

使用 LLM 分析市場情緒之股價預測模型

A Stock Price Prediction Model Using LLM-Based Market Attention Analysis

指導教師：溫育璋 博士

學生：陳凱華 陳宇帆 董子樂

國立聯合大學資訊工程學系

苗栗市南勢里聯大 2 號

摘要

本研究建立一套結合金融新聞情緒分析與時間序列模型（TCN）的股價預測系統，以 MAG-7 為研究對象。系統蒐集財經新聞，並使用 Financial-RoBERTa 將新聞文本分類，再聚合為每日情緒分數。接著，將情緒特徵與股價、技術指標等技術面相關數據整合，輸入 TCN 模型預測 T+1、T+2 與 T+3 的股價。整體而言，本研究驗證新聞情緒可作為輔助股價預測的重要參考資訊。

Abstract

This study develops a stock prediction system that integrates financial news sentiment analysis with a time-series model, focusing on MAG-7 stocks. The system collects financial news and uses Financial-RoBERTa to classify news text, which is then aggregated into daily sentiment scores. These sentiment features are combined with stock prices, technical indicators, and other technical market data, and are fed into a TCN model to predict stock prices for T+1, T+2,

and T+3. Overall, this study verifies that news sentiment can serve as an important reference for supporting stock price prediction.

一、緒論

金融市場的價格變化不僅受到歷史股價、成交量與技術指標影響，也會受到新聞報導、財報資訊、產業消息與市場情緒等非結構化資訊影響。傳統股價預測模型多以價格與成交量等結構化資料為主要輸入，雖然能反映過去市場走勢，但無法處理新聞文字中隱含的語意與情緒訊號。因此，若能將金融新聞轉換為可量化的情緒特徵，並與股價資料共同建模，便有機會補足傳統模型缺乏消息面資訊的限制。

本研究建立一套由新聞蒐集、情緒推論、特徵整合至股價預測的完整流程。系統以 MAG-7 作為研究對象，包含 AMZN、NVDA、GOOG、AAPL、TSLA、META 與 MSFT。研究流程先蒐集與目標股票相關之財經新聞，再使用金融語言模型進行情緒分析，將新聞內容轉換為每日股票情緒分數。接著，系統將情

緒分數、股價資料及技術指標整合，並輸入 TCN 模型進行未來 T+1、T+2 與 T+3 的股價報酬預測。

本研究的目的是在於驗證新聞情緒特徵是否能提升短期股價預測表現。為了進行比較，本研究建立加入新聞情緒特徵的 news_v18 模型，以及未加入新聞情緒特徵的 no_news_v18 模型，並以相同股票範圍與相同樣本外期間進行評估。透過模型間的比較，可觀察新聞情緒資訊是否能在相同資料條件下提供額外的預測輔助。

二、文獻探討

近年來，自然語言處理技術逐漸被應用於金融文本分析與股價預測任務中。金融市場常受到新聞事件與市場情緒等非結構化資訊影響，因此如何將新聞文字轉換為可供模型使用的量化特徵，成為金融預測研究中的重要議題。

在金融情緒分析方面，本研究採用 Financial-RoBERTa 作為新聞情緒推論模型。該模型以 RoBERTa 架構為基礎，可針對金融文本進行情緒分類，將新聞內容判斷為正面、中性或負面〔1〕。RoBERTa 是由 Liu 等人提出的改良式 BERT 預訓練模型，透過更充分的訓練策略提升語言理解能力〔2〕。本研究利用此類金融語言模型將財經新聞轉換為情緒訊號，使原本無法直接輸入時間序列模型的文字資料，能被整理為每日股票情緒分數。

在新聞情緒量化方面，Shapiro、Sudhof 與 Wilson 提出將新聞文本轉換為情緒指標的方法，並分析新聞情緒與經濟變化之間的關聯〔3〕。此概念說明新聞內容並非只能作為文字資料保存，也能透過量化方式成為可分析的時間序列特徵。本研究參考此作法，將新聞依股票代號與日期進行聚合，建立每日情緒分數。

此外，技術指標仍是股價預測中常見的重要特徵。Un 與 Li 的研究指出，可透過大量技術交易規則建立技術分析特徵空間，再進一步篩選較具代表性的指標〔4〕。本研究同樣建立大量候選技術指標，並透過 IC、FDR 與 JMIM 等方法進行過濾與排序，以降低冗餘特徵對模型訓練的干擾。

在時間序列預測方面，Bai、Kolter 與 Koltun 提出的 Temporal Convolutional Network (TCN) 使用因果卷積、膨脹卷積與殘差連接處理序列資料，能在避免未來資料洩漏的情況下學習較長時間範圍的變化關係〔5〕。因此，本研究採用 TCN 作為股價預測模型，用於整合過去股價、技術指標與新聞情緒特徵，預測未來 T+1、T+2 與 T+3 的股價變化。綜合上述研究，本專題以金融新聞情緒分析補足傳統技術面資料的不足，並結合 TCN 時間序列模型，建立一套整合消息面與技術面資訊的股價預測方法。

三、系統架構

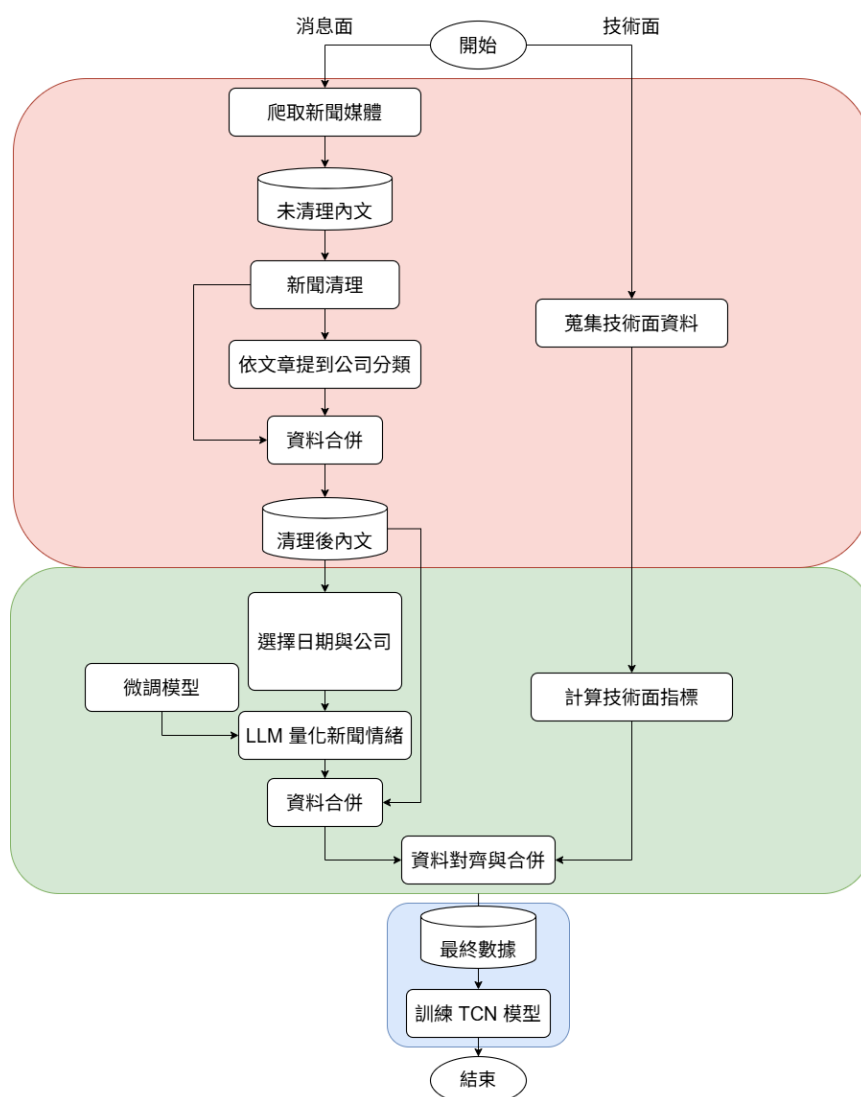


圖 1 系統流程圖

(一) 金融新聞蒐集與情緒量化

本研究系統首先針對 MAG-7 股票蒐集財經新聞資料。系統依據股票代號、公司名稱與相關關鍵字搜尋新聞，並擷取新聞標題、內文、發布日期、新聞來源與相關股票代號等資訊。為降低資料雜訊，系統會移除重複文章及內容過短文章。完成清理後，新聞會依日期與股票代號儲存，建立後續情緒分析所需的基礎資料。

在情緒推論階段，系統將新聞標題與

內文切分為句子，再輸入 Financial-RoBERTa 進行情緒分類。金融新聞常同時包含正面、中性與負面資訊，若直接以全文進行單一判斷，可能忽略文章內部不同句子的情緒差異。完成句子層級推論後，系統會將句子情緒結果彙整為文章層級情緒分數，再依照股票代號與日期聚合為每日情緒訊號。

每日情緒訊號可反映特定股票在某一交易日所受到的新聞情緒影響。系統會將

每日情緒分數與股價資料依日期對齊，並在建模時避免使用未來資訊，以降低前視偏誤。透過此流程，原本無法直接輸入模型的新聞文字，能被轉換為可供時間序列模型使用的量化特徵。

(二) 股價資料與技術指標整合

在股價資料處理方面，系統讀取 MAG-7 股票的 OHLCV 資料，包含開盤價、最高價、最低價、收盤價與成交量。這些資料會進一步轉換為對數報酬率、價格區間變化、成交量變化與滾動波動率等特徵，以描述股票在不同時間尺度下的價格變化與波動狀態。

除了個股本身的股價資料外，本研究也加入技術指標、市場基準與總體經濟資料。技術指標用於捕捉趨勢、動能、突破與波動特徵；市場基準資料則用於區分個股自身變化與整體市場波動；總體經濟資料則提供市場環境的背景資訊。所有資料會依股票代號與日期合併，形成模型訓練所需的多變量時間序列資料集。

在技術指標篩選方面，本研究共產生 1009 個候選技術指標。為避免模型輸入重複或無效特徵，系統先以資訊係數 (Information Coefficient, IC) 衡量各技術指標與未來三日股價變化之間的相關程度 [6]，並以季度為單位檢查其預測能力，再進一步檢查指標在不同季度間的穩定性。接著透過 JMIM 進行排序。實驗比較後，本研究採用 JMIM 排名前 30 的技術指標作為模型輸入 [7][8]。

(三) TCN 股價趨勢預測模型

本研究使用 TCN 作為股票趨勢預測模型。模型輸入為過去 60 個交易日的多維時間序列特徵，包含股價特徵、技術指

標、市場基準特徵、總體經濟資料與新聞情緒分數。模型輸出為未來 T+1、T+2 與 T+3 的股價報酬與波動範圍。

TCN 模型主要由因果卷積、膨脹卷積、正規化層、啟用函數、Dropout 與殘差連接組成。因果卷積可確保模型只使用過去資訊進行預測，避免未來資料洩漏；膨脹卷積可在不大幅增加參數量的情況下擴大感受野，使模型能學習較長時間範圍內的變化關係；殘差連接則有助於穩定深層模型的訓練。

模型訓練時會同時預測 T+1、T+2 與 T+3 的報酬率與波動，使系統不只提供點預測，也能提供不確定性範圍。完成訓練後，模型會將預測結果寫入資料庫，供後端系統與前端網頁查詢使用。前端可顯示不同股票的歷史股價資訊、未來三日預測趨勢，以及特定日期對應的新聞資料與模型預測結果。

四、實驗

(一) 實驗設定

本研究訓練的模型輸入為過去 60 個交易日之時間序列特徵，包含量價特徵、技術指標、市場基準指數與每日新聞情緒分數。技術指標取 JMIM 排名前 30 者作為輸入特徵。資料切分依時間順序，避免訓練與驗證資料間產生時間資料洩漏。

為驗證新聞情緒是否能提升短期股價預測表現，本研究建立兩組對照模型：一組為加入新聞情緒特徵的 news_v18 模型，另一組為未加入新聞情緒特徵的 no_news_v18 模型，兩組模型除新聞情緒特徵外採用相同的網路架構、資料集與訓練參數。兩組模型皆以相同的樣本外期間進行測試，資料涵蓋 2025 年 1 月 2

日至 2026 年 5 月 20 日，包含 MAG-7 共 2422 筆預測資料。透過相同訓練資料之比較，可公平地觀察新聞情緒特徵對模型預測結果的實際貢獻。在評估指標方面，本研究除使用平均絕對誤差

(MAE) 等點預測誤差指標外，亦納入多項相對於基準方法的比較指標，包含衡量預測報酬與實際報酬線性相關程度的 Pearson-r、相對於 naive baseline 改善幅度的 MAE Improve、模型與基準方法 RMSE 比值的 Theil-U [9]、逐筆擊敗基準方法比例的 Beat Naive，以綜合判斷模型是否真正優於基準預測方法。

(二) 模型結果與分析

在表 1 中，news_v18 於三個預測天期皆取得正向改善。此結果顯示加入新聞情緒特徵後，模型在三個預測天期的平均誤差皆較基準方法略有改善，且在 T+2 明顯優於未加入新聞特徵的模型。

表 1 MAE Improve 指標數值

model	T+1	T+2	T+3
no_news_v18	0.13%	-0.08%	0.11%
news_v18	0.16%	0.05%	0.14%

在表 2 的逐筆勝率方面，顯示 news_v18 在 T+1 與 T+2 的短期預測中有較高比例優於基準方法。

表 2 Beat Naive 指標數值

	T+1	T+2	T+3
no_news_v18	50.52%	49.42%	52.19%
news_v18	51.55%	49.71%	51.02%

在表 3 中，兩組模型整體皆與基準方法接近，但 news_v18 在 T+1 與 T+2 的相對誤差略低，且三個天期表現較為平均。

表 3 Theil-U 指標數值

	T+1	T+2	T+3
no_news_v18	1.0001	1.0014	0.9997
news_v18	0.9995	1.0010	1.0003

在表 4 中，news_v18 在 T+1 的相關性優於 no_news_v18，代表新聞情緒對隔日報酬方向具有較直接的弱訊號，但在 T+2 與 T+3 兩組模型的 Pearson-r 皆接近零或為負值，顯示模型對較長天期的報酬排序與方向掌握仍有限。

表 4 Pearson-r 指標數值

	T+1	T+2	T+3
no_news_v18	-0.0054	-0.0237	0.0002
news_v18	0.0133	-0.0311	-0.0331

綜合而言，news_v18 整體表現優於 no_news_v18，證明新聞情緒特徵在本研究架構中具有實驗價值；惟模型尚未能穩定預測報酬方向，反映短期股價本身具有高度雜訊，單一模型難以穩定捕捉所有市場變化。

五、結論與建議

本研究建立一套結合金融新聞情緒分析與時間序列模型的股價預測系統，整體流程包含財經新聞蒐集、新聞文本情緒分類、每日情緒分數聚合、股價與技術指標整合，以及 TCN 股價預測模型建置。透過此流程，本研究將原本屬於非結構化資料的新聞文字轉換為可量化的情緒特徵，並與技術面資料共同作為模型輸入，用於預測 MAG-7 股票未來 T+1、T+2 與 T+3 的股價變化。

實驗結果顯示，加入新聞情緒特徵的 news_v18 模型在三個預測天期的 MAE Improve 皆為正值，整體表現較未加入新聞特徵的模型穩定，表示新聞情緒對短期

股價預測具有一定輔助效果。然而，從 Pearson-r 與部分預測天期的結果來看，模型對實際報酬方向與變化幅度的掌握仍有限，顯示短期股價受到大量市場雜訊影響，僅依靠新聞情緒與技術指標仍難以穩定預測所有市場變化。

未來可從資料品質、情緒模型與預測模型三方面進行改進。首先，可擴充新聞來源並改善新聞與股票代號的關聯判斷，以降低不相關新聞造成的雜訊。其次，可針對金融文本進一步微調情緒模型，或加入事件類型、財報消息、產業趨勢等更細緻的分類特徵。最後，除 TCN 外，亦可嘗試 Temporal Fusion Transformer 等模型，以提升多來源時間序列資料的整合能力與模型解釋性〔10〕。整體而言，本研究已完成一套以新聞情緒輔助股價預測的系統雛形，並驗證其作為市場趨勢分析參考資訊的可行性。

六、參考文獻

- 〔1〕 M. Soleimanian, “financial-roberta-large-sentiment,” Hugging Face, 2022. [Online]. Available: <https://huggingface.co/soleimanian/financial-roberta-large-sentiment>
- 〔2〕 Y. Liu, M. Ott, N. Goyal, J. Du, M. Joshi, D. Chen, O. Levy, M. Lewis, L. Zettlemoyer, and V. Stoyanov, “RoBERTa: A robustly optimized BERT pretraining approach,” arXiv preprint arXiv:1907.11692, 2019, doi: 10.48550/arXiv.1907.11692.
- 〔3〕 A. H. Shapiro, M. Sudhof, and D. J. Wilson, “Measuring news sentiment,” *Journal of Econometrics*, vol. 228, no. 2, pp. 221–243, 2022, doi: 10.1016/j.jeconom.2020.07.053.
- 〔4〕 K. S. Un and H. Li, “Technical analysis and stock returns: An aggregate approach,” SSRN preprint, 2022. [Online]. Available: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4277254
- 〔5〕 S. Bai, J. Z. Kolter, and V. Koltun, “An empirical evaluation of generic convolutional and recurrent networks for sequence modeling,” arXiv preprint arXiv:1803.01271, 2018.
- 〔6〕 F. Zhang, R. Guo, and H. Cao, “Information coefficient as a performance measure of stock selection models,” arXiv preprint arXiv:2010.08601, 2020.
- 〔7〕 Y. Benjamini and Y. Hochberg, “Controlling the false discovery rate: A practical and powerful approach to multiple testing,” *Journal of the Royal Statistical Society: Series B*, vol. 57, no. 1, pp. 289–300, 1995, doi: 10.1111/j.2517-6161.1995.tb02031.x.
- 〔8〕 M. Bennasar, Y. Hicks, and R. Setchi, “Feature selection using joint mutual information maximisation,” *Expert Systems with Applications*, vol. 42, no. 22, pp. 8520–8532, Dec. 2015, doi: 10.1016/j.eswa.2015.07.007.
- 〔9〕 H. Theil, *Applied Economic Forecasting*. Amsterdam, Netherlands: North-Holland, 1966.
- 〔10〕 B. Lim, S. O. Arik, N. Loeff, and T. Pfister, “Temporal fusion

transformers for interpretable multi-horizon time series forecasting,”
International Journal of Forecasting,
vol. 37, no. 4, pp. 1748–1764, 2021,
doi: 10.1016/j.ijforecast.2021.03.012.