roll_min	rolling max/min (TEJ 收盤價)	5/10/20/60
區間位置	pct_pos	$\frac{P_t - \min(P)}{\max(P) - \min(P) + \epsilon}$ (TEJ 收盤價)	5/10/20/60
位置差距	pct_to_high, pct_to_low	距離/低點比例 (以區間 max / min 與 P_t 計；TEJ 收盤價)	20
標準化	zscore_close	$\frac{P_t - \mu_k}{\sigma_k + \epsilon}$ (TEJ 收盤價)	20/60

產業分類 (Industry classification)：產業欄位採用 TEJ 提供之 TEJ 00 - 00。該分類於樣本期間內具相對穩定性，適合用於建構產業層級圖結構；實驗中不額外進行主觀重分群。

3.2.2 圖結構設計與建構流程

產業圖 (Industry Graph)：以「同產業股票完全連結」為原則，將同一產業內之股票視為相互連結，並為每檔股票加入自環 (self-loop)，以捕捉產業內共通訊號與個股自身資訊。圖 7 以三個示例產業為例，展示同產業節點形成高密度子圖 (cluster)，且每個節點皆具有自環結構。

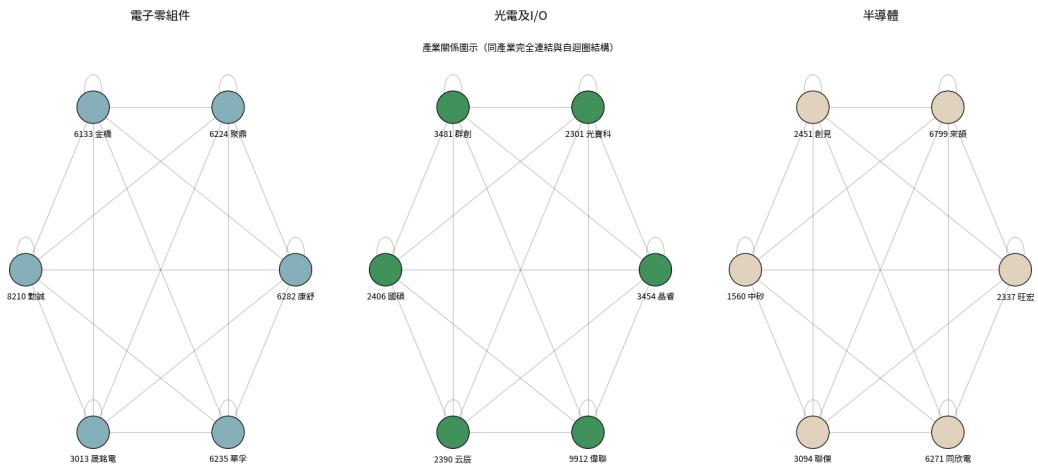


圖 7: 產業圖示意 (Industry Graph)：同產業股票完全連結並加入自環 (self-loop)。圖中每個子圖對應一個產業群集，群集內節點兩兩相連以表示產業內共同訊號傳遞。

全市場圖 (Universe Graph)：全市場圖採完全圖 (complete graph) 設計，所有股票之間皆存在連結，並同樣包含自環。此設計目的在於表徵跨產業之市場共同因子影響，使模型可在截面層級整合全市場訊息。圖 8 顯示完全連結圖結構下，節點之間呈現高度稠密的連線關係。

全市場圖示 (Universe Graph : 完全連結與自迴圈)

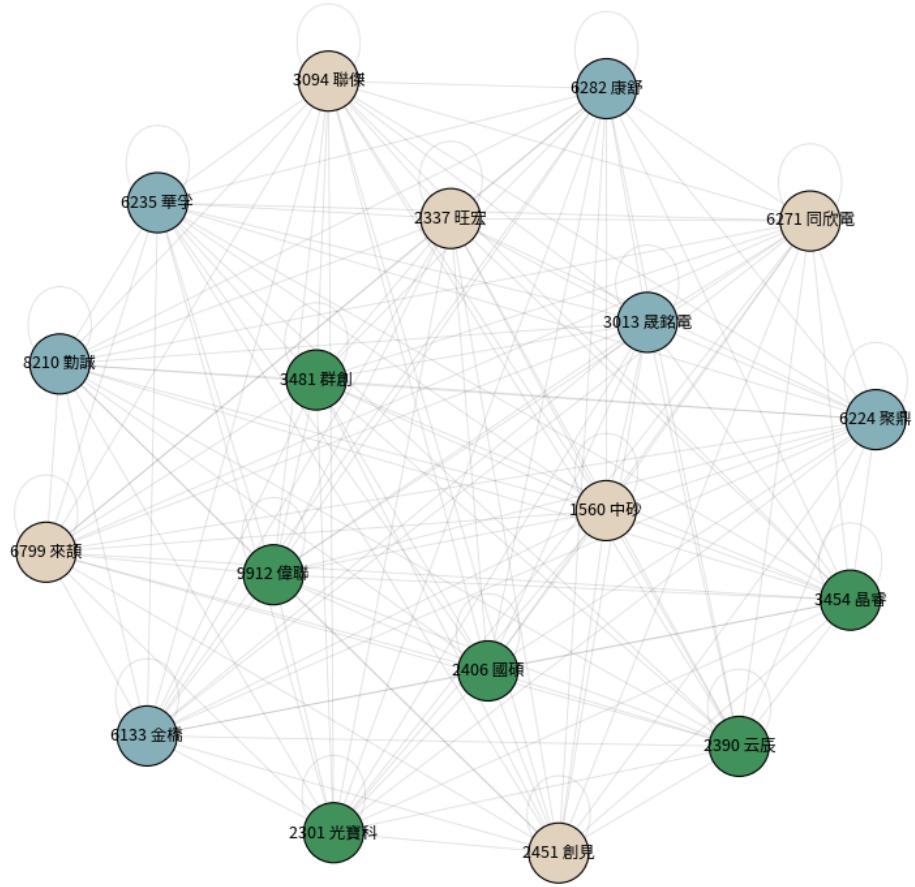


圖 8: 全市場圖示意 (Universe Graph)：完全連結並包含自環 (self-loop)。在此圖結構下，每個節點可與所有其他節點交換訊息，以表徵跨產業的市場共通因子影響。

表示方式(Representation)：圖結構以 PyTorch Geometric (PyG) 之 `edge_index` 表示，其為一個形狀為 $[2, E]$ 的整數張量，其中 E 為邊 (edges) 數量。具體而言，`edge_index` 的每一欄對應一條邊，第一列為起點節點 (source / start node)，第二列為終點節點 (target / end node)。在實作上，本文先依預定規則建立節點間的連結關係（例如同產業完全連結、全市場完全連結與自環），再將邊集合轉換為 `edge_index`，以供 GAT 層進行訊息傳遞與注意力權重計算。

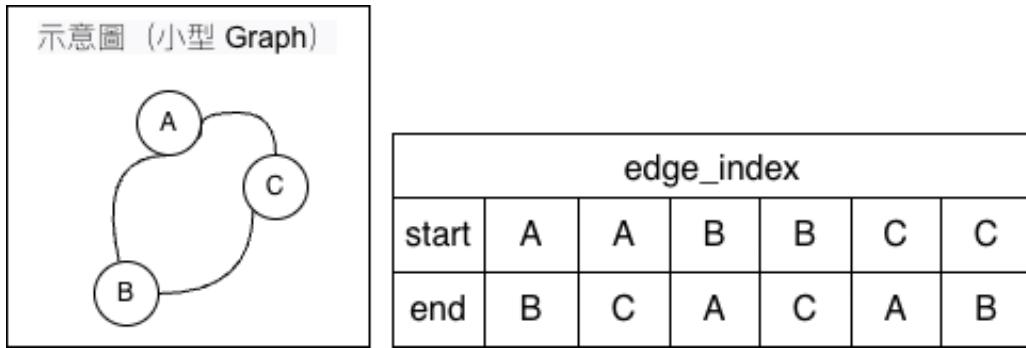


圖 9: edge_index 表示方式示意圖：左側為小型圖結構（節點 A、B、C），右側為其對應之 edge_index。edge_index 以 $[2, E]$ 形式紀錄所有邊，每一欄代表一條由起點（start）指向終點（end）的連結關係。

建構流程（Construction）：由輸入 CSV 讀取產業欄位並依股票代碼分組建立產業圖；若產業欄缺漏，則以單一產業處理。全市場圖依股票清單建立完全連結。圖結構與特徵張量共同保存為 artifacts，以支援訓練與評估之重複使用。

3.3 模型設計

3.3.1 基準模型：GATRegressor

基準模型採用 GATRegressor 作為對照：以兩層 Graph Attention Network (GAT) 於產業圖上進行訊息傳遞，並以線性層輸出預測值。相較於 DMFM，GATRegressor 不包含「產業中性化 + 全市場中性化」之階層式結構，亦不含因子注意力模組，因此可作為較簡化之圖模型基準。超參數設定盡量與 DMFM 對齊（例如 hidden dimension、heads、dropout），以提升可比性。

3.3.2 DMFM 模型架構 (Wei et al. 2022)

DMFM 之核心流程可概括為：「特徵編碼 \rightarrow 產業中性化 \rightarrow 全市場中性化 \rightarrow 階層式拼接 \rightarrow 深度因子學習 \rightarrow 因子注意力」。模型在每個交易日 t 以截面方式處理所有股票之特徵，並輸出每檔股票之預測因子值。

步驟 1：特徵編碼 (Feature encoder)：原始特徵 F^t 經 BatchNorm 後輸入 MLP，映射為 C^t （維度為 `hidden_dim`），作為股票之語境表示。

步驟 2：產業中性化 (Industry neutralization)：以產業圖進行 GAT 運算得到產業影響 H_I^t ，並定義產業中性特徵為 $C_I^t = C^t - H_I^t$ 。

步驟 3：全市場中性化 (Universe neutralization)：以全市場圖在 C_I^t 上進行 GAT 運算得到 H_U^t ，並定義全市場中性特徵為 $C_U^t = C_I^t - H_U^t$ 。

步驟 4：階層式特徵拼接 (Hierarchical concatenation)：拼接三層表示形成 $H^t = [C^t \parallel C_I^t \parallel C_U^t]$ ，以同時保留原始語境與兩層中性化資訊。

步驟 5：深度因子學習 (Deep factor learning)：將 H^t 輸入 decoder MLP，輸出深度因子 f^t （每檔股票一個預測值）。

步驟 6：因子注意力 (Factor attention)：模型同時學習特徵注意力權重 a^t ，形成 $\hat{f}^t = F^t \odot a^t$ ，用以估計驅動因子輸出之重要特徵，並作為監督與解釋性分析依據。

超參數配置 (Hyperparameters)：預設為 `hidden_dim=64`、`heads=2`、`dropout=0.1`、`epochs=200`、`learning rate=1e-4`、`weight decay=0.01`、 $\lambda_{attn} = 0.1$ 、 $\lambda_{IC} = 1.0$ 、`patience=30`

3.4 損失函數

3.4.1 總損失函數

採用 DMFM 對應之損失函數：

$$L = \lambda_{attn} \cdot \|f - \hat{f}\| + \lambda_{IC} \cdot (1 - IC) - \lambda_b \cdot b \quad (3.3)$$

其中 $\|f - \hat{f}\|$ 為注意力估計誤差， IC 為資訊係數 (Information Coefficient, IC)， b 為截面回歸得到的因子收益項。實作中 λ_b 固定為 0.01，以降低因子收益項之不穩定性。

IC (Information Coefficient)：以 Pearson correlation 計算，基於每個交易日之截面預測與真實 forward- k 日報酬，衡量預測排序有效性。

優化器 (Optimizer)：使用 AdamW，並加入 weight decay 以抑制過度擬合。

3.5 評估指標

3.5.1 預測品質指標

IC / Daily IC / ICIR：IC 為整體樣本之 Pearson correlation；Daily IC 為逐交易日 IC 之平均；ICIR (Information Coefficient Information Ratio, ICIR) 定義為 Daily IC 平均除以其標準差，用以衡量穩健性。

產業中性 IC (Industry-neutral IC)：為排除產業共同波動之影響，先在每交易日、每產業內分別對預測與真實報酬去均值 (industry de-meaning)，再對殘差計算 Pearson correlation。令 $\hat{y}_{i,t}$ 與 $y_{i,t}$ 分別為股票 i 在日 t 的預測與真實報酬， $g(i)$ 表示產業別，則

$$\tilde{y}_{i,t} = \hat{y}_{i,t} - \frac{1}{|G_{g(i),t}|} \sum_{j \in G_{g(i),t}} \hat{y}_{j,t}, \quad (3.4)$$

$$\tilde{y}_{i,t} = y_{i,t} - \frac{1}{|G_{g(i),t}|} \sum_{j \in G_{g(i),t}} y_{j,t}, \quad (3.5)$$

其中 $G_{g(i),t}$ 為日 t 時產業 $g(i)$ 的有效股票集合 (排除標籤為 NaN 者)。產業中性 IC 定義為

$$IC_t^{\text{ind-neutral}} = \text{Corr} \left(\tilde{y}_{\cdot,t}, \tilde{y}_{\cdot,t} \right). \quad (3.6)$$

誤差指標 (Error metrics)：以 MSE、RMSE 與 MAE 衡量預測誤差。

3.5.2 投資組合績效指標

年化報酬率 (Annualized return)：以每次再平衡選取預測分數最高之 top_pct 股票等權倂多，計算截面平均報酬作為策略報酬，並依再平衡頻率年化。

Sharpe Ratio：以平均報酬除以波動後再乘年化因子，衡量風險調整後報酬。

勝率 (Hit ratio)：正報酬期間比例，用以衡量策略穩定性。

3.5.3 基準比較與流程摘要

天真基準 (Naive baseline)：令所有股票預測值為 0，計算 MSE/RMSE/MAE 作為誤差下界參考。

台灣 50 ETF 基準：以 0050 之 forward- k 日報酬作為市場基準，與策略累積報酬比較。

評估流程為：載入 artifacts 與模型權重後，分別於訓練期與測試期計算預測與投資組合指標；並輸出 Daily IC、IC 分佈與累積報酬等視覺化結果，以利比較不同模型之表現。對於 DMFM，另輸出注意力權重分佈作為解釋性分析依據。

第四章、實驗結果

本章整理各模型在相同資料切分下的測試結果，從(1)預測能力、(2)投組績效、(3)DMFM 特徵權重三個面向進行比較。

4.1 模型對比

4.1.1 評估指標說明

本研究比較五個模型(Linear、XGBoost、LSTM、GAT、DMFM)在三種時間窗(Short-Medium/Long)之測試表現。指標設計分成兩類：RMSE/MAE 用於衡量點預測誤差，IC/ICIR/DirAcc 用於衡量排序與方向判斷是否可用於選股決策。

- **RMSE** : $\sqrt{\frac{1}{M} \sum_{t,i} (\hat{y}_{i,t} - y_{i,t})^2}$ 。
- **MAE** : $\frac{1}{M} \sum_{t,i} |\hat{y}_{i,t} - y_{i,t}|$ 。
- **IC** : 以每日截面相關係數 $IC_t = \text{corr}(\hat{y}_{:,t}, y_{:,t})$ 計算，再彙整其平均值。
- **ICIR** : $\frac{\text{mean}(IC_t)}{\text{std}(IC_t)}$ ，衡量排序訊號的穩定性。
- **DirAcc** : $\frac{1}{M} \sum_{t,i} \mathbf{1}[\text{sign}(\hat{y}_{i,t}) = \text{sign}(y_{i,t})]$ 。

其中 M 為測試集中「交易日 \times 股票」樣本總數， $\hat{y}_{i,t}$ 與 $y_{i,t}$ 分別為模型預測值與實際報酬。

4.1.2 模型比較結果

表 3: 各模型於不同時間窗之預測指標比較（測試集）

時間窗	模型	RMSE	MAE	IC	ICIR	DirAcc
Short	Linear	0.0555	0.0340	0.0266	0.4313	0.5225
	XGBoost	0.0559	0.0344	0.0644	0.9551	0.5269
	LSTM	0.0661	0.0394	0.0181	0.3267	0.5562
	GAT	0.1206	0.1055	-0.0225	-0.2615	0.4169
	DMFM	0.0563	0.0338	0.0721	0.8228	0.5821
Medium	Linear	0.0502	0.0334	0.0249	0.2192	0.5092
	XGBoost	0.0503	0.0335	0.0688	0.5968	0.5232
	LSTM	0.0564	0.0367	0.0519	0.4525	0.5239
	GAT	0.1293	0.1171	-0.0013	-0.0107	0.4504
	DMFM	0.0627	0.0461	0.0815	0.6762	0.5565
Long	Linear	0.0557	0.0359	0.0414	0.3865	0.5096
	XGBoost	0.0556	0.0358	0.0837	0.7087	0.5271
	LSTM	0.0604	0.0377	0.0344	0.3874	0.5350
	GAT	0.0808	0.0605	-0.0096	-0.1032	0.5266
	DMFM	0.1213	0.0914	0.0834	0.7854	0.5663

表 3 先呈現完整數值，圖 10 進一步以視覺化方式呈現同一組比較結果，便於跨模型與跨時間窗對照。

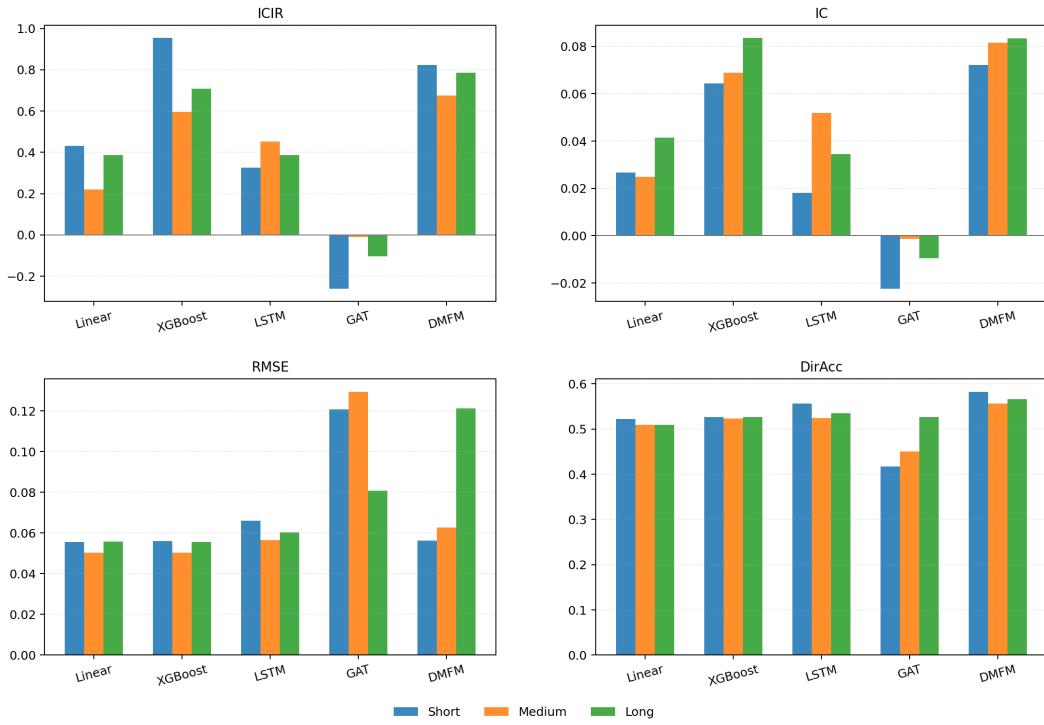


圖 10: 各模型在不同時間窗之關鍵預測指標比較 (ICIR、IC、RMSE、DirAcc)。

由表 3 與圖 10 可得：

- 誤差表現**：點預測誤差仍以 Linear/XGBoost 較低；DMFM 與 GAT 的 RMSE/MAE 較高，顯示排序訊號優勢不必然伴隨最小化誤差。
- 排序能力**：DMFM 在 Short/Medium 的 IC 與 DirAcc 皆為最佳；在 Long 視窗中，DMFM 與 XGBoost 的 IC 接近(0.0834 vs 0.0837)，且 DMFM 的 ICIR 最高(0.7854)。
- 模型定位差異**：DMFM 在排序相關指標與方向準確率上有明顯提升，而 Linear/XGBoost 仍在誤差指標上保有優勢。

其餘預測輸出（含 MSE、Daily IC 與完整模型彙整）見附錄 A.2。

4.2 投組績效驗證

本節比較 GAT、DMFM 與 0050 在不同時間窗的年化報酬、Sharpe 與勝率。回測選股流程為：在每次再平衡日，先依模型預測分數由高至低排序，再選取前 top_pct 股

票等權倂多，持有至下一次再平衡。表 4 整理績效指標；圖 11 對應呈現 Short/Medium/- Long 三種時間窗的累積報酬曲線比較（已收錄於圖目錄）。其中「對齊的天數」表示策略與 0050 在同一時間軸可直接比較的對齊後樣本天數。

表 4: GAT、DMFM 與 0050 於回測區間之績效比較

時間窗	組合	對齊的天數	年化報酬	Sharpe	勝率
Short	0050	64	0.7584	4.201	0.615
	GAT	64	0.4357	7.160	0.769
	DMFM	64	0.9298	8.002	0.846
Medium	0050	162	-0.0890	-0.341	0.424
	GAT	162	0.1158	0.570	0.636
	DMFM	162	0.3734	1.274	0.667
Long	0050	292	-0.0335	-0.113	0.486
	GAT	292	-0.0614	-0.404	0.571
	DMFM	292	0.3131	1.369	0.629

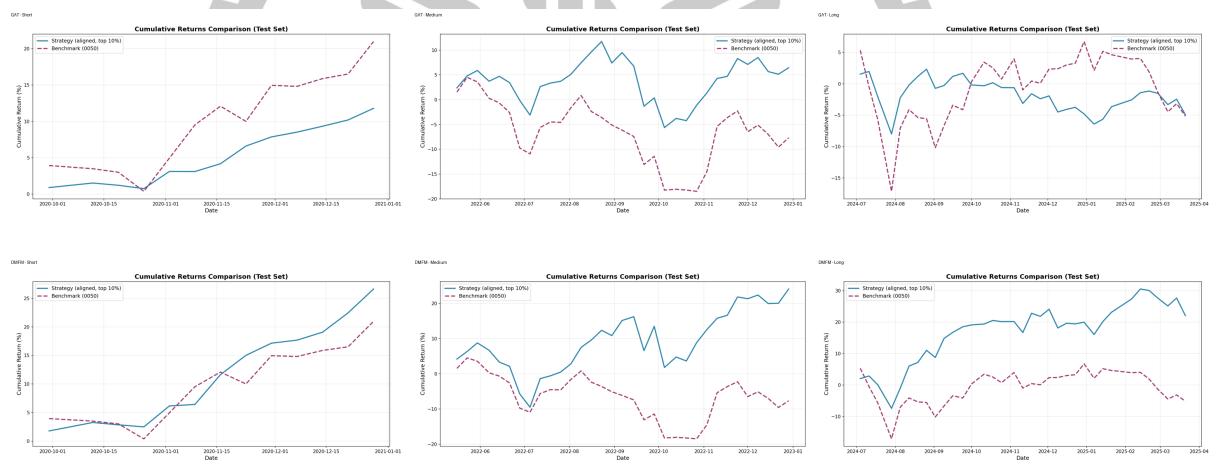


圖 11: GAT 與 DMFM 在 Short/Medium/Long 三種時間窗之累積報酬曲線比較。

由表 4 與圖 11 搭配比較可得：

- **Short :** DMFM 年化 0.9298、Sharpe 8.002、勝率 0.846，優於 GAT 與 0050。
- **Medium :** DMFM (年化 0.3734、Sharpe 1.274) 優於 GAT (0.1158、0.570) 與 0050 (負年化)。

- **Long** : DMFM 仍維持最佳（年化 0.3131、Sharpe 1.369），GAT 在長窗轉為負年化。

完整累積報酬圖（含各時間窗、各模型）見附錄 A.2；其餘回測輸出（Daily IC、IC 分佈、月度命中率等）亦整理於附錄。

4.3 特徵重要性排名

為了解 DMFM 在不同時間窗下偏好的訊號來源，本節依注意力權重統計前五特徵。

表 5 列出各時間窗的 Top-5 結果，圖 12 以圖形方式對照相同資訊。

表 5: DMFM 注意力權重前五特徵（依時間窗）

時間窗	Rank1	Rank2	Rank3	Rank4	Rank5
Short	down_vol_20 (0.0342)	ret_20 (0.0273)	roll_max_60 (0.0256)	amihud_20 (0.0247)	zscore_close_20 (0.0234)
Medium	vol_over_ma_5 (0.0287)	roll_min_20 (0.0275)	macd_signal (0.0251)	roll_max_5 (0.0250)	pct_to_high_20 (0.0229)
Long	pct_to_low_20 (0.0325)	roll_max_5 (0.0258)	roll_min_60 (0.0242)	macd_hist (0.0231)	amihud_5 (0.0228)

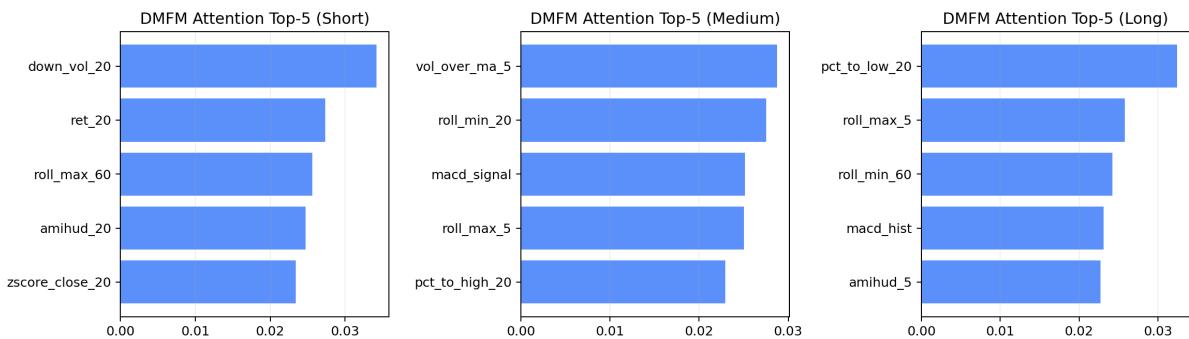


圖 12: DMFM 在不同時間窗的注意力前五特徵權重。

由表 5 與圖 12 可得兩項重點：

- 區間位置與量價特徵在多個時間窗持續出現（如 `roll_max_5`、`pct_to_low_20`、`down_vol_20`），顯示 DMFM 偏好結合價格區間與量能訊號。
- 流動性與動能訊號仍具影響力（如 `amihud_20`、`amihud_5`、`ret_20`、`macd_signal`），反映模型在不同時間窗會切換關注的主導因子。

更多注意力與輔助輸出亦整理於附錄 A.2。

第五章、結論

5.1 未來研究

未來研究將聚焦於下列三項工作：

- **補齊 benchmark model 的投組績效比較：**除既有預測指標外，將 Linear、XG-Boost、LSTM 也納入投組回測，與 GAT、DMFM 以同一口徑輸出年化報酬、Sharpe、勝率與累積報酬，提供完整數值比較結果。
- **增加資料筆數：**擴充訓練與測試樣本區間，提升模型估計穩健性，並降低單一期間市場狀態對結論的影響。
- **加入更多特徵：**在既有特徵基礎上增加更多不同型態訊號（如基本面、技術面、流動性與風險特徵），並透過特徵篩選與消融分析評估其對預測與投組績效的增益。

参考文献

- [1] E. F. Fama and K. R. French, “Common risk factors in the returns on stocks and bonds,” *Journal of Financial Economics*, vol. 33, no. 1, pp. 3–56, 1993. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/0304405X93900235>
- [2] N. Jegadeesh and S. Titman, “Returns to buying winners and selling losers: Implications for stock market efficiency,” *The Journal of Finance*, vol. 48, no. 1, pp. 65–91, 1993.
- [3] S. Gu, B. Kelly, and D. Xiu, “Empirical asset pricing via machine learning,” *The Review of Financial Studies*, vol. 33, no. 5, pp. 2223–2273, 2020.
- [4] C. Fieberg, A. Metko, T. Poddig, and T. Loy, “Machine learning techniques for cross-sectional equity returns’ prediction,” *OR Spectrum*, 2023.
- [5] S. Basak, S. Kar, S. Saha, and L. Khadem, “A comparison of tree-based ensemble methods for stock price direction prediction,” *Econometrics and Finance*, 2019, tree-based ensembles for stock direction prediction.
- [6] E. K. Ampomah, G. Nyame, F. Tawiah, and M. Baah, “Stock market prediction using ensemble machine learning,” *International Journal of Scientific & Technology Research*, 2020, ensemble models for stock market prediction.
- [7] I. K. Nti, A. F. Adekoya, and B. A. Weyori, “A comprehensive evaluation of ensemble learning for stock-market prediction,” *Journal of Big Data*, vol. 7, no. 1, pp. 1–40, 2020.

- [8] A. A. Saifan, M. Alshraideh, and M. Almasri, “Investigating algorithmic stock market trading using ensemble methods,” *IEEE Access*, vol. 8, pp. 123 042–123 055, 2020.
- [9] V. Vidmanát, J. Bartůňka, and L. Marek, “Stock market volatility forecasting using information-theoretic measures and machine learning,” *Journal of Forecasting*, vol. 37, no. 3, pp. 281–294, 2018.
- [10] T. Chen and C. Guestrin, “Xgboost: A scalable tree boosting system,” in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2016, pp. 785–794.
- [11] M.-C. Wu, S.-Y. Lin, and C.-H. Lin, “An effective application of decision tree to stock trading,” *Expert Systems with Applications*, vol. 31, no. 2, pp. 270–274, 2006. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0957417405002137>
- [12] t. b. v. Tang, “Machine learning in financial time series forecasting (survey/notes),” PDF manuscript / survey notes, 2022, please verify bibliographic metadata if replacing with a published version.
- [13] W. A. Brock, J. Lakonishok, and B. LeBaron, “Simple technical trading rules and the stochastic properties of stock returns,” *The Journal of Finance*, vol. 47, no. 5, pp. 1731–1764, 1992. [Online]. Available: <https://api.semanticscholar.org/CorpusID:8539335>
- [14] Q. Lin, “Technical analysis and stock return predictability: An aligned approach,” *Journal of Financial Markets*, vol. 38, pp. 103–123, 2018. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S1386418117300824>
- [15] A. W. Li and G. S. Bastos, “Stock market forecasting using deep learning and technical analysis: A systematic review,” *IEEE Access*, vol. 8, pp. 185 197–185 215, 2020.

- [16] R. Cervelló-Royo, F. Guijarro, and K. Michniuk, “Stock market trading rule based on pattern recognition and technical analysis: Forecasting the djia index with intraday data,” *Expert Systems with Applications*, vol. 42, no. 14, pp. 5963–5975, 2015. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0957417415002110>
- [17] M. Agrawal, P. K. Shukla, R. Nair, A. Nayyar, and M. Masud, “Stock prediction based on technical indicators using deep learning model,” *Computers, Materials & Continua*, vol. 70, no. 1, pp. 287–304, 2022.
- [18] M. Agrawal, A. U. Khan, and P. K. Shukla, “Stock price prediction using technical indicators: A predictive model using optimal deep learning,” *International Journal of Recent Technology and Engineering (IJRTE)*, vol. 8, no. 2, pp. 2297–2305, 2019.
- [19] O. M. Sezer, U. Gudelek, and A. M. Ozbayoglu, “Financial time series forecasting with deep learning: A systematic literature review: 2005–2019,” *arXiv preprint*, 2019.
- [20] S. Zaheer, N. Anjum, S. Hussain, A. D. Algarni, J. Iqbal, S. Bourouis, and S. S. Ullah, “A multi parameter forecasting for stock time series data using lstm and deep learning model,” *Mathematics*, vol. 11, no. 3, p. 590, 2023.
- [21] J. Qiu, B. Wang, and C. Zhou, “Forecasting stock prices with long-short term memory neural network based on attention mechanism,” *PLoS ONE*, vol. 15, no. 1, p. e0227222, 2020.
- [22] 鄭邦廷, “基於深度學習與技術分析指標預測股市買賣點,” Master’s thesis, 國立臺灣師範大學機電工程學系, 台北, 台灣, 2023, stock buy and sell points prediction based on deep learning and technical analysis indicators.

- [23] 廖俊翔, “應用深度學習結合自相關分析優化股票預測模型,” Master’s thesis, 國立中興大學資訊管理學系, 台中, 台灣, 2022, applying deep learning combined with autocorrelation analysis to optimize stock forecasting model.
- [24] R. Sonkavde, I. Ahmed, A. Sonkavde, S. Mittal, S. Jha, S. Sharma, and H. Y. Kim, “A review of machine learning and deep learning applications for stock market prediction,” *Expert Systems with Applications*, vol. 213, p. 119043, 2023.
- [25] F. Scarselli, M. Gori, A. C. Tsoi, M. Hagenbuchner, and G. Monfardini, “The graph neural network model,” *IEEE Transactions on Neural Networks*, vol. 20, no. 1, pp. 61–80, 2009.
- [26] Z. Wu, S. Pan, F. Chen, G. Long, C. Zhang, and P. S. Yu, “A comprehensive survey on graph neural networks,” *IEEE Transactions on Neural Networks and Learning Systems*, vol. 31, no. 1, pp. 4–24, 2020.
- [27] R. Feng *et al.*, “Spatio-temporal graph attention neural network for stock price prediction,” *Applied Sciences*, 2025, sTGAT variant for stock prediction.
- [28] t. b. v. Wei *et al.*, “Fstgat: A multi-factor spatio-temporal graph attention network for stock market prediction,” *Expert Systems with Applications*, 2025, fSTGAT variant for stock prediction (please verify final metadata).
- [29] A. Li *et al.*, “Structure over signal: A graph neural network approach for robust stock prediction,” arXiv preprint, 2025, graph-based robustness discussion (please confirm final bibliographic info).
- [30] G. Song, T. Zhao, S. Wang, H. Wang, and X. Li, “Stock ranking prediction using a graph aggregation network based on stock price and stock relationship information,” *Information Sciences*, vol. 643, p. 119236, 2023.

- [31] t. b. v. Xiang, "Dynamic attributes-driven graph attention network modeling on behavioral finance for stock prediction," Preprint / manuscript, 2022, please verify authors/journal metadata if you decide to keep this citation key.
- [32] t. b. v. Liu *et al.*, "Heterogeneous dual-dynamic attention network for stock prediction," *IEEE Access*, 2024, heterogeneous graph attention modeling for stock prediction.
- [33] Z. Ding, S. Martin, and J. Yang, "The statistics of information coefficients," *SSRN Electronic Journal*, 2011, iC distribution and related statistical properties.
- [34] ———, "Statistics of time-varying cross-sectional information coefficients," *SSRN Electronic Journal*, 2022, time-varying IC and volatility implications.
- [35] E. E. Qian and R. Hua, "Active risk and the information ratio," *The Journal of Investment Management*, 2004, fundamental law style discussion: IC, breadth, and IR.
- [36] J. Ye, "How variation in signal quality affects performance," *Financial Analysts Journal*, vol. 64, no. 4, pp. 48–61, 2008.
- [37] t. b. v. Zhang, "Turnover-adjusted information ratio and practical performance evaluation," Working paper / manuscript, 2021, turnover-adjusted IR discussion (please verify final bibliographic info).
- [38] P. Bermejo, I. Figuerola-Ferretti, C. Hevia, and M. S. Santos, "Improving factor investing in the european equity market," *Heliyon*, vol. 7, no. 9, p. e08168, 2021.
- [39] t. b. v. Yuan, "Machine learning-driven factor investing: Enhancing stock return prediction with xgboost," *Transactions on Economics, Business and Management Research*, 2024, please verify publisher and DOI if needed.

Appendix 附錄

本附錄彙整本研究模型輸入特徵與完整實驗輸出補充。

A.1 模型輸入特徵完整清單

表 6: 本研究實際使用之模型輸入特徵 (feature_cols, 共 54 項；全部由 TEJ 原始欄位衍生)

特徵名稱	類別	定義 (由 TEJ 欄位計算)	視窗
ret_1	報酬	$P_t/P_{t-1} - 1$, 其中 P_t 為 TEJ 收盤價 (元)	1
ret_3	報酬	$P_t/P_{t-3} - 1$ (TEJ 收盤價)	3
ret_5	報酬	$P_t/P_{t-5} - 1$ (TEJ 收盤價)	5
ret_10	報酬	$P_t/P_{t-10} - 1$ (TEJ 收盤價)	10
ret_20	報酬	$P_t/P_{t-20} - 1$ (TEJ 收盤價)	20
rev_1	反轉	-ret_1 (由 TEJ 收盤價衍生)	1
rev_5	反轉	-ret_5 (由 TEJ 收盤價衍生)	5
rev_10	反轉	-ret_10 (由 TEJ 收盤價衍生)	10
px_over_sma_5	趨勢	$P_t/\text{SMA}_5(P)_t - 1$ (TEJ 收盤價)	5
px_over_sma_10	趨勢	$P_t/\text{SMA}_{10}(P)_t - 1$ (TEJ 收盤價)	10

(續下頁)

特徵名稱	類別	定義（由 TEJ 欄位計算）	視窗
px_over_sma_20	趨勢	$P_t/\text{SMA}_{20}(P)_t - 1$ (TEJ 收盤價)	20
px_over_sma_60	趨勢	$P_t/\text{SMA}_{60}(P)_t - 1$ (TEJ 收盤價)	60
mom_diff_10	動能差	$\text{ret}_{10} - \text{ret}_1$ (由 TEJ 收盤價衍生)	-
mom_diff_20	動能差	$\text{ret}_{20} - \text{ret}_1$ (由 TEJ 收盤價衍生)	-
std_ret_5	波動	$\text{Std}(r_{t-4}, \dots, r_t)$, r_t 由 TEJ 收盤價計算	5
std_ret_10	波動	$\text{Std}(r_{t-9}, \dots, r_t)$ (TEJ 收盤價)	10
std_ret_20	波動	$\text{Std}(r_{t-19}, \dots, r_t)$ (TEJ 收盤價)	20
std_ret_60	波動	$\text{Std}(r_{t-59}, \dots, r_t)$ (TEJ 收盤價)	60
skew_20	分配	20 日報酬偏度 (由 TEJ 收盤價計算之 r_t)	20
kurt_20	分配	20 日報酬峰度 (由 TEJ 收盤價計算之 r_t)	20
vol_over_ma_5	量能	$V_t/\text{SMA}_5(V)_t - 1$, V_t 為 TEJ 成交量 (千股)	5
vol_over_ma_10	量能	$V_t/\text{SMA}_{10}(V)_t - 1$ (TEJ 成交量)	10
vol_over_ma_20	量能	$V_t/\text{SMA}_{20}(V)_t - 1$ (TEJ 成交量)	20
vol_over_ma_60	量能	$V_t/\text{SMA}_{60}(V)_t - 1$ (TEJ 成交量)	60
up_vol_20	量能	20 日內上漲日平均量 : $\frac{\sum V \mathbb{I}(r>0)}{\sum \mathbb{I}(r>0)+\varepsilon}$ (TEJ 成交量、TEJ 收盤價計算 r)	20
down_vol_20	量能	20 日內下跌日平均量 : $\frac{\sum V \mathbb{I}(r<0)}{\sum \mathbb{I}(r<0)+\varepsilon}$ (TEJ 成交量、TEJ 收盤價計算 r)	20

(續下頁)

特徵名稱	類別	定義（由 TEJ 欄位計算）	視窗
amihud_5	流動性	Amihud illiquidity : $\frac{1}{5} \sum \frac{ r }{\text{Amt} + \varepsilon}$, Amt 為 5 TEJ 成交值(千元)	5
amihud_20	流動性	同上 (TEJ 成交值、TEJ 收盤價計算 r)	20
rsi_14	技術指標	RSI (TA-Lib 預設參數；由 TEJ 收盤價計算)	14
stoch_k_14	技術指標	Stochastic %K (TA-Lib 預設；由 TEJ 最高/最低/收盤價計算)	14
stoch_d_3	技術指標	Stochastic %D (%K 的平滑；TA-Lib 預設)	3
macd	技術指標	MACD 主線 (TA-Lib 預設；由 TEJ 收盤價計算)	-
macd_signal	技術指標	MACD 訊號線 (TA-Lib 預設)	-
macd_hist	技術指標	MACD 柱狀差 ($\text{macd}-\text{macd_signal}$)	-
atr_14	風險	ATR (Average True Range；由 TEJ 最高/最低/收盤價計算)	14
mdd_20	風險	20 日最大回撤 : $\max_{\tau \leq t} \left(\frac{\max P - P}{\max P} \right)$ (TEJ 收盤價)	20

(續下頁)

特徵名稱	類別	定義（由 TEJ 欄位計算）	視窗
beta_60	風險	60 日 rolling beta (由 TEJ 資料計算) : 60 以個股日報酬 $r_{i,t}$ 對市場日報酬 $r_{m,t}$ 做 OLS, $\beta = \frac{\text{Cov}(r_i, r_m)}{\text{Var}(r_m)}$ 。本研究之市場報 酬 $r_{m,t}$ 使用 TEJ 匯出之 0050 (證券代碼 0050) 收盤價計算日報酬。	
idio_vol_60	風險	60 日特質波動：上述 rolling 回歸殘差 $\varepsilon_{i,t}$ 60 的標準差 $\text{Std}(\varepsilon)$ 。回歸所用市場報酬同 beta_60，亦由 TEJ 匯出之 0050 收盤價 計算。	
roll_max_5	區間極值	5 日 rolling maximum : $\max(P_{t-4}, \dots, P_t)$ 5 (TEJ 收盤價)	
roll_min_5	區間極值	5 日 rolling minimum : $\min(P_{t-4}, \dots, P_t)$ 5 (TEJ 收盤價)	
pct_pos_5	區間位置	$\frac{P_t - \text{roll_min_5}}{\text{roll_max_5} - \text{roll_min_5} + \varepsilon}$ (TEJ 收盤價)	5
roll_max_10	區間極值	10 日 rolling maximum (TEJ 收盤價)	10
roll_min_10	區間極值	10 日 rolling minimum (TEJ 收盤價)	10
pct_pos_10	區間位置	$\frac{P_t - \text{roll_min_10}}{\text{roll_max_10} - \text{roll_min_10} + \varepsilon}$ (TEJ 收盤價)	10
roll_max_20	區間極值	20 日 rolling maximum (TEJ 收盤價)	20
roll_min_20	區間極值	20 日 rolling minimum (TEJ 收盤價)	20
pct_pos_20	區間位置	$\frac{P_t - \text{roll_min_20}}{\text{roll_max_20} - \text{roll_min_20} + \varepsilon}$ (TEJ 收盤價)	20

(續下頁)

特徵名稱	類別	定義（由 TEJ 欄位計算）	視窗
roll_max_60	區間極值	60 日 rolling maximum (TEJ 收盤價)	60
roll_min_60	區間極值	60 日 rolling minimum (TEJ 收盤價)	60
pct_pos_60	區間位置	$\frac{P_t - \text{roll_min_60}}{\text{roll_max_60} - \text{roll_min_60} + \varepsilon}$ (TEJ 收盤價)	60
pct_to_high_20	位置差距	距 20 日高點比例 : $\frac{P_t}{\text{roll_max_20}} - 1$ (TEJ 收盤價)	20
pct_to_low_20	位置差距	距 20 日低點比例 : $\frac{P_t}{\text{roll_min_20}} - 1$ (TEJ 收盤價)	20
zscore_close_20	標準化	20 日 z-score : $\frac{P_t - \mu_{20}}{\sigma_{20} + \varepsilon}$ (TEJ 收盤價 ; μ, σ 為 rolling mean/std)	20
zscore_close_60	標準化	60 日 z-score : $\frac{P_t - \mu_{60}}{\sigma_{60} + \varepsilon}$ (TEJ 收盤價)	60

A.2 完整實驗輸出補充

A.2.1 各模型於不同時間窗之完整預測輸出補充（測試集）

表 7: 各模型於不同時間窗之完整預測輸出補充（測試集）

時間窗	模型	樣本數（筆）	MSE	Daily IC
Short	Linear	49280	0.003078	0.0281
	XGBoost	49280	0.003130	0.0646
	LSTM	49280	0.004373	0.0211
	GAT	49280	0.014555	-0.0175
	DMFM	49280	0.003168	0.0717
Medium	Linear	121734	0.002522	0.0225
	XGBoost	121734	0.002533	0.0688
	LSTM	121734	0.003179	0.0488
	GAT	121734	0.016715	-0.0007
	DMFM	121734	0.003930	0.0826
Long	Linear	221277	0.003101	0.0363
	XGBoost	221277	0.003094	0.0804
	LSTM	221277	0.003650	0.0323
	GAT	221277	0.006532	-0.0085
	DMFM	221277	0.014716	0.0823

A.2.2 完整累積報酬圖

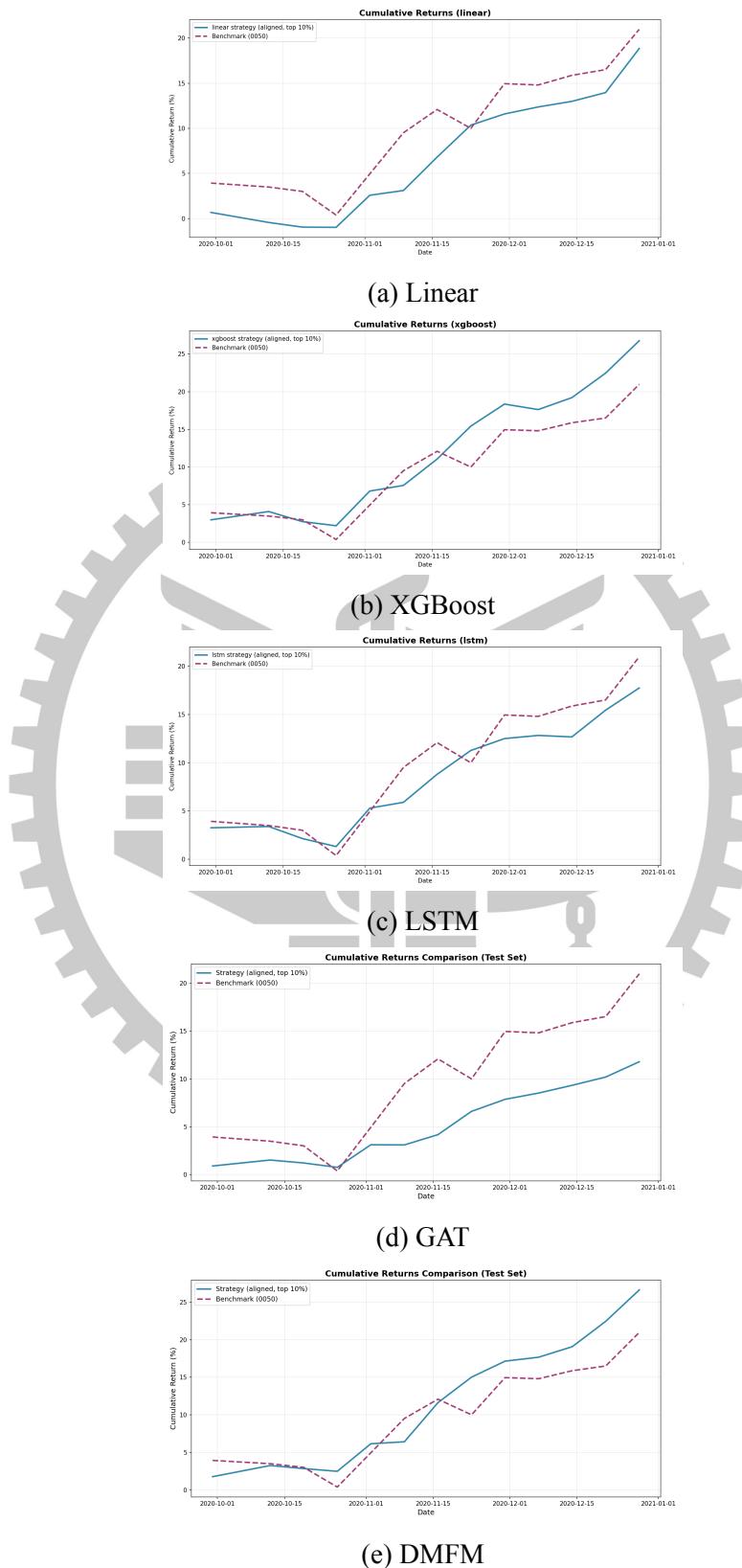


圖 13: short 時間窗各模型累積報酬曲線。

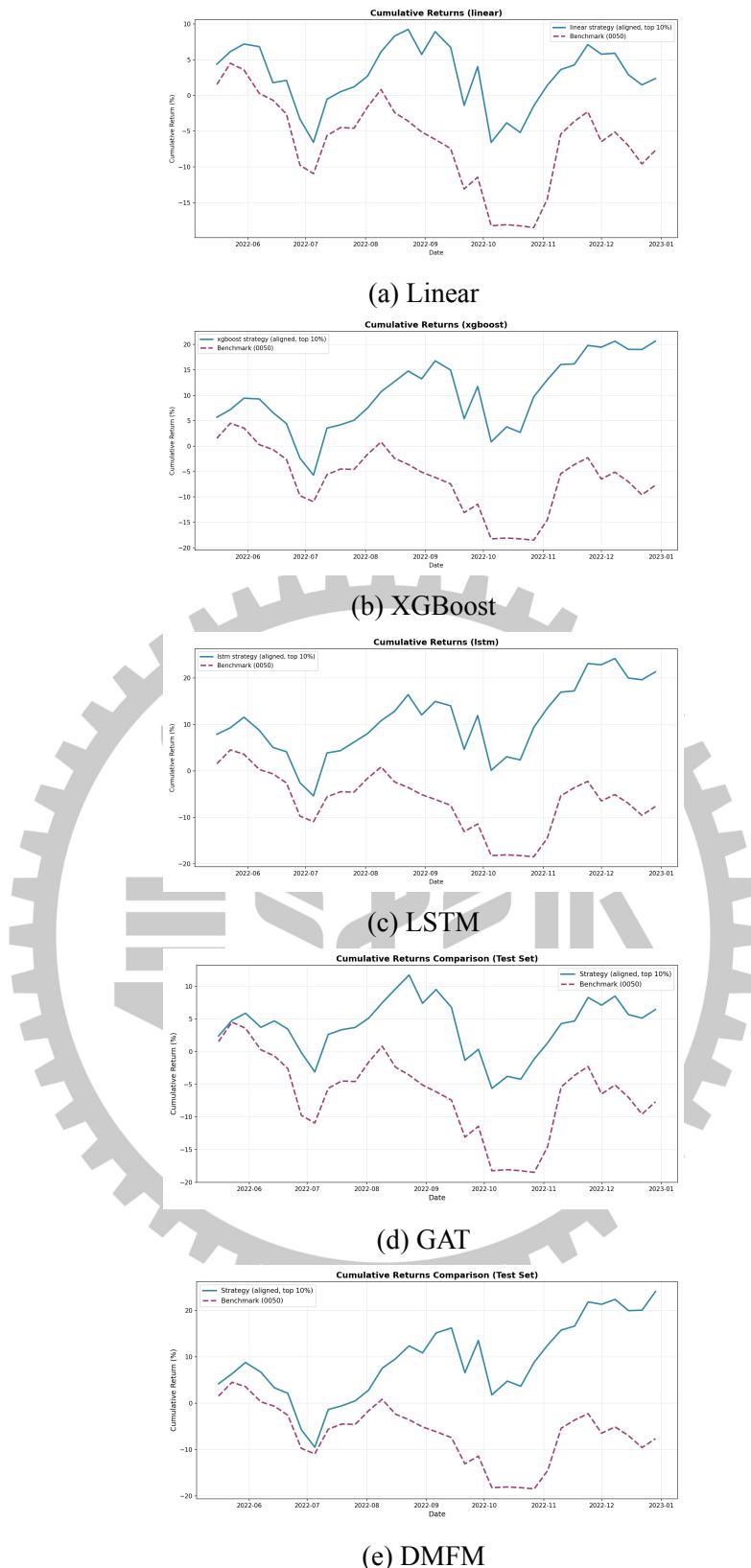


圖 14: medium 時間窗各模型累積報酬曲線。

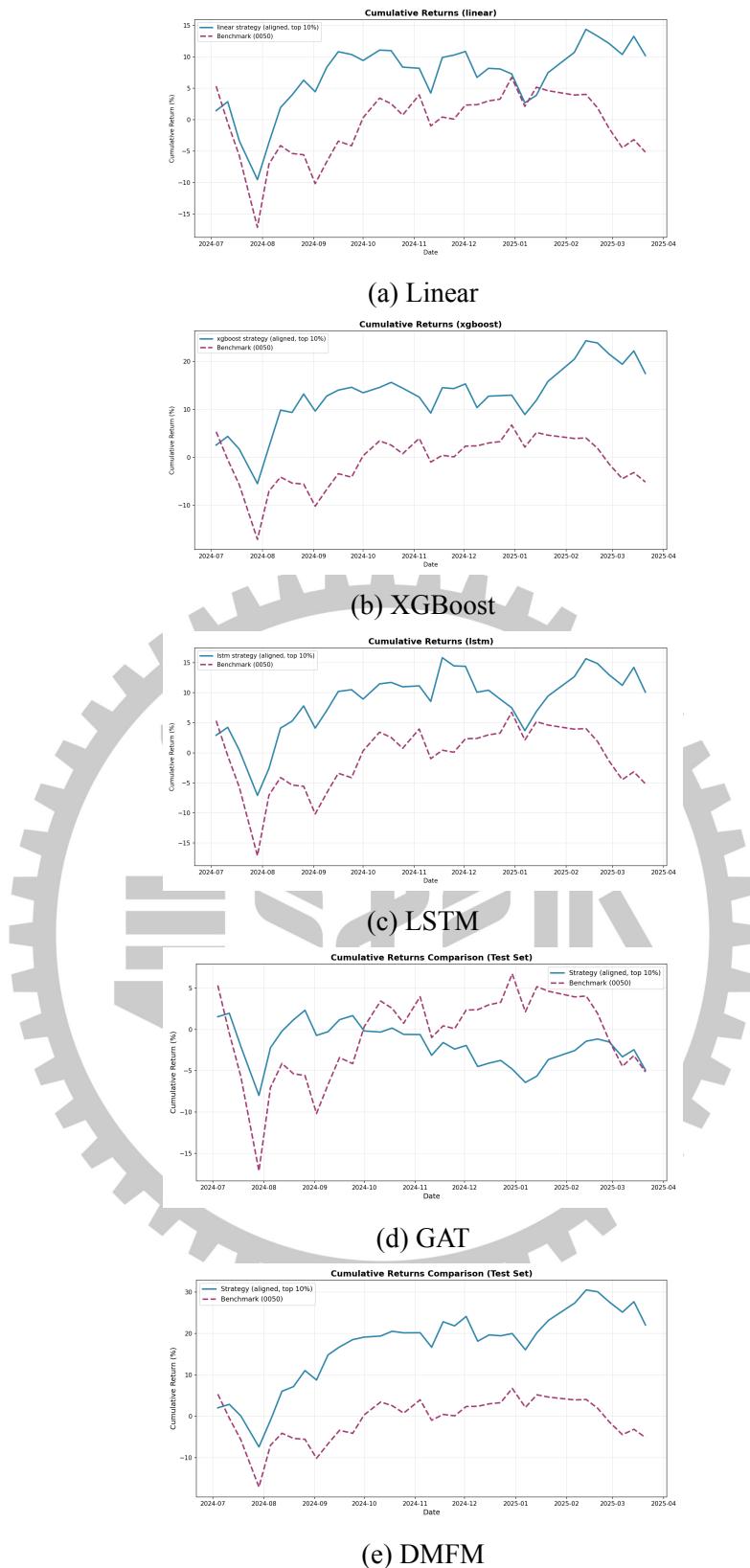


圖 15: long 時間窗各模型累積報酬曲線。

A.2.3 Daily IC 補充結果

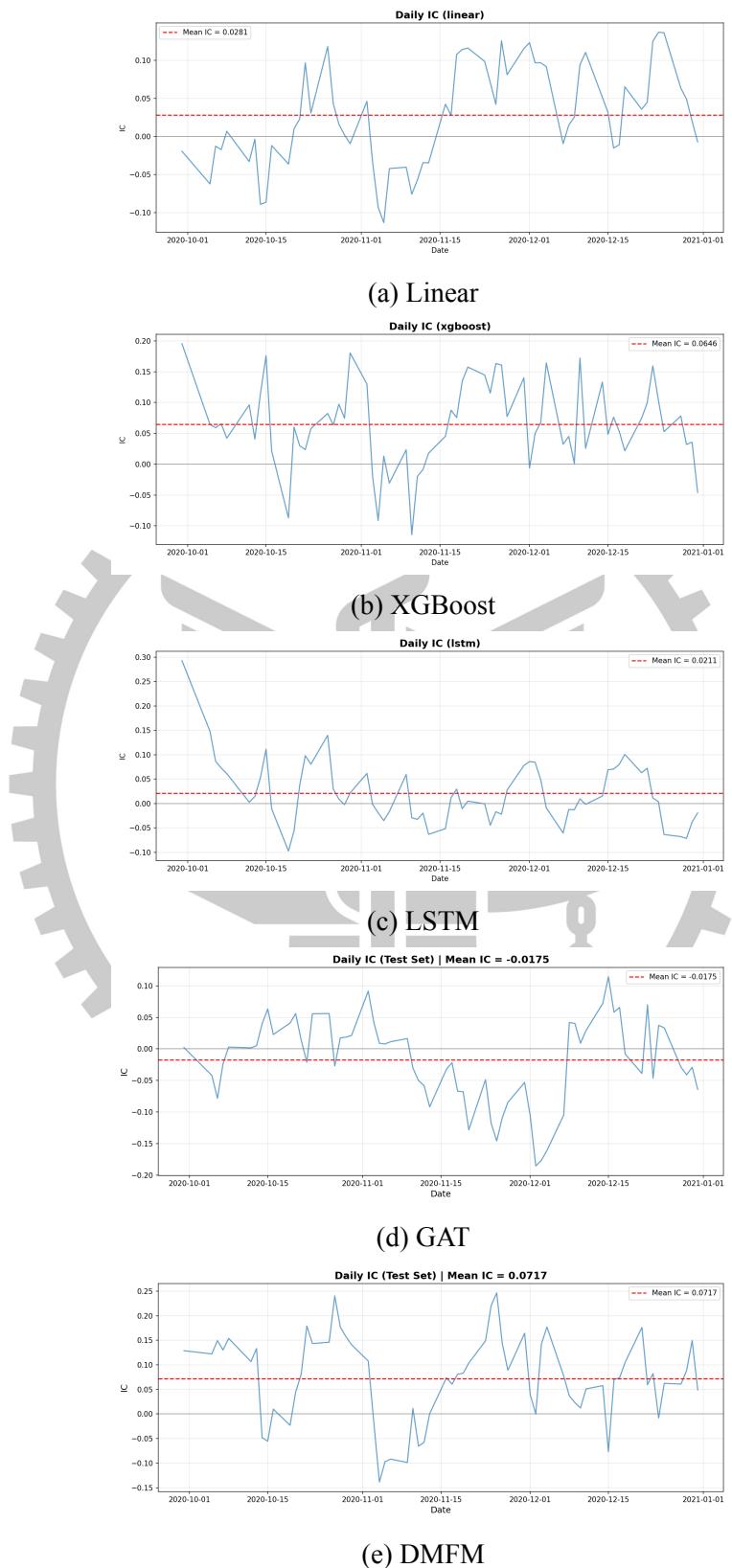


圖 16: short 時間窗各模型 Daily IC。

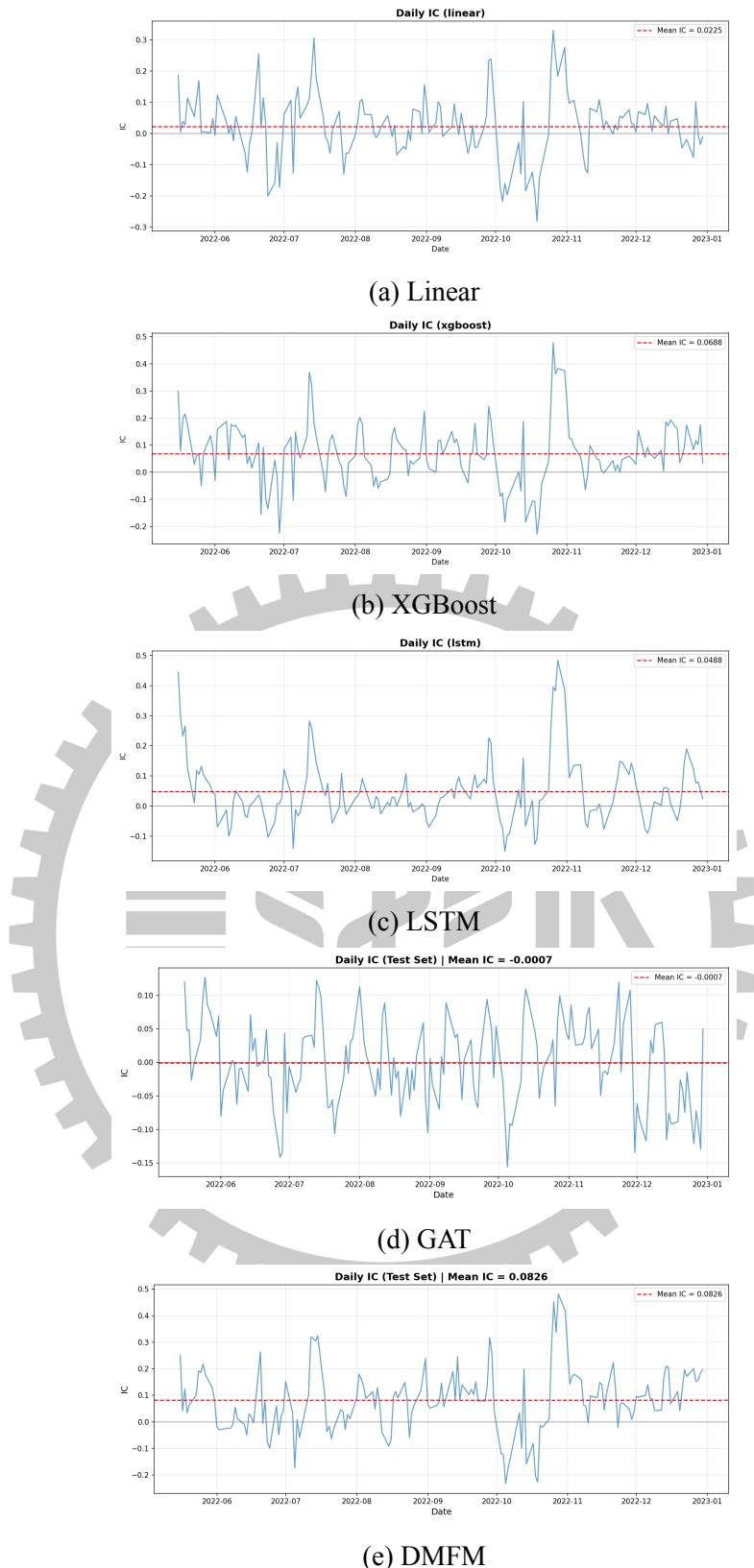


圖 17: medium 時間窗各模型 Daily IC。



圖 18: long 時間窗各模型 Daily IC。

A.2.4 IC 分佈補充結果

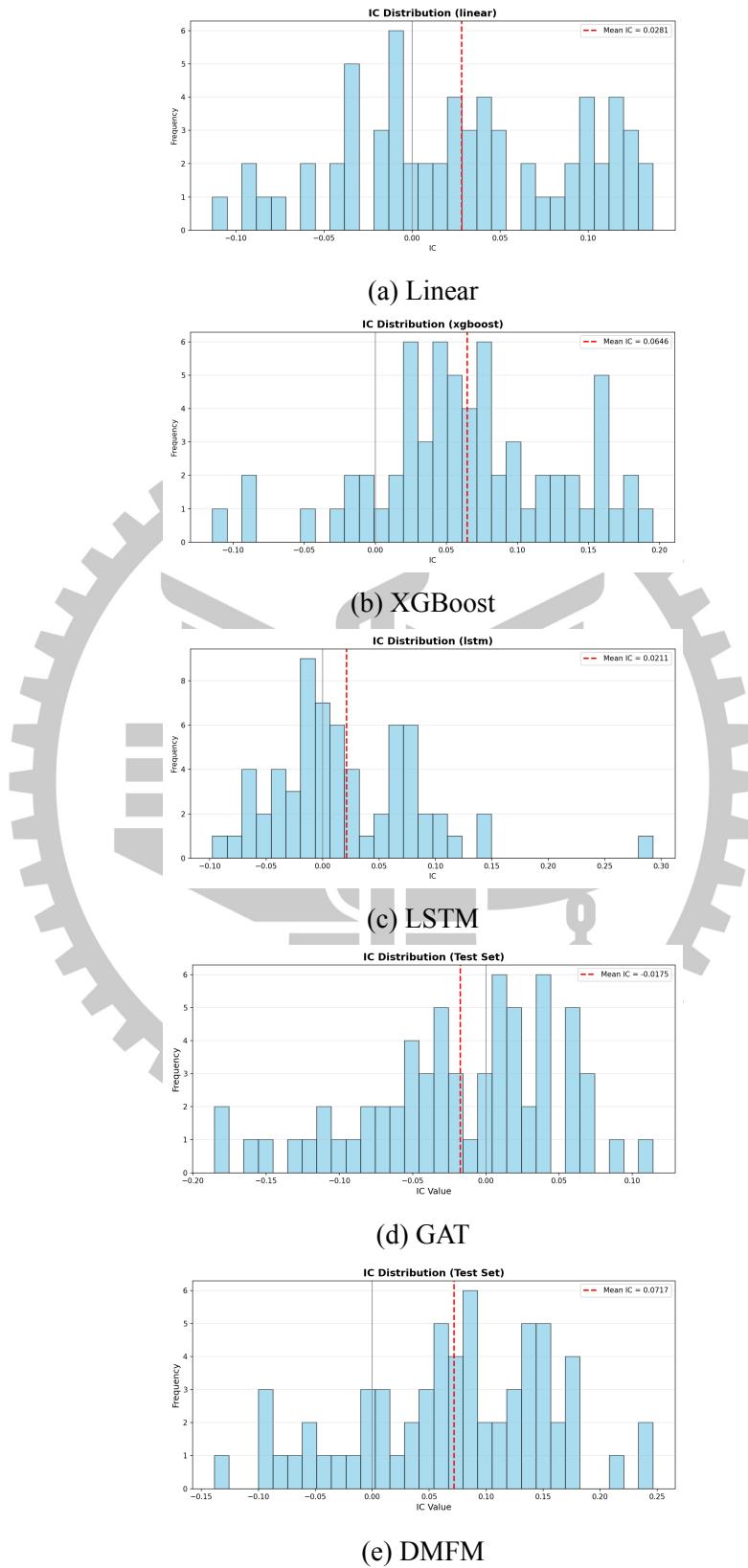


圖 19: short 時間窗各模型 IC 分佈。

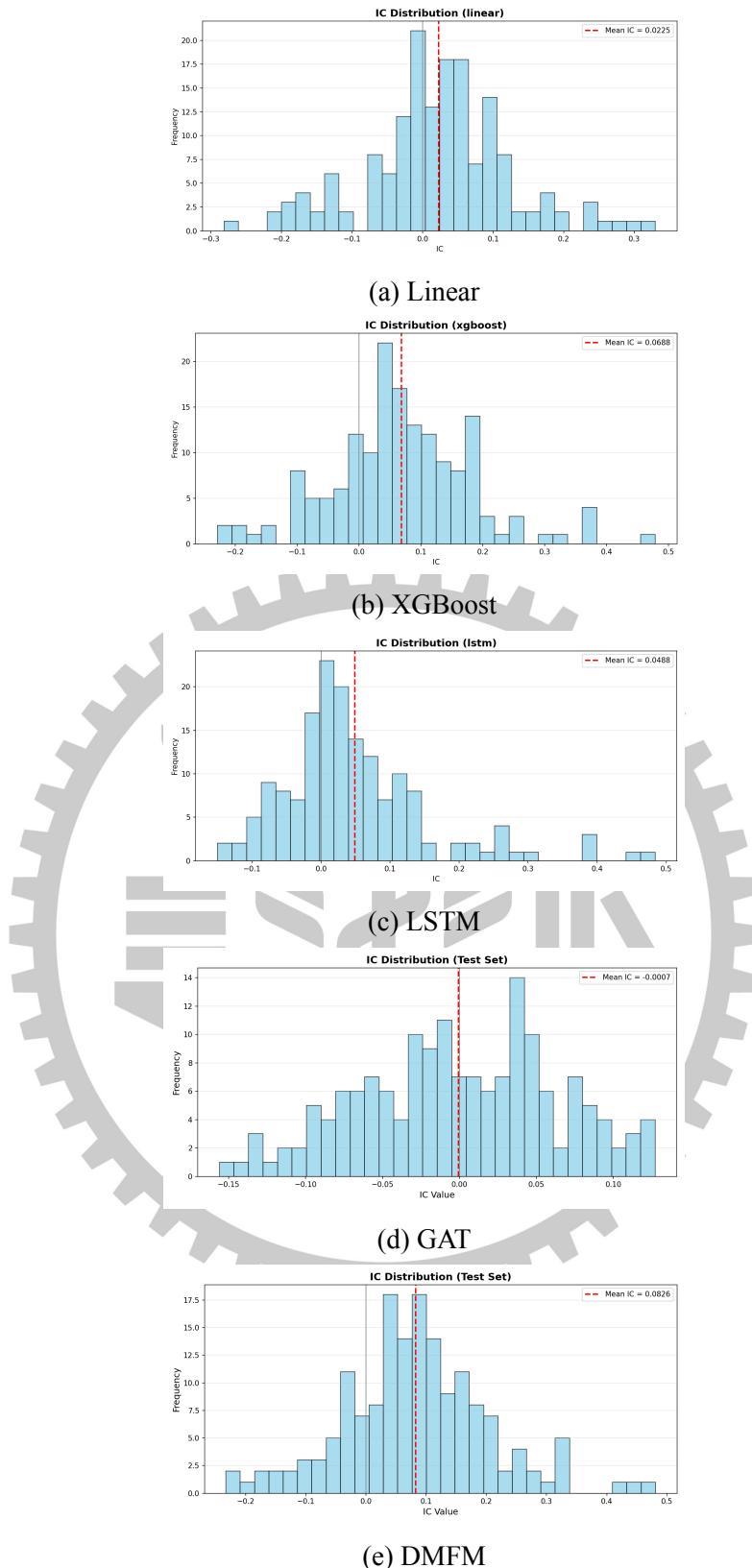


圖 20: medium 時間窗各模型 IC 分佈。

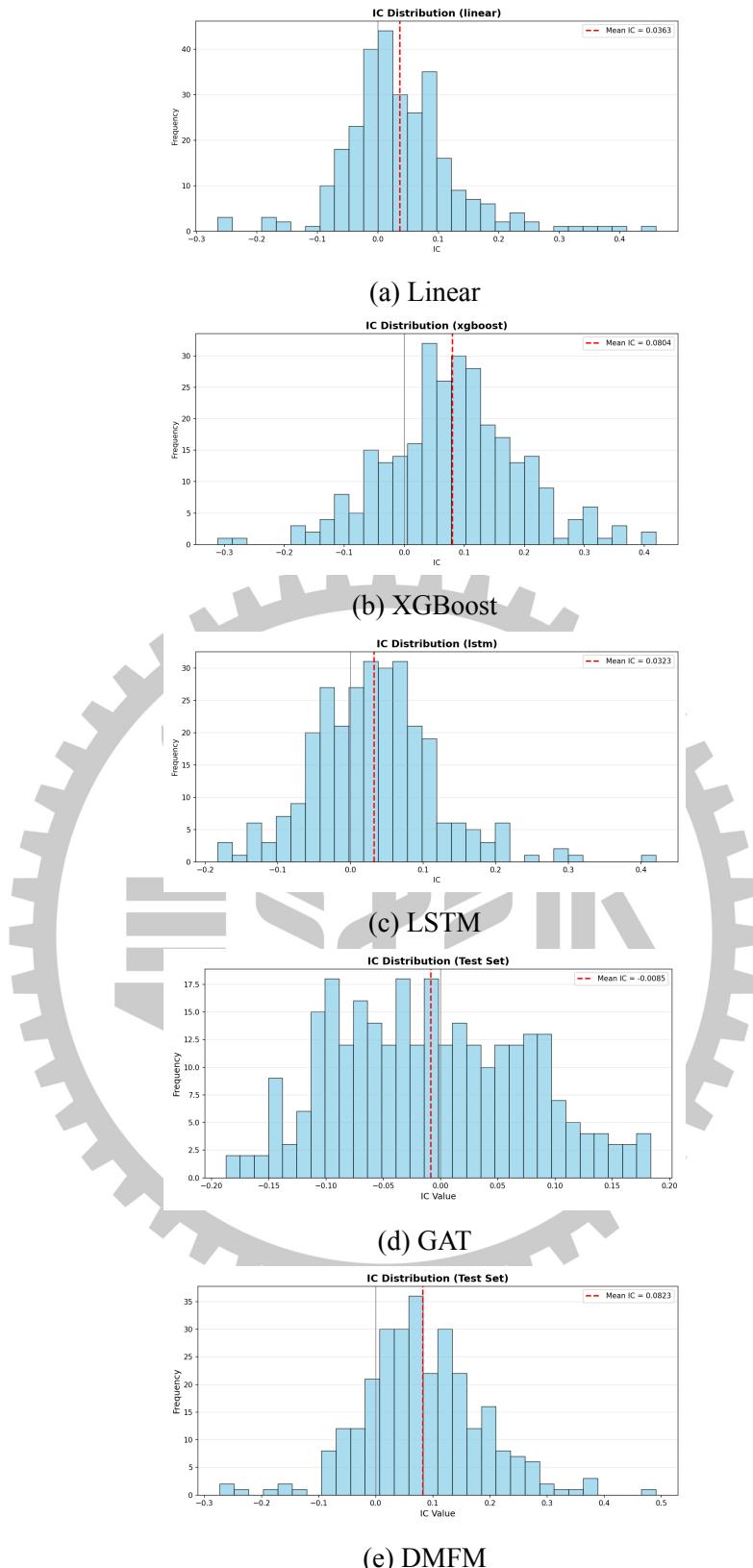


圖 21: long 時間窗各模型 IC 分佈。