



股票價格預測

李柏漢	資訊碩一	113753218
林靖淵	資管碩一	113356040
陳昶安	資訊碩一	113753121
林祐祥	資訊碩一	113753114
廖偉哲	資訊碩一	113753222
陳彥融	資管四乙	110306018



第一部分

主題發想



研究動機

台灣科技產業在全球供應鏈中扮演關鍵角色，其股價波動不僅受國內經濟影響，更與國際市場高度連動。對投資者與研究人員而言，能夠準確預測這些企業的股價變化，有助於提升決策品質與降低投資風險

傳統的技術分析與經濟模型常難以精確掌握，本研究因此結合資料科學與機器學習技術，期望發展出更具準確性與前瞻性的股價預測方法



研究概述

本研究以台灣科技類股為對象，蒐集其歷史股價、成交量與技術指標等項目，進行時間序列建模。透過長短期記憶模型(LSTM)預測股價走勢，以掌握其潛在的非線性與長期依賴特徵

LSTM 是一個具時間記憶能力的遞迴神經網路，能有效處理股價資料的動態變化，目標是建構一套具實務應用價值的深度學習架構，協助投資人制定前瞻性決策



第二部分

資料介紹



資料蒐集

本研究採⊠網路爬蟲技術，⊠TWSE 臺灣證卷交易所網站擷取台灣科技優質股的歷史資料，蒐集項⊠包含

- 1.⊠收盤價、開⊠低價、成交量等常⊠項⊠
- 2.技術指標（動能指標、價格變動率等）
- 3.時間範圍涵蓋近五年資料，以利捕捉⊠期趨勢



臺灣證券交易所

TWSE

台灣證卷交易所資料

日期	成交股數	成交金額	開盤價	最高價	最低價	收盤價	漲跌價差	成交筆數
114/03/03	2,449,933	451,945,999	187.00	190.50	181.50	181.50	-6.50	2,940
114/03/04	2,030,139	362,395,757	179.00	182.50	174.50	182.00	+0.50	2,307
114/03/05	2,441,900	453,633,463	182.50	189.00	181.00	185.50	+3.50	2,484
114/03/06	1,842,269	337,244,714	186.50	188.00	180.00	180.50	-5.00	2,304
114/03/07	1,005,109	180,872,440	180.50	182.50	178.00	178.00	-2.50	1,500
114/03/10	1,182,342	211,983,008	180.50	182.00	177.00	177.50	-0.50	1,444
114/03/11	1,726,247	296,081,963	172.00	173.50	169.00	171.50	-6.00	2,318
114/03/12	1,119,650	192,346,186	173.00	173.50	171.00	171.00	-0.50	1,450
114/03/13	1,835,895	310,554,552	172.50	173.50	166.50	167.00	-4.00	2,846
114/03/14	1,099,961	184,001,426	168.00	169.00	165.50	168.50	+1.50	1,559



技術指標

在股票預測中，技術指標是從歷史價格與成交量中計算出的統計數據，用以判斷市場趨勢與買賣時機，例如

1. 價格變動率

正表示價格上升，為負表示下跌

2. 動能指標

反映了價格上漲或下跌的速度，有助於捕捉趨勢強度與反轉信號



第三部分

資料前處理



步驟

為提升模型預測準確性與穩定性，本研究對原始股價資料進行以下前處理步驟

1. Missing Value

補缺漏期與價格資訊
確保沒有空值

2. Feature Selection

挑選對預測最有幫助的特徵，去除無關冗餘的資料，提升模型效能



步驟

3.Standardization

- Min-Max Scaling 將各欄位特徵壓縮 □ $[0, 1]$ 範圍
- 確保不同尺度資料在模型中具同等重要性

4.分割資料

依照時間順序切分訓練集與測試集，避免未來資料洩漏
保持時間序列連貫性，模擬實際預測情境

部分資料展示

透過這些處理步驟，建立乾淨且具預測意義的特徵資料集，使 LSTM 模型更有效學習時間依賴性與價格趨勢

1	date	target	open	high	low	close	volume	rsi14	ma5	macd_diff
2	2021/3/2	1	-0.05465	-0.06236	-0.0685	-0.06997	0.157358	0.041221	-0.06251	-0.07916
3	2021/3/3	0	-0.06647	-0.07398	-0.07035	-0.06815	-0.00215	0.091015	-0.06762	-0.09634
4	2021/3/4	0	-0.07374	-0.08023	-0.08055	-0.08274	0.119694	-0.32442	-0.07236	-0.29044
5	2021/3/5	0	-0.09102	-0.09452	-0.0889	-0.09369	0.094368	-0.60577	-0.07948	-0.53832
6	2021/3/8	0	-0.08556	-0.09363	-0.0954	-0.1019	0.041581	-0.80501	-0.08313	-0.77127
7	2021/3/9	1	-0.10921	-0.11596	-0.1158	-0.11467	0.352903	-1.09375	-0.09208	-1.04573
8	2021/3/10	1	-0.10375	-0.10792	-0.1056	-0.11194	0.122354	-0.99507	-0.10084	-1.11845
9	2021/3/11	1	-0.10739	-0.08827	-0.10282	-0.08183	0.251106	-0.06017	-0.10066	-0.67537
10	2021/3/12	0	-0.07829	-0.0838	-0.07963	-0.08092	0.192084	-0.03545	-0.0981	-0.33271

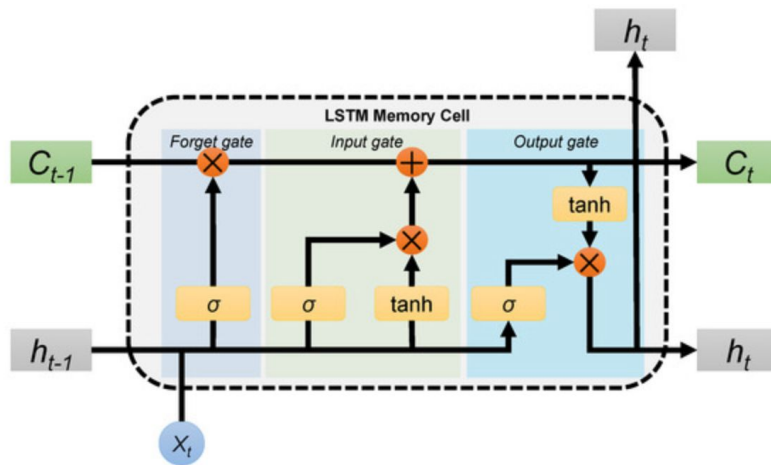


第四部分

模型架構

介紹

LSTM(Long Short-Term Memory)是一種改良的遞迴式神經網路，具有記憶長期時間序列資訊的能力，與傳統 RNN 不同，LSTM 能控制資訊的保留與更新，有效避免長序列中常見的梯度消失問題





優勢

由於股票價格本身具有時間依賴性與非線性波動，LSTM 特別適合捕捉長期趨勢與短期變化的複雜模式，並且能利用過去價格、成交量與技術指標資料，建立動態預測模型，提升預測的準確度與穩定性，並為投資決策提供重要參考依據



模型

輸入層：三十天歷史資料



1D 卷積層：用以抽取局部時序特徵



雙向 LSTM 層：可同時捕捉前向與後向的長期依賴



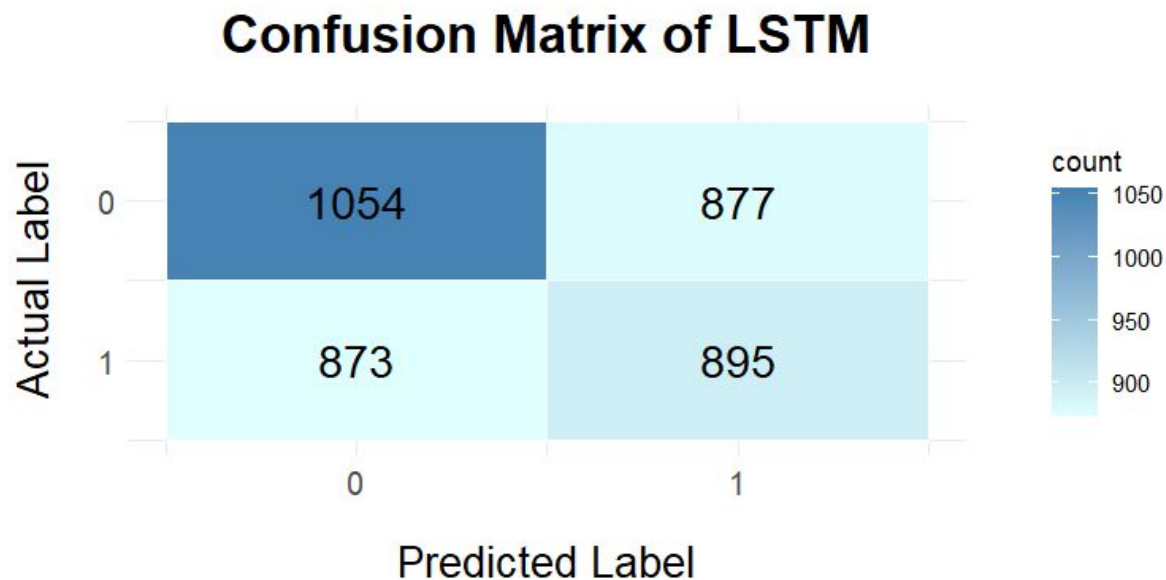
輸出層：股票價格上漲或下跌



第五部分

結果

Confusion Matrix





結論

- 1.成功以 LSTM 模型預測股價，捕捉時間序列中的□線性與趨勢特徵
- 2.整合各項技術指標，提升模型在股價預測上的準確率與穩定性
- 3.從資料蒐集、前處理到模型訓練，形成可複□的預測架構



第六部分

未來展望



未來規劃

未來，我們計畫引入更多外部變數，如國際股市指數與新聞情緒，進一步強化模型對市場動態的反應能力。同時，也將探索Transformer 等先進模型架構，提升預測效能與可解釋性。

此外，預計部署自動化爬蟲機制，定期更新資料集，實現即時、智慧的股價預測平台，協助投資做出更具前瞻性的決策



感謝您的聆聽

Thanks for listening