

AI 市場雷達：以大型語言模型解析新聞對股價影響之預測系統

AI Market Radar: A News-Driven Stock Movement Prediction System Using Large Language Models

組別：C-4

組員：劉詠欣

指導教授：李政德

1 研究動機與目的

一則財經新聞，會讓某檔股票「漲」還是「跌」？LLM 擅長讀文字，是否能用它即時判斷，並做成人人可用的工具？但這類工具常只能給「看起來很有信心」的預測，卻無法判斷是否可靠。

目的

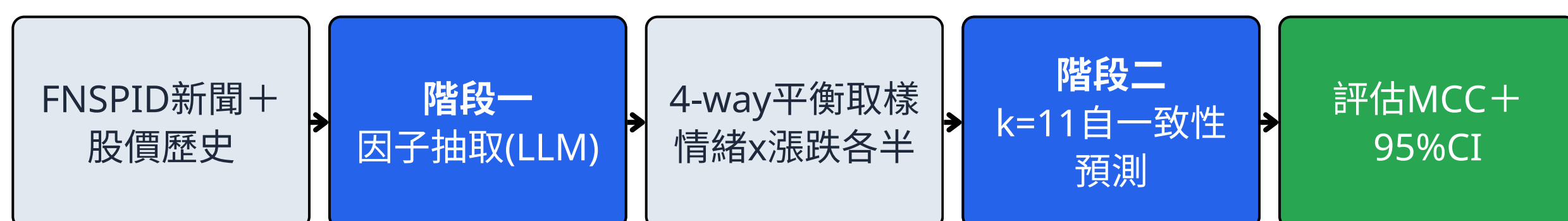
- 打造可用工具：Chrome 擴充功能，一鍵把當前新聞分頁轉成個股漲跌預測。
- 讓工具誠實：加入以統計為依據的「可靠度標章」，標示每個股票×時間窗能不能信。
- 為支撐標章，建立嚴謹的離線評估（50/50 平衡、MCC+信賴區間、與基準及更大模型比較），量化真實預測力。

2 實作技術與環境

- 前端：JavaScript (ES Modules)、Chrome Manifest V3、HTML / CSS
- LLM：Gemini API (gemini-2.5-flash-lite, Batch)；本地 Ollama + Qwen2.5-7B (備援)
- 評估管線：Python、pandas、NumPy、scikit-learn、PyArrow、Matplotlib
- 基準對照：FinBERT (PyTorch)、20 日價格動量、gemini-2.5-pro、claude-opus
- 資料行情：FNSPID、Polygon.io、yfinance
- 開發環境：macOS、Git (雙 repo：Extension／評估)

3 系統架構圖

Evaluation Pipeline



利用gemini-2.5-flash-lite
評估結果產生可靠度指標標章

Chrome Extension



模型邏輯 (兩階段，兩端共用)

階段一：相關性+3-5 因果因子

階段二：因子+價格摘要+公司資訊+時間窗提示

→ UP/DOWN+信心

評估端

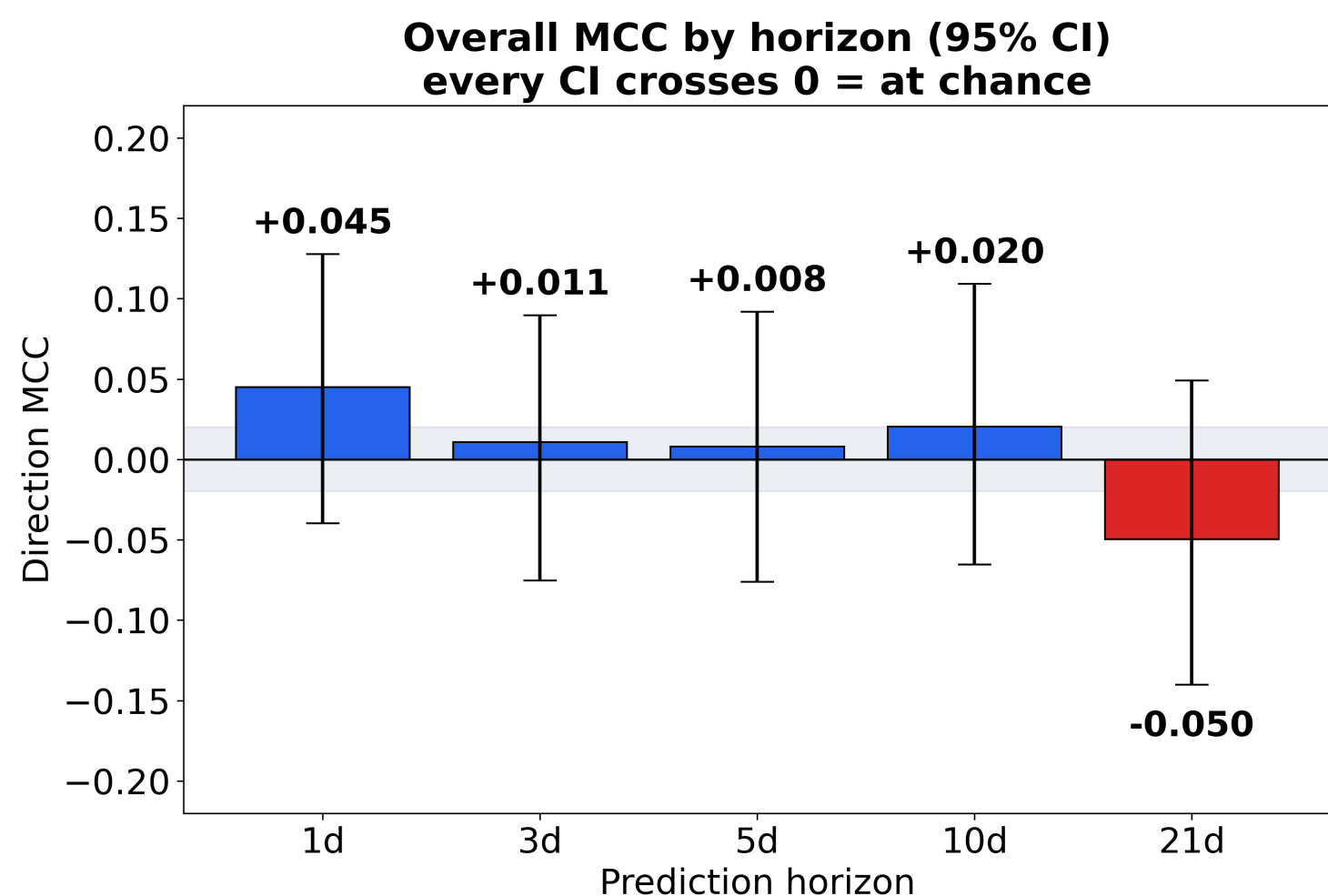
- K=11 自一致性、4-way 平衡 (50/50)、MCC+bootstrap 95% CI
- 資料：FNSPID (2023-25)、7 檔美股權值股、~3,000 篇

擴充功能端

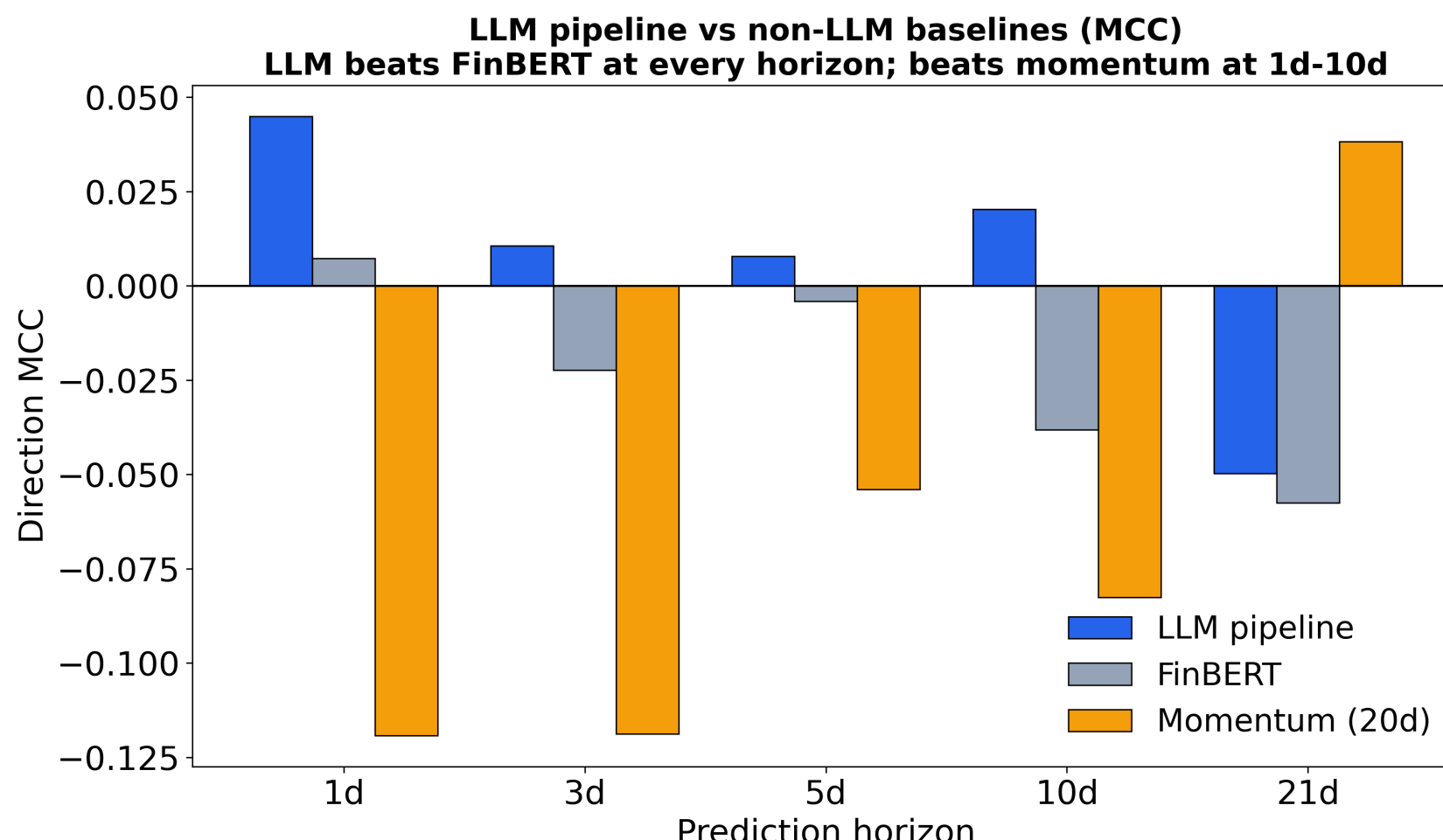
- 即時：開啟任意新聞分頁一鍵分析；現價／公司資訊用 Polygon.io
- 線上版單次預測 (K=1)，求速度 (評估用 K=11)
- 可靠度標章：把評估的 MCC 信賴區間 → 「有訊號／與隨機無異／比隨機差」

4 成果展示

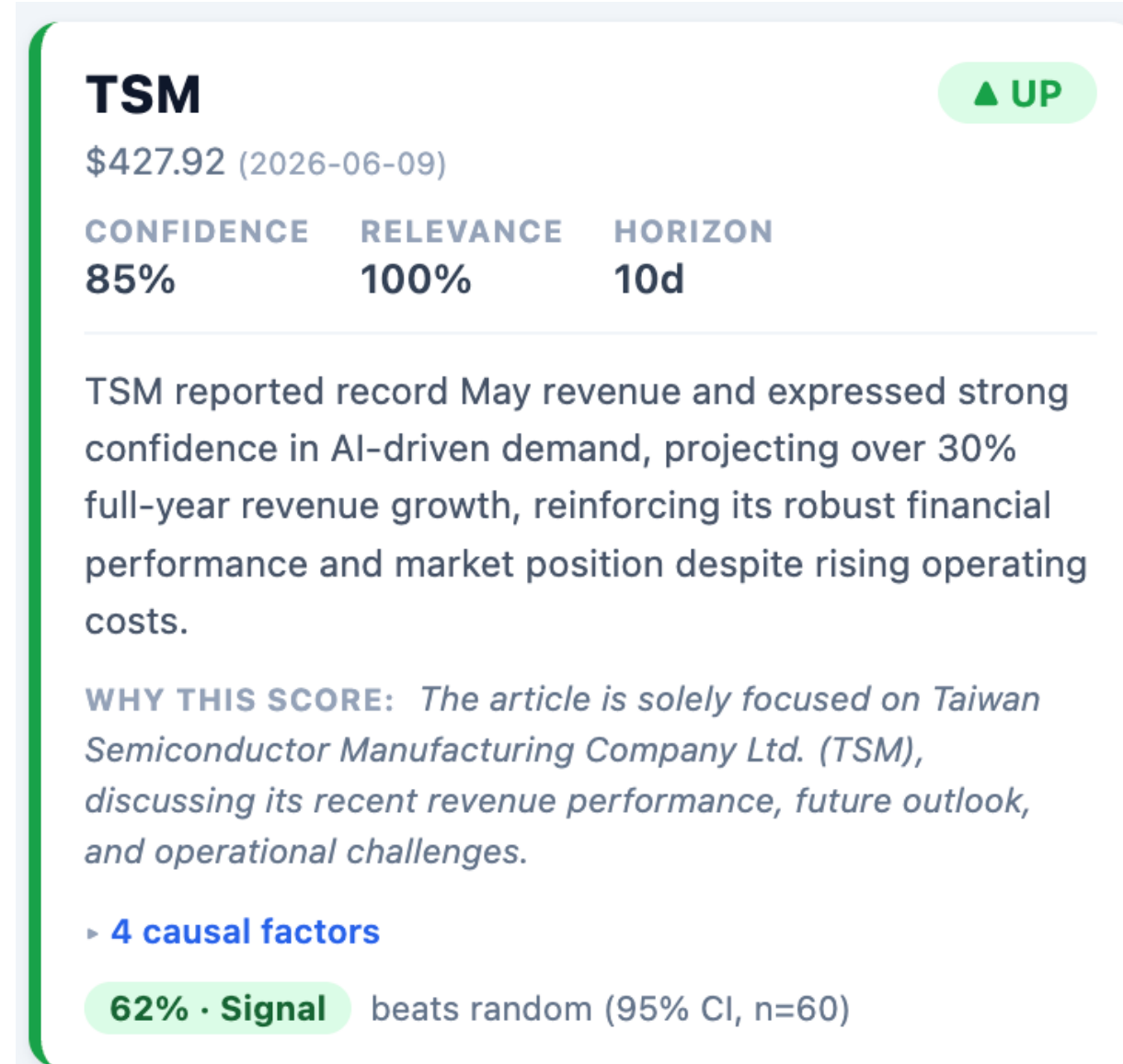
1d 準確率 52.2% · MCC +0.045 (95% CI 跨越 0 → 接近隨機)
35 組 (股票×時間窗) 多重比較校正後 → 0 組顯著 · LLM 每個時間窗都勝過 FinBERT，放大模型無幫助



各時間窗整體 MCC 與 95% CI
所有 CI 跨越 0 = 與隨機無異



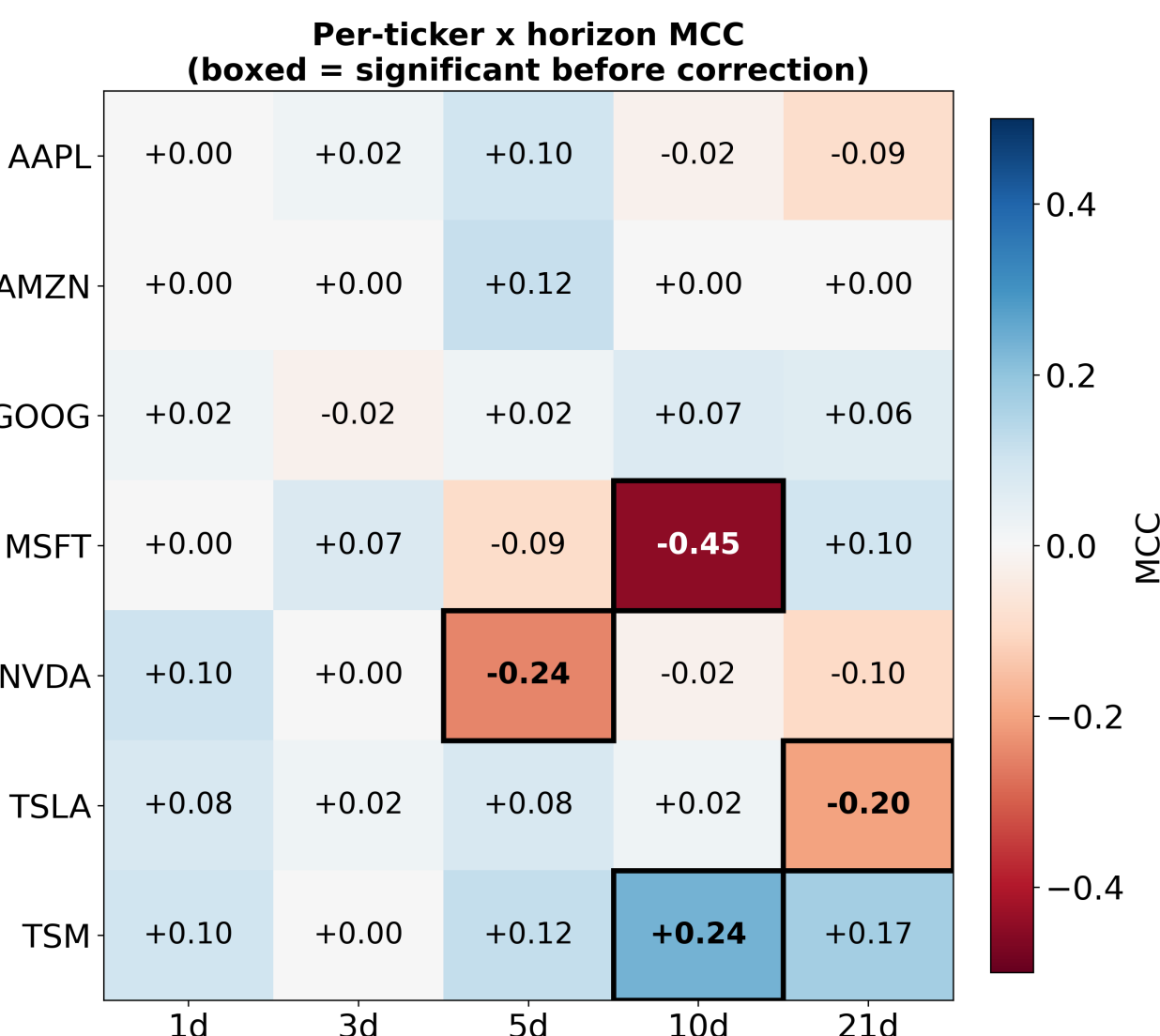
LLM vs FinBERT vs 動量 (MCC)
1d 時 LLM 高 +0.038 / +0.16



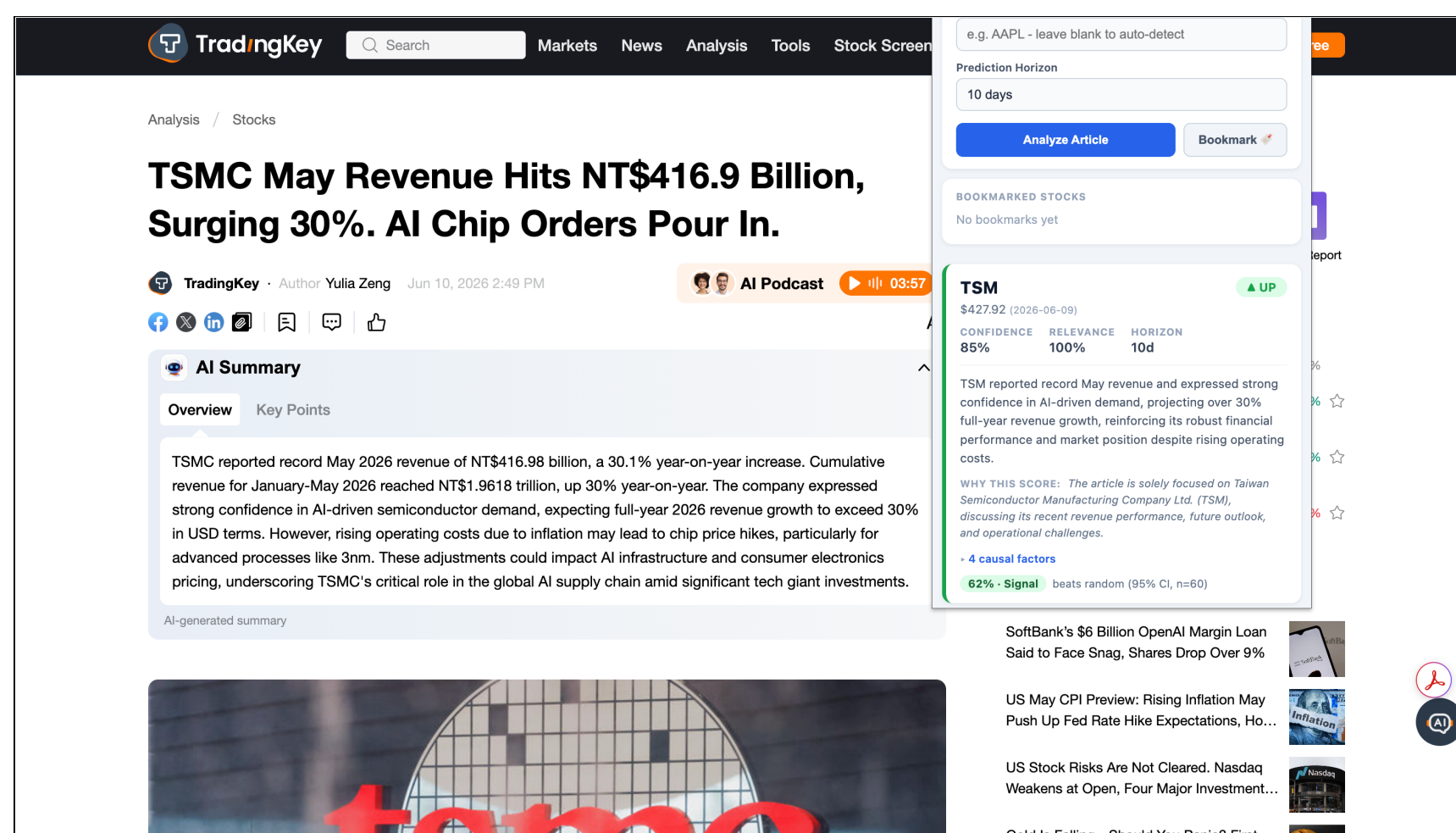
預測卡 (方向／信心／因子) 與
「誠實可靠度標章」

可靠度標章 - 圖例

有訊號 (Signal)	MCC 95% CI 完全 > 0 → 真的有預測力
與隨機無異 (At chance)	CI 跨越 0 → 無法與隨機區分
比隨機差 (Underperforms)	CI 完全 < 0



各股票×時間窗 MCC 熱力圖
黑框＝校正前顯著，校正後全部不顯著



Chrome Extension 介面

5 結論與未來展望

結論 LLM 從新聞預測大型權值股方向接近隨機，放大模型沒有幫助。瓶頸是市場效率，不是模型能力。本流程仍勝過 FinBERT 與價格動量，並以 50/50 平衡、MCC+信賴區間、多重比較校正，誠實量化此極限

未來 各股票分別校準閾值、以訓練截止後的新聞做留出測試、同日多篇文章整合、擴大標的至中小型股／非美股

Demo Video

