蕭國倫、劉柏辰、蔡泊均 (2024)，「以時間卷積網路結合長短期記憶模型預測股價：臺股預測實證」，*資訊管理學報*，第三十一卷，第二期，頁 177-207。

以時間卷積網路結合長短期記憶模型預測股價：臺股預測實證

蕭國倫 *
國立臺中科技大學資訊管理學系

劉柏辰
國立臺中科技大學資訊管理學系

蔡泊均
國立臺中科技大學資訊管理學系

摘要

股價預測一直是一個很棘手的問題，由於許多因素都會影響股價，因此簡單的模型無法準確預測。但添加太多的特徵將增加模型的複雜度，若能找到關鍵的特徵，模型的準確性將會更好。在股票預測領域中，許多使用長短期記憶(Long Short-Term Memory; LSTM)的研究顯示了良好的結果。時間卷積網絡(Temporal Convolution Network; TCN)在時間序列研究中也取得了不錯的成果。因此，本研究採用了時間卷積網絡(TCN)與長短期記憶(LSTM)相結合的方法，並與 RNN-LSTM、CNN-LSTM 和長短期記憶(LSTM)等三種深度學習模型進行比較。為了比較不同模型的效能，本研究使用了不同的損失函數進行比較。結果表明，本研究提出的 TCN-LSTM 模型相比其他三種模型的結果更好。除了使用過去的數據集回測，本研究還基於所提出的方法開發了一種實際操作方法，並使用當沖交易進行驗證。根據股價預測趨勢圖發現，當遇到大的價格趨勢波動時，本研究提出的 TCN-LSTM 模型的預測結果與真實值的差距比 LSTM 模型更小。根據結果，我們可以得出結論，即 TCN-LSTM 更適用於相對活躍且對沖量較高的股票。本研究的發現，結果與討論可為該領域有所貢獻，並啟發相關領域的從業人員。

關鍵詞：股票預測、深度學習、時間卷積網路(TCN)、長短期記憶(LSTM)

* 本文通訊作者。電子郵件信箱：hsiao.kuolun@gmail.com
2023/04/03 投稿；2023/07/19 修訂；2023/11/01 接受

Hsiao, K.L., Liu, B.C., & Tsai, P.C. (2024). Predicting Stock Prices by Combining Long- and Short-Term Memory with Time Convolutional Networks: An Empirical Study on the Stocks in Taiwan. *Journal of Information Management*, 31(2), 177-207.

Predicting Stock Prices by Combining Long- and Short-Term Memory with Time Convolutional Networks: An Empirical Study on the Stocks in Taiwan

Kuo-Lun Hsiao*
Department of Information Management, National Taichung University of Science and Technology

Bo-Chen Liu
Department of Information Management, National Taichung University of Science and Technology

Po-Chun Tsai
Department of Information Management, National Taichung University of Science and Technology

Abstract

Predicting stock prices has always been a challenging task due to the multitude of factors that can influence them. Adding too many features can make the model overly complex, so identifying key features is crucial for accuracy. In the field of stock price prediction, many studies have shown that Long Short-Term Memory (LSTM) models perform well. Similarly, the Temporal Convolution Network (TCN) has achieved good results in time series research. Therefore, this study combines LSTM and TCN models and compares them with RNN-LSTM, CNN-LSTM, and LSTM models for stock price prediction using various loss functions. The results indicate that the proposed TCN-LSTM model performs better than the other models. This study not only tested the proposed method with historical data sets but also validated it through day trading. The TCN-LSTM model proposed in this study outperforms the LSTM model in predicting stock prices during large price trend fluctuations, making it more suitable for active

* Corresponding author : Email:hsiao.kuolun@gmail.com
2023/04/03 received; 2023/07/19 revised; 2023/11/01 accepted

stocks with high hedging volume. These findings can contribute to the field and inspire practitioners in related fields.

Keywords: Stock prediction, Deep learning, Temporal Convolutional Network (TCN), Long Short-Term Memory (LSTM)

壹、緒論

一、研究背景

金融市場被認為是世界經濟的心臟，每天有數十億美元的交易。金融市場包含外匯市場、股票市場和債券市場等等，而股票市場在經濟增長中有著重要作用。顯然，如何良好的預測金融市場未來的狀態在各個領域都將非常有價值。加上近年來，投資理財的觀念越來越受重視，基金和股票等金融商品已被當作投資理財的工具，投資理財已經是大眾最關心的話題，而投資股票是一般大眾最矚目的投資管道。根據台灣證券交易所 (2021)的數據顯示開戶人數有逐年上升的趨勢，如圖 1 所示。

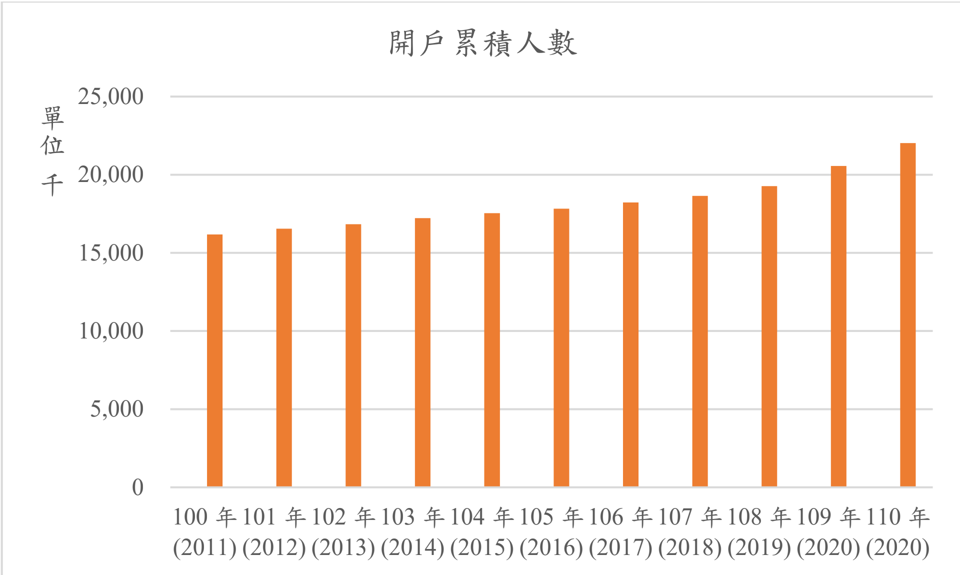

圖 1：台灣證券交易所證券累積開戶人數

現有的股票預測方法大致可分為基本面分析、技術面分析和時間序列分析 (Devadoss & Ligori 2013)，基本面分析：是一種對企業進行財務分析和經濟學上的研究來評估企業的價值或預測證券價值的走勢；技術面分析：研究市場過去的資料來預測未來的走向，使用股票的歷史價格來確定未來價格，而金融市場中最常使用的技術分析，像是 K 線、KD 指標或是 KD 黃金交叉等等；時間序列分析：時間序列是指在某一段特定時間的順序集合，時間序列模型運用在很多的領域，無論是在金融、醫療或是天氣等多種領域皆有重要的應用。簡單來說，時間序列預測是一種通過分析過去的數據行為或趨勢來預測未來事件的技術。因此，有效的運用將可以發現股票市場的趨勢。

在傳統中股票預測使用統計方式來進行預測，隨著演算法的進步，再加上統計和學習模型相互結合，得以完善了幾種機器學習演算法，機器學習顧名思義是電腦利用數據自主學習並優化性能，根據其訓練資料、產出判別的過程與結果的不同。隨著近年來大數據時代的來臨，如何從龐大資料中取得資料並進行分析是一件很重要的事，使得深度學習崛起並開始被大量運用在像是農業(Chen, Kung,

& Hwang 2019)、醫療 (Rahimy 2018)和金融市場 (Sezer, Ozbayoglu, & Dogdu 2017)等等各個領域中。

二、研究動機

投資從來都不是一件容易的事,金融市場的不確定性和突發性事件對股票都有所影響，再加上影響股票的因素非常的多，使得簡單的模型無法更準確地預測未來資產的價值。也因此讓許多研究人員想以更複雜的模型來獲得更好的預測結果，以增加收益。現今已經有許多研究將深度學習方法運用於金融領域，例如：有學者使用生成對抗網路(Generative Adversarial Network; GAN)對中國股市進行價格預測 (Zhou et al. 2018)，還有學者將長短期記憶(Long Short-Term Memory; LSTM)用於外匯交易系統中的金融趨勢預測 (Rundo 2019)。雖然影響股票價格的因素很多，若是添加過多的時事數據，像是政治因素和物價指數等等。隨著時間和數據的增加，會使得預測過於複雜，反而會增加預測的難度。因此如何從數據中篩選出關鍵特徵，並運用合適的數據進行價格的預測至關重要。

Bai, Kolter, & Koltun (2018) 提出一種深度學習的新模型，為時間卷積網路(Temporal Convolutional Network; TCN)，可以用來解決時間序列預測的算法，近年來也被運用在很多領域且都有不錯的表現，如 Hewage et al. (2020) 使用時間序列數據進行天氣預測，發現 TCN 在時間序列數據的預測任務中表現較好，也有學者使用時間卷積網路(TCN)用於估計機率密度 (Chen et al. 2020)。在金融領域 Deng et al. (2019) 通過時間卷積網路(TCN)進行知識驅動的股票趨勢預測，還有學者使用時間卷積網路(TCN)結合 Stock2Vec 來預測股票價格 (Wang et al. 2020)。

Hochreiter & Schmidhuber (1997) 提出了長短期記憶(Long Short-Term Memory; LSTM)，主要是為了解決遞歸神經網路(Recurrent Neural Network; RNN)梯度消失的問題，適合於處理時間序列相關的預測任務。LSTM 也被運用在各個領域中，Google 使用 LSTM 進行 Google 翻譯 (Wu et al. 2016)， Zhao et al. (2017)提出在 LSTM 中加入時間加權函數，其結果優於其他模型。雖然 LSTM 等深度學習方法在股票預測都有不錯的表現，但有學者認為混合方法比單一方法預測的效果來的更好，如 Pawar, Jalem, & Tiwari (2019) 提出的 LSTM-RNN 混合方法進行股票預測，結果都比 LSTM 或是 RNN 單一方法都來的更好。

本研究主要使用深度學習方法進行股票價格的預測，使用時間卷積模型(TCN)結合長短期記憶(LSTM)和其他常用在股票價格預測的深度學習模型進行比較。希望能透過每日收集的數據找出更多的隱藏資訊，並獲得相對較準確的價格預測，期望找出合適的操作方式。

三、研究目的

本研究利用網路爬蟲技術來搜集台灣股票市場的每日價格數據集進行實驗研究，數據集中主要的內容為股票數據。在過去 2008 年金融海嘯對經濟市場造

成很大的影響 (Hasan & Mohammad 2015)，而近年來不論是新型冠狀病毒(COVID-19)、2021 年中國大陸停電限電事件或是元宇宙的崛起，隨著時間的增加，對於某些股票的漲跌都特別大，有研究證實新型冠狀病毒(COVID-19)對股票市場產生重大的影響 (Ashraf 2020)，因此本研究選擇一些漲跌幅度較大的股票進行研究。為了預測 2021 年股票市場的趨勢，數據集使用常見的 80/20 來做劃分，所以選擇近五年的時間序列長度來進行實驗設計。雖然 LSTM 的技術已用於財務預測中 (Wu et al. 2021)，而 TCN 的技術也被時間序列數據預測上 (Hewage et al. 2020)，但是這兩種深度學習工具尚未結合使用，以增強預測模型。因此本研究使用 TCN-LSTM 混合模型並運用歷史數據來對每日的收盤價格進行預測。

綜上所述，本研究的目的如下：

1. 從台灣股市中挑選至 2021 年底漲跌幅度較大的幾支個股並蒐集其近五年的每日價格數據，本研究將透過時間卷積網路(TCN)結合長短期記憶(LSTM)並運用個股的歷史價格數據來進行時間序列的價格預測。
2. 本研究使用相同的數據集，與常用來進行時間序列預測的長短期記憶模型(LSTM)、時間卷積網路(TCN)、RNN-LSTM 和 CNN-LSTM 等四種深度學習模型進行預測結果比較，來選取出最好的預測模型。
3. 找出合適的操作規則並運用在實際交易操作上，以驗證模型的效益。

四、研究流程

本研究流程主要區分為兩個階段進行，其各階段研究流程分別敘述如下。

研究流程的第一階段為在確定研究主題和目的後，將會擬定研究流程，並彙整相關文獻資料，對其進行統整後，找出本研究所適用的開發工具以及提出研究架構模型，針對研究模型中的各項變數進行定義。

研究流程的第二階段為根據先前所提出的研究模型，並從彙整完成的參考文獻，找出適合本研究模型的資料和評估模型方式，並對資料進行預處理，再使用模型進行訓練。接著進行資料分析與研究模型的驗證，最後根據分析結果來進行探討與研究結果的撰寫。

貳、文獻回顧

一、股票價格預測

在股票市場中像是管理層的變動和股利的發放等等，有學者認為這些方面都會影響股價(Bharambe & Dharmadhikari 2017)，在過去相關文獻中預測金融時間序列的方法主要分為統計學和人工智能等 (Lin et al. 2021)。由於股票數據包含大量噪聲和不確定因素，隨著預測週期的增長，線性模型的局限性更加明顯(Wu et al. 2019)。學者開始朝向使用非線性模型來進行股票預測，近年來，隨著資訊硬體設備的進步，大數據時代的崛起，深度學習被廣泛運用，Heaton et al. (2017)使用深度學習解決金融預測和分類問題。Wu et al. (2020)使用連續趨勢標記和門控

單元(Gate Recurrent Unit; GRU)預測股票價格。學者們不斷嘗試將各種方法運用在股票預測的應用上，以獲取更加好的結果。

在股票價格預測方面， Usmani & Shamsi (2021) 對使用新聞和股票價格的股票趨勢預測進行了廣泛的研究，統整了新聞事件對於股票趨勢預測的三個領域的研究回顧。Gandhmal & Kumar (2019) 所提出的研究是針對人工神經網路(Artificial Neural Network; ANN)、卷積神經網路(Convolutional Neural Network; CNN)和支援向量機(Support Vector Machine; SVM)等幾種方法去探討其在股票預測相關應用。 Jin et al. (2020) 使用支持向量迴歸(Support Vector Regression; SVR)來預測股票。Han & Kim (2021) 使用人工神經網路(ANN)的多種估值方法對股價進行預測。Hoseinzade & Haratizadeh (2019) 使用卷積神經網路(CNN)對股票市場預測。可以發現，已經有大量學者研究將各種方法運用在股票價格預測相關議題。

二、深度學習

深度學習是機器學習的一種方法，以人工神經網路為架構，對資料進行表徵學習的演算法。是一種具備至少一個隱藏層的神經網路，將一堆資料輸入神經網路中，透過模仿生物神經系統的數學模型，進行不同階層與架構的多次運算和訓練，找出最佳化、最有效的深度學習模型。

(一) 遞歸神經網路(Recurrent Neural Networks; RNN)

遞歸神經網路(RNN)是一種用來建立模型序列化資料的深度學習模型(Medsker & Jain 1999)，透過與過去資訊有相關的狀態來迭代處理序列資料。在過去的研究中，遞歸神經網路(RNN)在語音識別、翻譯和圖像描述等領域已有不錯的成果。雖然 RNN 可以記憶先前的資訊，但 RNN 無法捕捉長期時間之間的關聯。而且會有梯度消失(Vanishing gradient problem)或是梯度爆炸(Exploding gradient problem)的問題 (LeCun, Bengio, & Hinton 2015)。

圖 2 為 RNN 的架構圖，χ 為輸入層、y 為輸出層、h 為隱藏層、U 表示輸入層到隱藏層的權重，它將原始輸入進行運算作為隱藏層的輸入。隱藏層到隱藏層的權重為 W，它是網路的記憶控制者，負責調度記憶。隱藏層到輸出層的權重為 V，從隱藏層學習到的將通過它再一次運算，並作為最終的輸出。圖 2 中的 t 為時間， x_{t-1} 為上一層的輸入層、 y_{t-1} 為上一層輸出值和 h_{t-1} 為上一層的隱藏層。W 為前一次隱藏層的輸出值傳至本次的隱藏層。當前的 y 不只受當層輸入的影響，也會受到上一層輸入的影響。

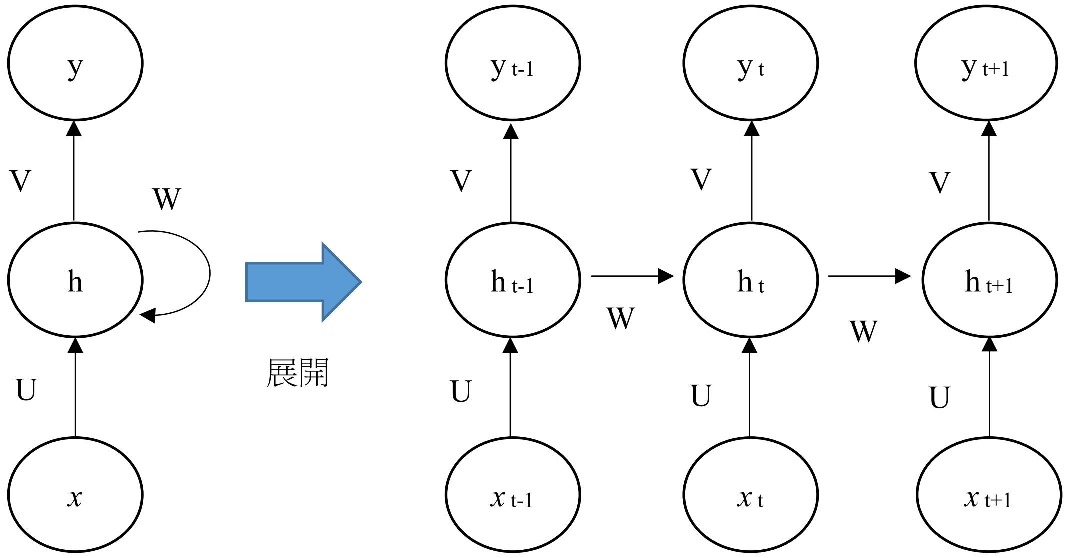

圖 2：遞歸神經網路結構圖

三、混合模型預測財務時間序列價格的相關研究

越來越多研究在財務時間序列的預測。像是深層信任網路（Deep Belief Nets; DBN）、遞歸神經網路（RNN）和門控循環單元（GRU）等模型已在許多股票和金融商品市場的研究中廣泛地被使用。Serrano (2018) 運用深度學習方法對債券和衍生性金融商品進行價格預測。然而由於金融時間序列具有非平穩、非線性、高噪聲的特點，近幾年來，越來越多研究學者嘗試使用混合方法來進行時間序列相關研究。Vidal & Kristjanpoller (2020) 提出 CNN-LSTM 預測黃金波動價格，這項研究的重要性在於應用一種新型架構，該架構能夠為任何時間序列預測任務處理各種信息源。Wu et al. (2021) 使用 CNN-LSTM 進行股票價格預測。Liwei et al. (2021) 基於貝葉斯優化的 LSTM-XGBoost 股價預測，發現比單一 LSTM 網路模型和 RNN 網路模型性能好。Kumar & Haider (2021)在 RNN-LSTM 網路上使用元啟發式優化增強日內股票市場預測，通過使用元啟發式方法獲得大約 4-6%的最大預測精度。Zhang et al. (2021)使用 DBN-RNN 預測中國股價走勢，與其他學習模型相比，所提出的模型在預測性能方面提供了顯著的改進。Anuradha (2021) 使用具有強化 LSTM 模型的深度 CNN 基於大數據的股票趨勢預測，模型性能通過不同的實驗進行評估，發現在各個評估指標都有較好的表現。Chen et al. (2021) 基於 CNN-BiLSTM-ECA 模型的股價預測，結合了卷積神經網路(CNN)、雙向長短期記憶(BiLSTM)網路和注意力機制，與現有方法相比驗證了該方法的有效性和可行性。這些研究成果證明能夠捕捉到更多的隱藏資訊且可獲得更準確的預測，可以發現大多數研究都是使用 CNN 結合 LSTM 的混合模型，而時間卷積網路(TCN)目前較少被應用在財務時間序列，因此本研究將使用時間卷積網路(TCN)結合長短期記憶(LSTM)進行預測，來探討是否會比單一模型和常用於價格預測的混和方法來的準確。其中下表 1 與表 2 為過去學者所採用之混合方法以及單一方法的優缺點說明。這些研究成果證明混合深度學習模型的優勢主要體現在

以下幾個方面。首先，它能夠捕捉到多層次的特徵，從而提高預測的準確性。其次，它可以考慮不同時間尺度的模型組合，充分考慮市場的時間特性，使得預測更全面和準確。此外，混合模型結合了不同模型的優點，強化了預測能力，例如結合 CNN 和 LSTM 可以同時捕捉空間和時間特徵，相比使用單一深度學習模型可以獲得更準確的預測。

在國內相關研究方面，為了更準確地預測股價的漲跌趨勢，郝沛毅等人(2018)，使用新聞文章的隱含主題和情緒資訊，並結合模糊支持向量機，相比傳統支持向量機模型，在股價預測上具有明顯優勢。陳振東、謝政翰(2019)通過篩選技術指標並應用模糊推論方法，建構了一個股價漲跌預測系統，以此提高預測準確率並預測股價漲跌幅度區間。郝沛毅、龔千芬(2022)提出由卷積神經網路(CNN)、雙向長短期記憶(BiLSTM)與注意力機制(AM)組成的混合深度模型，結合社群推文和股價技術指數資料，預測股價的變化。

表 1：混合模型-價格預測文獻整理

題目	研究方法	研究學者
使用 CNN-LSTM 方法預測黃金波動率	CNN-LSTM	Vidal & Kristjanpoller (2020)
一種基於圖的具有領先指標的 CNN-LSTM 股價預測算法	CNN-LSTM	Wu et al. (2021)
基於貝葉斯優化的 LSTM-XGBoost 股價預測	LSTM-XGBoost	Liwei et al. (2021)
在 RNN-LSTM 網路上使用元啟發式優化增強日內股票市場預測	RNN-LSTM	Kumar & Haider (2021)
使用 DBN-RNN 預測股價走勢	DBN-RNN	Zhang et al. (2021)
使用具有強化 LSTM 模型的深度 CNN 基於大數據的股票趨勢預測	DCNN-RLSTM	Anuradha (2021)
基於 CNN-BiLSTM-ECA 模型的股價預測	CNN-BiLSTM-ECA	Chen et al. (2021)
用於美元/人民幣匯率預測的 CNN-TLSTM 預測模型	CNN-TLSTM	J. Wang et al. (2021)
基於 CNN-BiSLSTM 的股票收盤價預測模型	CNN-BiSLSTM	H. Wang et al. (2021)

表 2：不同方法優缺點比較

	優點	缺點	應用	限制
CNN	對於圖像和視覺數據處理效果好，適合處理空間上相鄰的特徵，能夠捕捉局部和全局特徵。	缺乏對時間序列的長期依賴性建模能力，不適合處理序列數據或具有時間相關性的數據。	圖像分類、物體檢測、圖像生成等。	無法捕捉時間序列中的時間依賴性，對於時間相關的數據處理效果不佳。
RNN	能夠捕捉時間序列的時間依賴性，適合處理具有時間相關性的數據。	長期依賴性問題：對於長時間依賴的序列，RNN 往往難以捕捉到長期記憶。 梯度消失和梯度爆炸問題：在長期序列中，梯度在反向傳播過程中可能會出現過小或過大的問題。	語言模型、機器翻譯、語音識別等。	對於長期依賴性問題敏感，梯度消失和梯度爆炸問題。
LSTM	克服了 RNN 的長期依賴性問題，能夠更好地捕捉長期記憶。	計算量較大，相對於 RNN 而言比較複雜。	語言模型、情感分析、語音識別等。	複雜度較高，需要更多的計算資源。
TCN	可以捕捉序列數據中的長期依賴性，模型結構簡單，容易理解和實現，可以並行計算，適用於大規模數據。	對於處理長期記憶的能力較弱。	時間序列預測、機器翻譯、語音識別等。	對於長期記憶的建模能力相對較弱。

參、研究方法

一、時間卷積網路 (Temporal Convolutional Networks; TCN)

Bai et al. (2018) 提出了時間卷積網路(TCN)，該架構由因果卷積層(Causal Convolution)、擴張卷積層(Dilated Convolution)和殘差連接(Residual Connections)所組成，簡單來說 TCN 是一維全卷積網路(Fully Convolutional Network; FCN)加上因果卷積層。

(一)因果卷積層(Causal Convolutions)

因果卷積可以解決序列問題，可以將序列問題轉化為根據 $x_1, x_2, \ldots, x_t$ 去預測 $y_1, y_2, \ldots, y_t$。因為其屬於單向結構，所以不會考慮未來的信息，表示 y_t 只會受歷史數據 $x_1, x_2, \ldots, x_t$ 的影響。當其要追溯的歷史信息越久遠，其隱藏層就會越多，但卷積層數的增加會有梯度消失、訓練複雜和擬合效果不好的問題。圖 3 參考 Bai et al. (2018) 所繪製。

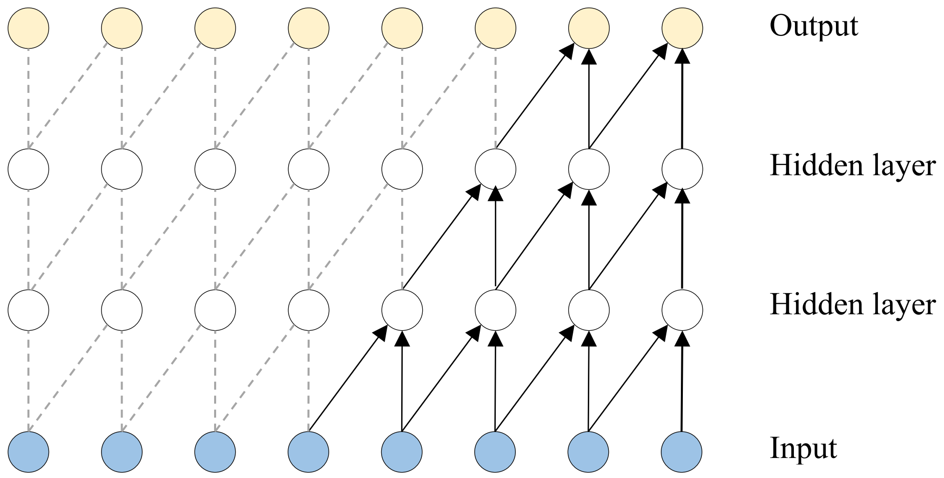

圖 3：因果卷積架構
原始出處：https://arxiv.org/pdf/1803.01271.pdf

(二)擴張卷積(Dilated Convolutions)

由於因果卷積會因歷史數據變多，導致卷積層數量隨著數據而增加，會有梯度消失、訓練複雜和擬合效果不好的問題。對於因果卷積，存在的一個問題是需要很多層或者很大的 filter 來增加卷積的感受野(Receptive Field Size)，因此使用擴張卷積，通過增加零來從原始 filter 中生成更大的 filter，使模型在層數不大的情況下有非常大的感受野，而感受野的增加，就可以看到更多的訊息。以公式表示，對於 1-D 序列輸入$x \in R^{\tilde{n}}$和過濾器 f：{0,…,k-1}→R，擴張卷積對序列的元素 s 的操作 F 定義為

$$F(s) = (x *_d f)(s) = \sum_{i=0}^{k-1} f(i) \cdot x_{s-d \cdot i}$$

其中 d 是擴張因子，k 是濾波器大小，s-d·i 解釋了過去資料搜尋的方向。當跳躍步數為 1(d = 1)時，擴張卷積簡化為一般正規卷積運算，使用更大的擴張可以使頂層的輸出能夠代表更廣泛的輸入範圍，從而有效地擴大 ConvNet 的感受野。圖 4 是參考 Bai et al. (2018) 所繪製。

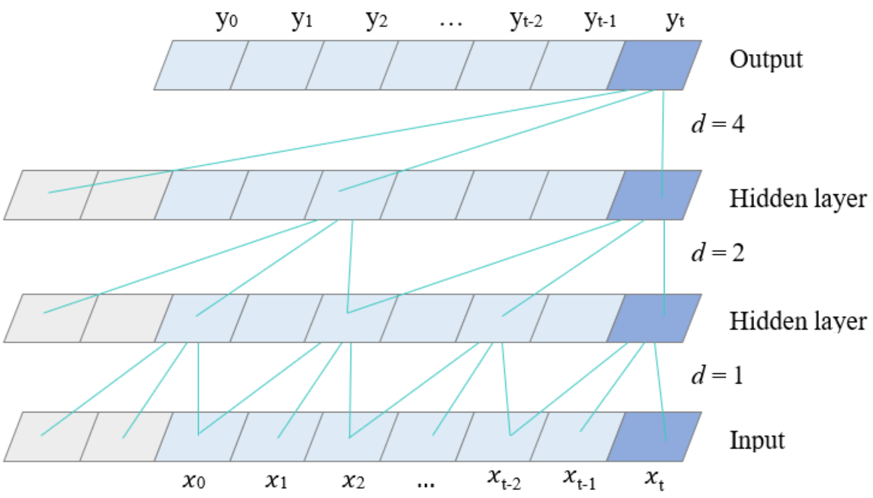

圖 4：擴張卷積架構

原始出處：https://arxiv.org/pdf/1803.01271.pdf

(三)殘差連接(Residual Connections)

隨著深度學習模型深度的增加，雖然可以進行更複雜的特徵選取，但也帶來了許多的問題，像是梯度消失等等問題，導致模型不容易訓練，因此可以使用殘差連接來解決。在殘差網路中，會將之前的信息和轉換後的信息，都作為當前的輸入，即使網路層數非常深，仍然能得到不錯的效果。其整個網路採用跳躍連線，以加快訓練過程，避免深度模型的梯度消失問題。圖 5 參考 Bai et al. (2018) 所繪製。

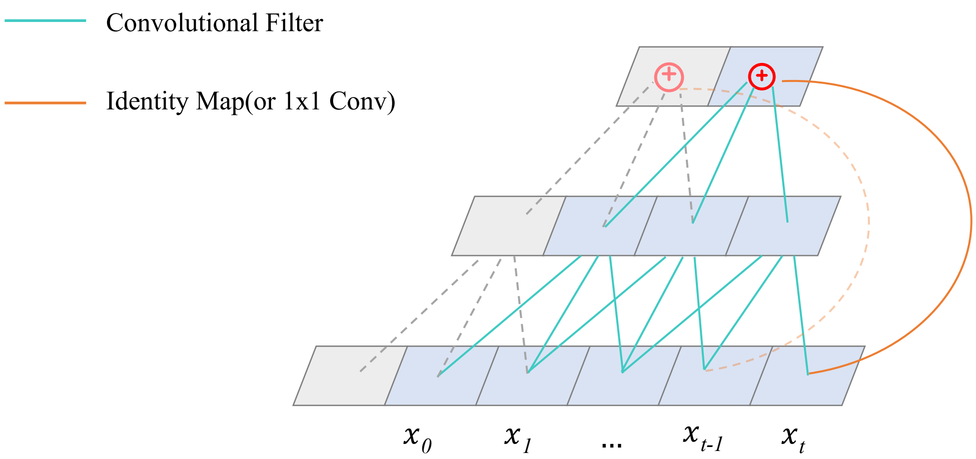

圖 5：殘接連接架構

原始出處：https://arxiv.org/pdf/1803.01271.pdf

二、長短期記憶 (Long Short-Term Memory, LSTM)

雖然 RNN 在長期依賴方面也有不錯的表現，但它有梯度消失的問題，為了解決這個問題，Hochreiter & Schmidhuber (1997)提出了(Long Short-Term Memory;

LSTM)。是一種特殊的 RNN，能夠學習長期依賴關係，主要由四個元件組成：輸入門(Input gate) 、記憶單元(Memory Cell)、輸出門(Output gate)和遺忘門(Forget gate)，透過三個門控來決定資料是否要保存和丟棄。

圖 6 為 LSTM 結構圖，x_t 為當前輸入的特徵向量、f_t為遺忘門的向量、C_t為當前的狀態向量、i_t為輸入門的向量、$\tilde{C}_t$為輸入的狀態向量、O_t為輸出門的向量、h_t為輸出結果的向量、tanh 為激活函數用在了狀態和輸出上，是對數據的處理、σ為 Sigmoid 激活函數，會選用 sigmoid function 選用 sigmoid function 是因為它的值介於 0 到 1 之間，而這 0 到 1 的值代表說這個門的打開程度，如果 1 是代表打開 0 則反之。

每一個 LSTM 中的 memory 都存了一個值，輸入的 x_t 經過運算並通過遺忘門(f_t)來決定是哪些資訊會被保留或丟棄。而輸入門(i_t)控制是否要輸入本次的資訊和其輸入狀態，最後輸出門(O_t)將所有狀態和資訊結合，決定最終輸出的訊息和之後要傳遞的狀態。

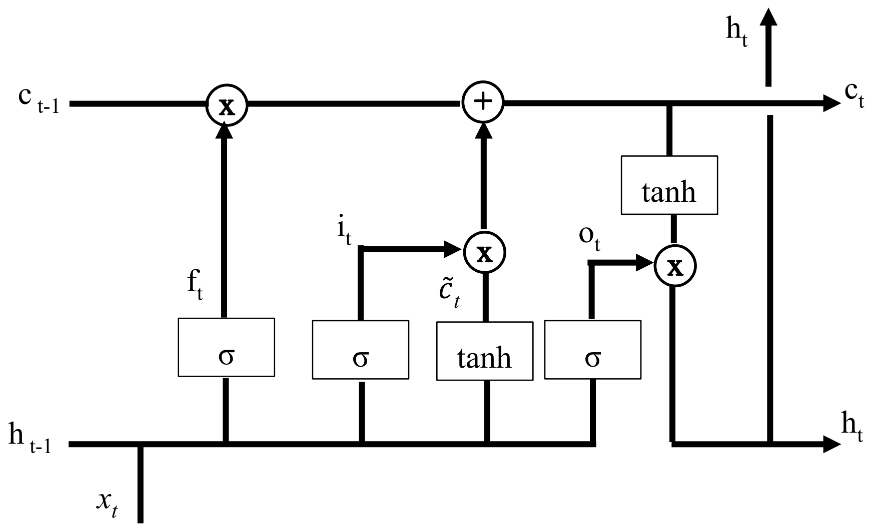

圖 6：長短期記憶結構圖

三、TCN-LSTM

TCN 在時間序列相關應用有良好的成果 (Zhu, Liao, & Wang 2020)，LSTM 在股票預測應用也有不錯的表現 (Jin, Yang, & Liu 2020)。有學者認為混合模型比單一模型有更好的預測結果 (Kumar & Haider 2021)。為了有效地提取特徵並提高預測精度，因此本研究將使用 TCN 結合 LSTM 來進行股票價格預測。該模型結構圖如

圖所示。主要結構為 TCN、LSTM，包含輸入層、TCN 層、LSTM 層和輸出層。首先利用 TCN 模型從輸入的原始時間序列數據中提取深度特徵向量。然後，使用 LSTM 模型從由深度特徵向量構建的新時間序列數據中學習時間特徵。最後使用全連接層組成的 Dense 模型來執行預測任務，本研究的時間窗格設置為

10 天，模型相關設定參考 (Lu et al. 2020)，將 CNN 層改成 TCN 層 Units 一樣設置隱藏層為一層 64 個神經元，由於 LSTM 預測會有預測延遲一天的問題，因此在第二層 LSTM 層設置的神經元和時間窗格使用同樣大小來解決預測延遲的問題，因此將 LSTM 層 Units 更改設置隱藏層為一層 10 個神經元。最後所有模型皆設置一層為一個神經元的全連接層作為輸出層，輸出的結果為模型所預測的收盤價格。

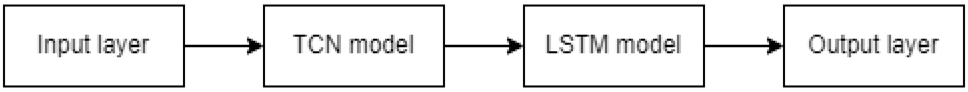

圖 7：TCN-LSTM 模型結構圖

肆、 實驗設計

一、實驗準備

(一)系統環境

實驗環境是使用 Anaconda 3，是 Python 及 R 語言的一個免費開源版本的開發平台，主要用於資料科學(Data Science)、機器學習(Machine Learning)和預測分析(Predictive Analytics)等等。支援多種套件包括 Conda、Python 以及許多安裝好的工具包。本研究使用 Anaconda 3 建立了一個 Tensorflow GPU 版本並結合 Keras 且 Python 版本為 3.7 的開發環境，並使用 Jupyter Notebook 執行和編輯程式，接下來引入或安裝所需的套件，如 Numpy(提供大量的數學函式庫)、Pandas(用來對資料進行重構、切割、聚合及選擇子集合等操作)、Matplotlib(進行圖形繪製)等等，最後透過該環境進行建模與實驗設計。

(二)數據蒐集

數據集除了挑選 2021 年度前 10 大權值股中的股票以外，還挑選了 2021 年度價格漲幅較高且當沖量較大的台股數據集作為研究，產業別為電子股和傳統產業，分別為台積電、智原、長榮和長榮航。選擇權值股是因為它和台股指數連動性較高。而台積電是台灣加權指數最具代表性的標的，智原、長榮和長榮航則是 2021 年度價格漲幅較高且當沖量較大，再加上其股價和三大法人買賣超連動性較高。本研究的數據是透過爬蟲技術從臺灣證券交易所取得的，該網站擁有台灣金融市場的資訊，包含股票、債券、權證和基金等等數據。表 3 為股票數據集欄位內容，但每支股票會因其上市上櫃時間不同，導致歷史數據長度會有所不同，因此本研究將資料範圍統一使用近五年的資料來進行研究，時間從 2017 年 1 月 1 日至 2021 年 12 月 31 日。

表 3：數據集描述

特徵名稱	描述
日期	時間日期
成交股數	當日所有交易的成交股數總和
成交金額	當日所有交易的成交金額總和
開盤價	當日開市時股票的價格
最高價	當日交易中最高的價格
最低價	當日交易中最低的價格
收盤價	當日收盤後的價格
漲跌價差	當日收盤價格和前日收盤價格差額
成交筆數	當日所有交易的成交筆數總和
三大法人買賣超股數	當日三大法人買賣股數總和

二、資料預處理

在模型的訓練中，要將時間的資料格式轉換成特徵值並輸入到模型訓練相對是複雜的，為了避免讓模型訓練變得更加複雜，因此將時間資料轉換成資料索引來進行排序，再將時間相關欄位資料移除。依照其排序讓模型能夠每日地輸入新數據並反覆進行模型的訓練和調整。

剩餘的資料有成交股數、成交金額、開盤價、最高價、最低價、收盤價、漲跌價差、成交筆數、三大法人買賣超股數。首先移除成交金額和成交筆數的資料欄位，根據過去的研究發現這兩個欄位對於股票價格預測的效益並不大，接下來再移除漲跌價差，雖然其漲跌幅度容易發現價格趨勢，但容易產生過度擬合(Over fitting)的情況發生。本研究將收盤價取出並進行標記，使模型在訓練過程中反覆調整至最佳的參數，將所有的資料使用最小最大正規化法方法(Min-Max Normalization)，將資料縮放至 0 和 1 之間，使輸入的特徵都能有所貢獻，讓其梯度下降減少訓練的時間，以提高模型的準確度。

(一)資料整理與標準化

良好的資料處理，可以使模型不會因為過多的雜訊，導致模型訓練效果不好。由於時間資料格式會增加模型的複雜度，因此將時間資料轉換成資料索引來進行排序，並移除該欄位，將成交金額和成交筆數不重要欄位移除，漲跌幅可能讓模型過度擬合，也將其移除。最後將特徵縮放避免數值過大而影響訓練時間和增加複雜度。在數據集的部分根據(Lu et al. 2020)使用八比二法則將其分為訓練集和驗證集，為了驗證模型未來可行性並添加了測試集。實際資料切分情形如表 4 所示。

表 4：本研究實驗數據集

數據集名稱	資料總筆數	訓練集筆數	驗證集筆數	測試集筆數
台積電	1193	877	219	97
智原	1193	877	219	97
長榮	1193	877	219	97
長榮航	1193	877	219	97

最後將所有資料使用最小最大正規化方法(Min-Max Normalization)(公式 1)，X為該筆資料，x_{min}為資料集裡最小的數值，x_{max}為資料集裡最大的數值。所有資料等比例縮放數值至 0 和 1 之間，再將縮放完的資料輸入模型進行訓練。使用特徵縮放主要是為了讓模型訓練更加容易，且不會因其數值過大或過小而有所影響，而導致預測失準。

$$x_{norm} = \frac{x - x_{min}}{x_{max} - x_{min}} \in [0,1] \qquad \text{(公式 1)}$$

三、模型相關參數設置

本研究使用 Tensorflow 和 Keras 的函式庫建立 TCN-LSTM、CNN-LSTM、RNN-LSTM 和 LSTM 四種模型。各個模型的輸入層皆輸入六個特徵值，模型相關設定參考 (Lu et al. 2020)，模型參數與設置如表 5，Epoch 設為 100、Batch size 改為 256、Optimizer 使用 adam、kernel_initializer 使用 RandomNormal、activation 使用 relu，其中 N 代表一次輸入的天數，本研究使用 10 天。所有模型第一層 Units 皆設置隱藏層為一層 64 個神經元，由於 LSTM 預測會有預測延遲一天的問題，因此在第二層 LSTM 層設置的神經元和時間窗格使用同樣大小來解決預測延遲的問題，因此第二層 Units 都設置隱藏層為一層 10 個神經元，所有模型都加入一個平坦層(Flatten layer)將多維轉為一維，最後所有模型皆設置一層為一個神經元的全連接層作為輸出層，輸出的結果為模型所預測的收盤價格。

表 5：模型參數設定

參數	說明
Feature	N(每日收盤價格、每日開盤價格、每日最高價格、每日最低價格、每日成交量、三大法人買賣超)
Label	(N+1)每日收盤價格
Epoch	100
Batch size	256
Optimizer	adam
TCN-LSTM、CNN-LSTM、RNN-LSTM、LSTM	input_shape(N,6), return_sequences =True, Units=64
TCN-LSTM、CNN-LSTM、RNN-LSTM、LSTM	Units=10
Flatten	平坦層:將多維陣列轉換為一維陣列
Dense	Units=1, kernel_initializer='RandomNormal', activation='relu'

四、評估方法

為了良好的比較各個不同模型之間預測性能的好壞，因此參考 Lu et al. (2020) 所使用的損失函數來進行比較。利用了 3 個最常見的度量標準，即平均絕對誤差（MAE）、平均絕對百分比誤差(MAPE)和均方根誤差（RMSE）來比較。

(一)平均絕對誤差(Mean Absolute Error; MAE)

平均絕對誤差(MAE)又被稱為 L1 損失或 L1 範數損失，是實際值和預測值之差的絕對值之和，能更好地反映預測值誤差的實際情況。y_i^p是預測值，y_i^r是實際值，n 為樣本數，數值越小代表平均誤差越小，如(公式 2 所示。

$$MAE = \frac{\sum_{i=1}^{n} |y_i^p - y_i^r|}{n} \qquad \text{(公式 2)}$$

(二)平均絕對百分比誤差(Mean Absolute Percentage Error; MAPE)

平均絕對百分比誤差(MAPE)可以用來衡量一個模型預測結果的好壞，為每個絕對誤差的和除以實際值，也就是誤差百分率的平均值。數值越小代表平均預測的準確性越好，如(公式 3 所示。

$$MAPE = \frac{100}{n} \sum_{i=1}^{n} \frac{|y_i^r - y_i^p|}{y_i^r} \qquad \text{(公式 3)}$$

(三)均方根誤差(Root Mean Squared Error; RMSE)

均方根誤差(RMSE)，又稱標準誤差，表示預測值與真實值的平均偏離程度，實際值和預測值的差，然後平方後的總和，再取平均值，最後再取根號，數值越小代表預測的準確性越好，如(公式 4 所示。

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n} (y_i^r - y_i^p)^2}{n}} \qquad \text{(公式 4)}$$

伍、當沖交易測試與實驗結果

過去部分研究採用當沖交易實驗以測試預測模型的準確性，同時證實交易策略獲利的可行性。舉例來說，(Bitvai & Cohn 2015) 利用技術分析開發了一種以線性模型(linear model)為形式的新型隨機交易(stochastic trading)演算法幫助當沖交易者做出更明確的預測。此外，還提出了多任務學習的概念提高了預測的精確度，進一步增加了當沖交易的潛在收益；在國內交易的部分，游騰芳、何怡滿(2022)探討了在臺灣上市股票在收盤漲停的次日，實施先賣後買的當沖交易策略，將不同因素納入考慮，包括產業分類、漲停前股價階段、當沖成交比重以及成交量是否突破新高，研究結果顯示，在考慮交易成本後的勝率為 52%、平均淨報酬率為 0.24%、獲利平均報酬率高於虧損報酬率、盈虧比為 1.06。其策略即使在考慮交易成本的情況下，仍可以獲利。

在當沖交易實證的部分，除了使用過去數據集進行回測之外，還根據本研究所提出的方法制定了一個實際操作方法，使用當沖的交易來進行驗證，當沖交易被稱為無本交易，之所以會選用當沖來進行實驗是因為其成本相對比隔日沖或是波段操作低且不需要先付龐大的交割金，若收益為負的才需要附交割金。本研究所有模型都設定 10 天的時間窗格，之所以不選擇 1 天或是 5 天的時間窗格，因為在本研究實驗環境下 1 天時間窗格雖然評估指標數值都非常小，但將預測值和實際值作比對發現會有延遲一天的問題，而 5 天時間窗格的評估指標數值相對比 10 天時間窗格來的高。

模型訓練完成後為了驗證模型未來的可行性和本研究所制定的實際操作方法並每天更新數據，因此當沖交易中所選取的時間為 2021 年 12 月到 2022 年 4 月共 97 個交易日，為了進行當沖操作，因此將 Label 分別標記為最高價、最低價、開盤價和收盤價，來進行最高、最低、開盤和收盤價的預測。透過 5 年的歷史數據來學習個股特性，並運用前 10 天的數值預測明天的數值。當沖交易的操作規則如下所述，首先根據明天預測的收盤價判斷要做多還是做空，做多和做空的操作規則，詳細操作規則如表 6 和表 7 所示。無論做多或做空都須符合三個條件，第一步都需先排除異常值，因預測出來的最高價可能比預測的收盤價、最低價和開盤價來的低是不合理的，預測出來的最低價可能比預測的收盤價、最高價和開盤價來的高也是不合理的，接下來根據操作規則來判斷最後操作與否。在做多時要符合開低走高，當預測收盤價大於昨日收盤價和預測開盤價符合規則二趨勢向上；當預測開盤價和實際開盤價小於昨日收盤價符合規則三開低。在做空時要開高走低，當預測開盤價和實際開盤價小於昨日收盤價符合規則二趨勢向下，當預測開盤價大於昨日收盤價符合規則三開高。在收益計算的部分，個股操作的單位數都為一張，而下單手續費採用券商給予的折扣為 2.8 折，詳細的計算方式會在後續來作說明。本研究假設無論是做多或做空都能以實際開盤的價格進行買賣，根據表 8 和表 9，在做多的部分，實際操作時在開盤前掛單以漲停價買進。在做多時會先以預測的最高價進行賣出在做多時賣出價格先以預測的最高價若收盤前 30 分鐘沒有成功賣出會以預測的收盤價來進行賣出，若都沒有成功賣出會在收盤前 5 分鐘以跌停價賣出，也就是最後以實際收盤價賣出；在做空的部分，實際操作時在開盤前掛單以跌停價賣出。在做空時會先以預測的最低價買進，若收盤前 30 分鐘沒有成功買進會以預測的收盤價來進行買進，若都沒有成功買進會在收盤前 5 分鐘以漲停價買進，也就是最後以實際收盤價買進。

表 6：做多的操作規則

規則	說明		
規則一	預測最高價	>	預測的收盤價、開盤價和最低價
規則二	預測收盤價	>	昨日收盤價和預測開盤價
規則三	預測開盤價	<	昨日收盤價

表 7：做空的操作規則

規則	說明		
規則一	預測最低價	<	預測的收盤價、開盤價和最高價
規則二	預測收盤價	<	昨日收盤價和預測開盤價
規則三	預測開盤價	>	昨日收盤價

表 8：當沖做多的買賣時機

買賣時機	操作方式
買進時機	開盤前掛單以漲停價買進。
賣出時機	1.以預測的最高價進行賣出。 2.若收盤前 30 分鐘沒有成功賣出會以預測的收盤價來進行賣出。 3.若都沒有成功賣出會在收盤前 5 分鐘以跌停價賣出。

表 9：當沖做空的買賣時機

買賣時機	操作方式
賣出時機	開盤前掛單以跌停價賣出。
買進時機	1.以預測的最低價進行買進。 2.若收盤前 30 分鐘沒有成功買進會以預測的收盤價來進行買進。 3.若都沒有成功買進會在收盤前 5 分鐘以漲停價買進。

一、台積電數據集實驗結果

表 10 為台積電驗證集實驗結果，在 10 天時間窗格的實驗下，所得出的三項評估指標(MAE、RMSE、MAPE)，從評估指標發現本研究所提出的 TCN-LSTM 模型表現均優於另外三種模型。

表 10：台積電數據集 各模型評估指標結果

Model	MAE	RMSE	MAPE
TCN-LSTM	0.0327	0.0447	3.2568
RNN-LSTM	0.1313	0.1773	12.9263
CNN-LSTM	0.1319	0.1805	12.9757
LSTM	0.1319	0.1799	12.9854

圖 8 至圖 11 為台積電數據集在 10 天時間窗格下的驗證集價格預測趨勢圖，趨勢圖上 X 軸顯示到 250 是因為一格的單位為 50，但實際去對照圖上的趨勢線其實不到 250 筆只有 219 筆，可以看到各個模型在 10 天時間窗格下驗證集的預測程度。從趨勢圖可以發現 RNN-LSTM、CNN-LSTM 和 LSTM 模型在 0-50 天那段大漲大跌時，預測值和實際的價格偏差較大，而 TCN-LSTM 優於其他三種模型，這也和前面所獲得的三種評估指標有一致的結果。因此我們可以得到一個假設，當價格趨勢發生較大的波動變化，本研究所提出的 TCN-LSTM 模型相較於其他三種模型更能夠跟上實際的價格趨勢。

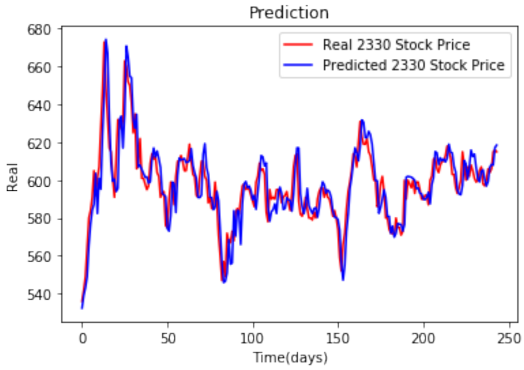

圖 8：TCN-LSTM 台積電價格趨勢圖

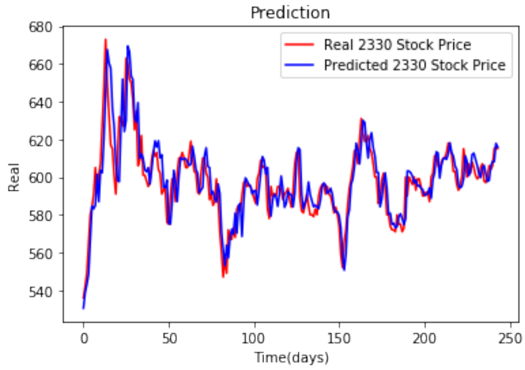

圖 9：RNN-LSTM 台積電價格趨勢圖

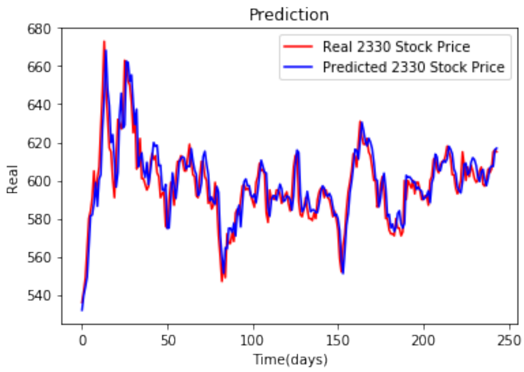

圖 10：CNN-LSTM 台積電價格趨勢圖

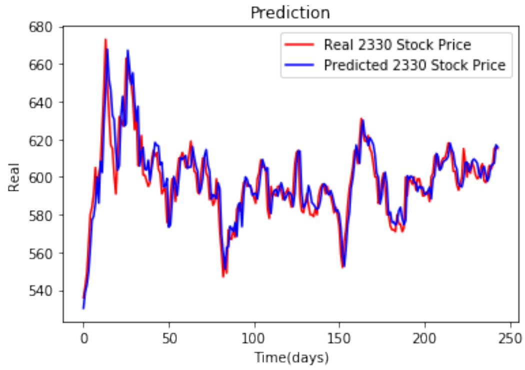

圖 11：LSTM 台積電價格趨勢圖

台積電數據集的當沖操作策略，首先會根據預測的收盤價判斷明日漲或跌來決定要做多或是做空，接下來根據本研究所制定的操作規則來決定操作與否，詳細操作規則如表 6 和表 7 所示，需符合特定條件時才進行交易，以降低交易時產生的風險。對於當沖客而言，每個交易日頻繁買賣將導致成本高昂且難以獲利，同時需要花費大量時間盯盤。然而，我們之所以採用當沖交易進行實驗，是因為相較於隔日沖或波段操作，當沖交易的成本相對較低且不需要支付龐大的交割金。

從 110 年 12 月 1 日至 111 年 4 月 29 日的 97 個交易日中，總共操作了 9 次，總收益為 22250 元，在做多的部分操作了 6 次，獲利的次數有 4 次，獲利的成功率大約有 66%；在做空的部分操作了 3 次，獲利次數為 2 次，獲利的成功率大約有 66%。在收益計算的部分，表 11 為買賣手續費和證交稅的計算方式。表 12 為當沖收益試算表，在 110 年 12 月 1 日根據操作規則要做多，接下來根據(公式 5 以開盤價 596 買進，買進手續費為 237 元，用預測最高價賣出但由於該值大於實際最高價，因此改用預測收盤價 600 順利賣出，賣出手續費和證交稅分別為 239 元和 900 元，收益為 2624；在 110 年 12 月 3 日根據操作規則要做空，根據(公式 5 以開盤價 614 賣出，賣出手續費和證交稅分別為 244 元和 921 元，接下來用預測最低價 613 順利買進，買進手續費為 244 元，收益為-409。根據結果發現收益的部分都是大賺小賠，因此得到一個結論本研究所提出的模型在制定的操作規則具有良好的績效。

收益=賣出價金-賣出手續費-證交稅-買進價金-買進手續費 (公式 5)

表 11：手續費和證交稅計算方式

項目	計算方式
買進手續費	買進價金*0.001425*折扣(預設券商提供 2.8 折)
賣出手續費	賣出價金*0.001425*折扣(預設券商提供 2.8 折)
證交稅	賣出價金*0.0015

表 12：當沖收益試算表

做多				做空			
	賣出	買入	收益		賣出	買入	收益
110/12/1	—	—	—	**110/12/23**	—	—	—
每股價格	600	596	—	**每股價格**	614	613	—
股數	1000	1000	—	**股數**	1000	1000	—
價金	600000	596000	—	**價金**	614000	613000	—
手續費	-239	237	—	**手續費**	-244	244	—
證交稅	-900	—	—	**證交稅**	-921	—	—
總計	598861	596237	2624	**總計**	612835	613244	-409

二、智原數據集實驗結果

表 2 為智原驗證集實驗結果，在 10 天時間窗格的實驗下，從所得出的三項評估指標(MAE、RMSE、MAPE)，本研究所提出的 TCN-LSTM 模型表現均優於另外三種模型。

表 2：智原數據集 各模型評估指標結果

Model	MAE	RMSE	MAPE
TCN-LSTM	<u>0.034</u>	<u>0.0538</u>	<u>3.3679</u>
RNN-LSTM	0.133	0.1983	12.8483
CNN-LSTM	0.1322	0.201	12.7457
LSTM	0.1356	0.2046	13.1515

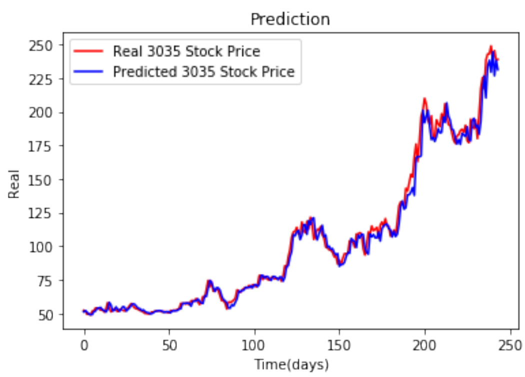

圖 12：TCN-LSTM 智原價格趨勢圖

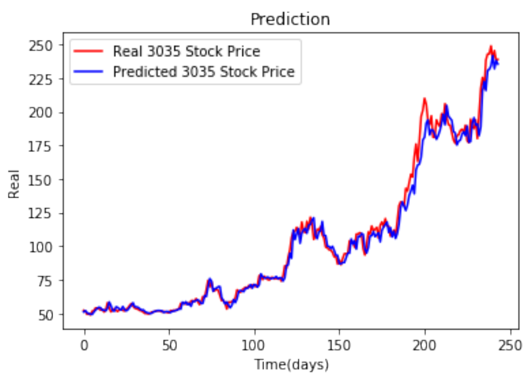

圖 13：RNN-LSTM 智原價格趨勢圖

圖 12 至圖 15 為智原數據集在 10 天時間窗格下驗證集價格預測趨勢圖，趨勢圖上 X 軸顯示到 250 是因為一格的單位為 50，但實際去對照圖上的趨勢線其實不到 250 筆只有 219 筆，可以看到各個模型在驗證集的預測程度。從趨勢圖

可以發現 TCN-LSTM 在後半段相較於其他三個模型，所預測的價格較能貼近真實值。

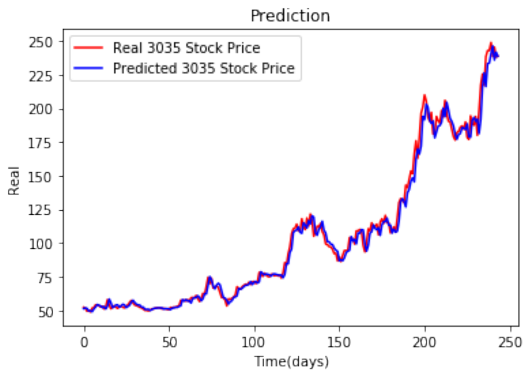

圖 14：CNN-LSTM 智原價格趨勢圖

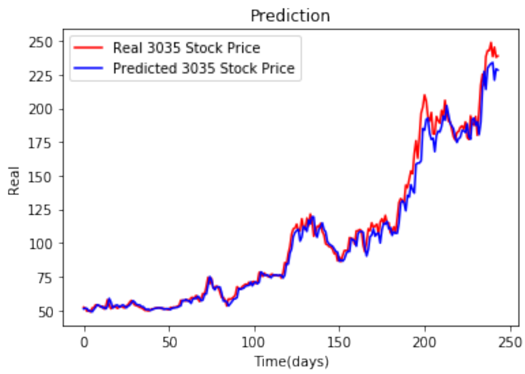

圖 15：LSTM 智原價格趨勢圖

智原數據集的當沖操作策略，首先會根據預測的收盤價判斷明日漲或跌來決定要做多或是做空，接下來根據本研究所制定的操作規則來決定操作與否，從 110 年 12 月 1 日至 111 年 4 月 29 日的 97 個交易日中，總共操作了 14 次，總收益為 199314 元，在做多的部分操作了 10 次，獲利的次數有 9 次，獲利的成功率大約有 90%；在做空的部分操作了 4 次，獲利次數為 4 次，獲利的成功率為 100%。在損益的部分也都是大賺小賠，在智原數據集中本研究所提出的模型在制定的操作規則也具有良好的績效。

三、長榮數據集實驗結果

表 14 為長榮驗證集實驗結果，在 10 天時間窗格的實驗下，從所得出的三項評估指標(MAE、RMSE、MAPE)，本研究所提出的 TCN-LSTM 模型表現均優於另外三種模型。

表 14：長榮數據集 各模型評估指標結果

Model	MAE	RMSE	MAPE
TCN-LSTM	0.0379	0.0647	3.7548
RNN-LSTM	0.1302	0.1983	12.5453
CNN-LSTM	0.1351	0.2059	13.0179
LSTM	0.1385	0.2081	13.3904

錯誤! 找不到參照來源。至**錯誤! 找不到參照來源。**為長榮數據集在 10 天時間窗格下驗證集價格預測趨勢圖，以看到各個模型在驗證集的預測程度。從趨勢圖可以發現在接近 100-150 天時，RNN-LSTM、CNN-LSTM 和 LSTM 等三種模型相較於 TCN-LSTM 模型，所預測的價格和真實價格有較大的偏差。

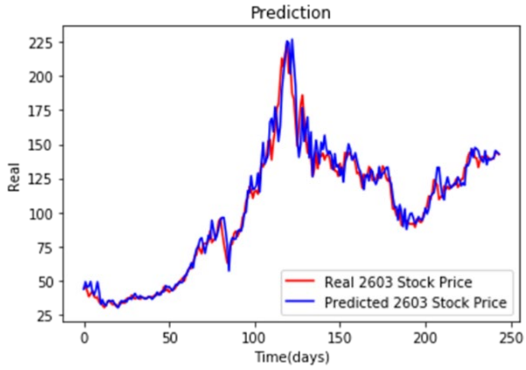

圖 16：TCN-LSTM 長榮價格趨勢圖

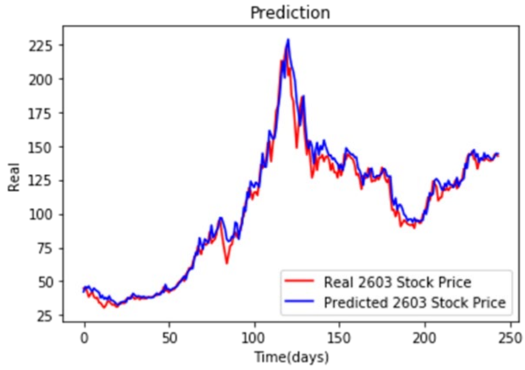

圖 17：RNN-LSTM 長榮價格趨勢圖

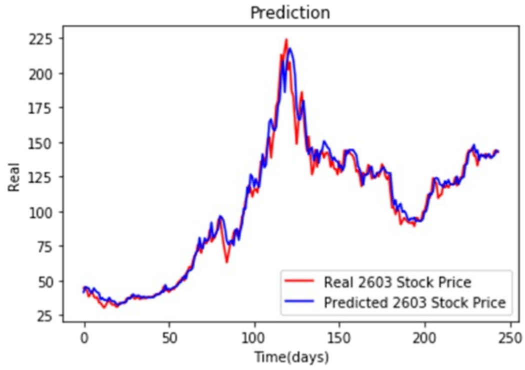

圖 18：CNN-LSTM 長榮價格趨勢圖

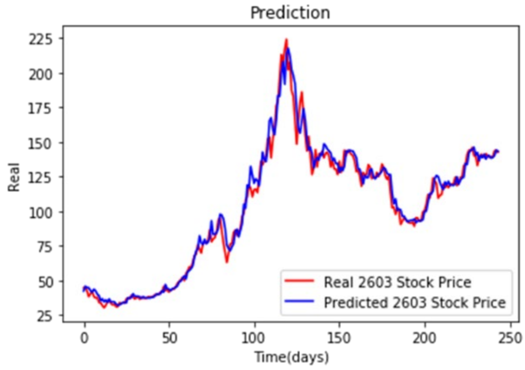

圖 19：LSTM 長榮價格趨勢圖

長榮數據集在 97 個交易日中，總共操作了 16 次，總收益為 34405 元，在做多的部分操作了 6 次，獲利的次數有 6 次，獲利的成功率為 100%；在做空的部分操作了 10 次，獲利次數為 8 次，獲利的成功率大約有 80%。在損益的部分也都是大賺小賠，在長榮數據集中本研究所提出的模型在制定的操作規則也具有良好的績效。

四、長榮航數據集實驗結果

表 35 為長榮航驗證集實驗結果，在 10 天時間窗格的實驗下，從所得出的三項評估指標(MAE、RMSE、MAPE)，本研究所提出的 TCN-LSTM 模型表現均優於另外三種模型。

表 35：長榮航數據集 各模型評估指標結果

Model	MAE	RMSE	MAPE
TCN-LSTM	0.0374	0.0739	3.6668
RNN-LSTM	0.1346	0.2025	13.0161
CNN-LSTM	0.1409	0.2119	13.6154
LSTM	0.1439	0.2169	13.9023

錯誤! 找不到參照來源。至**錯誤! 找不到參照來源。**為長榮航數據集在 10 天時間窗格下驗證集價格預測的趨勢圖，可以看到各個模型在驗證集的預測程度。從趨勢圖可以發現 TCN-LSTM 相較於其他三個模型，所預測的價格較能貼近真實值，尤其是在 200-250 天的時間區段。

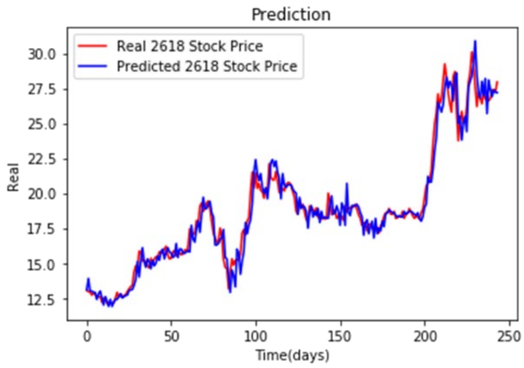

圖 20：TCN-LSTM 長榮航價格趨勢圖

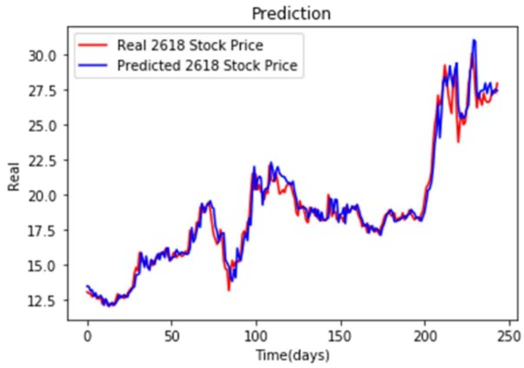

圖 21：RNN-LSTM 長榮航價格趨勢圖

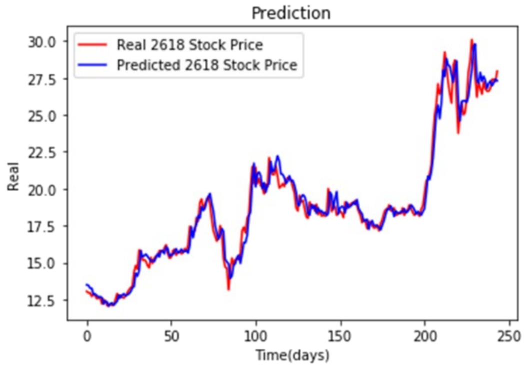

圖 22：CNN-LSTM 長榮航價格趨勢圖

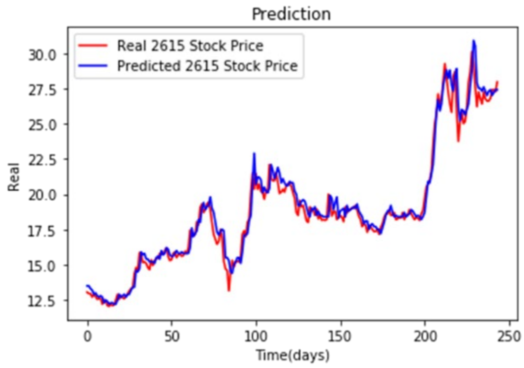

圖 23：LSTM 長榮航價格趨勢圖

長榮航數據集在 97 個交易日中，總共操作了 13 次，總收益為 11616 元，在做多的部分操作了 6 次，獲利的次數有 6 次，獲利的成功率大約有 91%；在做空的部分操作了 7 次，獲利次數為 6 次，獲利的成功率大約有 86%。在損益的部分也都是大賺小賠，在長榮航數據集中本研究所提出的模型在制定的操作規則也具有良好的績效。

陸、結論

一、結論與討論

現有的股票預測方法大致可分為基本面分析、技術面分析和時間序列分析 (Devadoss & Ligori 2013)，過去在這個主題的研究大多數使用技術分析和時間序列分析，通過分析市場活動、過去價格和交易量產生的統計數據來評估股票。尋找影響股票價格走勢的峰值、底部、趨勢和其他因素。股票未來的價格通常取決於過去值和其他相關變量的過去值。但通常用來預測較短期的時間，而本研究運用深度學習方法結合歷史價格數據在短時間的預測也和 Shen & Shafiq (2020) 學者的研究得到一致的結果都具有良好的性能。基本面分析假設股票價格（當前和未來）取決於其內在價值。為了找到特定股票的內在價值，需要檢查股票當前和未來的整體健康狀況以及經濟狀況。基本面分析的優勢在於其系統化的方法和在圖表出現之前預測變化的能力，基本面分析是長期穩定和增長的一種優越方法 (Agrawal et al. 2013)。Huang, Capretz, & Ho (2021) 基於基本面分析來預測股票的

長期表現。並嘗試了三種不同的機器學習方法來預測長期股票表現，其結果表明適合提供投資者進行長時間操作策略。無論是技術面分析、基本面分析和時間序列分析的方法都有其自身的局限性，未來可以嘗試將方法進行結合。

時間序列數據是在特定時間間隔內獲取的數據，在觀察期間可能會發生巨大變化，因此變得高度非線性。因此要準確地預測非常具有挑戰性，為了準確預測時間序列數據，許多研究人員使用不同的深度學習技術也獲得了不錯的成果。Hu, Zhao, & Khushi (2021) 使用深度學習進行外匯和股票價格預測的調查，結果表明使用基於深度學習的方法進行股票預測的趨勢呈指數增長。在時間窗格的選擇上，本研究在使用 1 天時間窗格時，雖然其誤差評估指標很小，但將預測價格趨勢圖和數值作比對發現只是在模仿前一天的數值，有學者表明根據當天價格預測第二天收盤價是不合適的 (Hota et al. 2017)，推測可能在股票價格預測領域使用 1 天時間窗格，因價格變化不夠大等等問題會學習不到趨勢，因此本研究根據 Chou & Nguyen (2018) 的研究使用 10 天的時間窗格去進行實驗。

本研究參考 Selvin et al. (2017) 所提出的 SLIDING WINDOW 方法，本研究使用不同的時間窗格、模型參數、數據集和增加了更多的特徵，本研究所提出的方法除了和其所使用的單一方法比較之外，還和其他結合 LSTM 的混合方法比較且都獲得較好的結果。數據集除了挑選 2021 年度前 10 大權值股中的股票以外，還挑選了 2021 年度價格漲幅較高且當沖量較大的台股數據集作為研究，統一使用 10 天的時間窗格法並比較不同模型之間的效能，從實驗結果的趨勢圖發現 TCN-LSTM 模型相較於 RNN-LSTM、CNN-LSTM 和 LSTM 模型這三種模型來的好，而在三種評估指標也都呈現較好的預測結果。過去研究大多數研究都是針對美國股市進行預測和對數據集進行回測並無提出一個實際操作的方法，而本研究則針對台灣股市進行預測，所提出的方法在歷史回測都比單一方法和其他混合方法都有較好的預測結果，還制定了一個實際可以操作的方法並將其運用在當沖交易中。

從驗證集的價格預測趨勢圖發現，RNN-LSTM、CNN-LSTM 和 LSTM 模型在遇到價格趨勢大幅度波動時，所預測的結果與真實值差距較大。而 TCN-LSTM 模型在價格趨勢大幅度波動時，預測結果相對和真實值差距較小，因此相較於其他模型，TCN-LSTM 模型較能抓出更多的隱藏訊息。

在實際操作的部分，操作單位數量都以一張為單位，而本研究制定了一個操作規則並運用當沖的方式來進行驗證，結果如表 16 所示，在台積電數據集獲利的成功率大約有 70%；在智原數據集中獲利的成功率大約有 93%；在長榮數據集中獲利的成功率大約有 88%；在長榮航數據集中獲利的成功率大約有 92%。在損益的部分也都是大賺小賠，由此可知，本研究所提出的模型在當沖交易中獲得了一個不錯的成效且在收益的部分也都是獲利的。

根據預計收益獲利最多的是智原數據集，雖然台積電股價是所有數據集中最高的，但由於其股性相對沒這麼活潑，使得預計收益和獲利成功率相對都比當沖量較大的智原低。根據結果我們可以得到一個結論是使用本研究所提出的模型和

當沖交易操作方式，較適合運用在股性相對較活潑且當沖量較大的個股。

雖然本研究成效顯著，但仍不鼓勵投資人從事台股當沖交易。在 Barber et al. (2004) 的研究中發現，當沖交易在台灣股市中佔據著相當大的份額，並且絕大多數參與者都是個人投資者，通過調查數據顯示，個人當沖交易者的盈虧表現受到多種因素影響，包括交易頻率、交易成本和持倉時間等。值得注意的是，大多數個人當沖交易者的表現不如市場平均水平，其表現與其交易頻率呈負相關，這可能是由於頻繁進行當沖交易帶來的高交易成本，可能會使部份當沖交易者不足以支付交易成本；Newall & Weiss-Cohen (2022) 強調投資與賭博之間的相似之處，提出投資者可能會因為高頻股票交易、過度自信、過度樂觀、缺乏知識與市場波動等因素遭受虧損。總體來說，投資者應該謹慎選擇策略，確保在不斷變化的金融環境中，知識、謹慎和風險管理仍然是投資者獲利的關鍵。

表 16：個股當沖操作成功率

個股	獲利的成功率	盈虧比	測試集收益
台積電	70%	8.38	NT$23,912
智原	93%	1231.1	NT$161,143
長榮	88%	22.04	NT$34,405
長榮航	92%	26.14	NT$11,616

根據上述結果，本研究結論為：

1. 從實驗結果得知， TCN-LSTM 模型相較於 RNN-LSTM、CNN-LSTM 和 LSTM 三種模型更加適合運用在時間序列的價格預測上，當價格有大幅度的波動時，根據預測結果發現本研究所提出的模型，預測值更能接近真實值。
2. 本研究所提出的模型，在制定的交易操作中具有不錯的成效，在所選擇的個股中操作的成功率都在 70%以上，在損益的部分也都是大賺小賠。
3. 本研究模型的盈虧比高(總獲利金額除以總虧損金額)，代表獲利遠高於虧損金額，雖交易次數較少，但相較於過去的研究成果表現佳（Bitvai & Cohn 2015; 游騰芳、何怡滿 2022)。

二、 研究限制與未來研究方向

(一)研究限制與建議

1. 單一的金融商品

本研究為了確保在受控制的環境下進行，只收集股票來作為實驗的數據集，也因此無法探討到像是基金或是外匯等金融商品，因此僅能代表特定的金融商品。

2. 實驗數據集較為侷限

本研究的數據集僅來自台灣股票市場中的個股，因此其研究結果僅能反映特定國家的股票，而不同國家影響股票價格的因素可能有所不同，建議未來的研究可以加入不同國家的股票數據集來進行探討。

3. 當沖交易實際操作的限制

首先要先符合以下三個條件：開戶滿三個月（不限單一券商）、近一年內的買賣成交筆數達 10 筆以上和簽署相關的「風險預告書」和「應付當日沖銷券差有價證券借貸契約書」。並不是所有的股票都能進行當沖交易，當個股遇到股東常會、分配收益、除息、除權息或是被列為處置股票都會導致無法進行當沖的交易。在選擇當沖標的也必須找成交量至少大於 1000 的個股來進行操作才比較不會有當沖沖不掉的問題出現，而導致需要付巨額的交割金。

4. 交易頻率較低

因為交易策略限制較多，雖本研究的交易策略成功率高，但交易頻率較低，不適合經常從事當沖交易者，若欲經常從事當沖交易且又期望提高成功率，建議搭配其它不同的交易策略，在不同的情況下採用不同的交易策略，方可同時提高勝率和交易頻率。

(二)未來研究方向

無論是全球性疫情災難、戰爭或是通貨膨脹等等因數，都會導致全球金融市場受到大幅度的影響，但是台灣某些個股卻呈現高幅度的逆勢成長，雖然本研究所提出的 TCN-LSTM 在預測價趨勢波動較大時得到良好的成果，在制定的操作規則也獲得了不錯的結果。但是距離準確的預測和交易操作應用上仍有發展的空間，未來可以加入其他關鍵的特徵數據或是方法上再進行改進，進而獲得更加準確的預測效果或是找出其他的交易操作方式以獲得更好的效益。

因此提供以下建議給未來研究此議題的學者：

1. 嘗試找出更多操作的規則並運用在實際交易操作中，如隔日沖或是波段操作等等。
2. 延續本研究的實驗結果加入或挖掘出其他關鍵的特徵，不單單只使用技術面的歷史價格資料也可結合基本面資料，如新聞資訊、財務報告等，進行更加準確的預測。
3. 結合或改善其他的資料預處理方法，如小波轉換等等，進行時間序列價格預測之研究。

參考文獻

台灣證券交易所 (2021)，證券經紀商投資人開戶統計表年報，https://www.twse.com.tw/zh/statistics/statisticsList?type=07&subType=261。

郝沛毅、歐仁彬、黃天受、林振穎、吳建生 (2018)，「透過新聞文章預測股價漲跌趨勢－結合情緒分析、主題模型與模糊支持向量機」，*資訊管理學報*，第二十五卷，第四期，頁 363-396。

陳振東、謝政翰 (2019)，「應用機器學習與模糊推論於股價漲跌預測之研究」，*資訊管理學報*，第二十六卷，第二期，頁 153-178。

游騰芳、何怡滿 (2022)，「台灣股市漲停次日之當沖績效分析」，*輔仁管理評論*，第二十九卷，第三期，頁 1-27。

郝沛毅、龔千芬 (2022)，「融合深度神經網路與深層模糊學生支持向量機於股價預測」，*資訊管理學報*，第二十九卷，第四期，頁 303-333。

Agrawal, J., Chourasia, V. S., & Mittra, A. K. (2013). State-of-the-art in stock prediction techniques. *International Journal of Advanced Research in Electrical, Electronics and Instrumentation Energy*, 2, 1360-1366.

Anuradha, J. (2021). Big data based stock trend prediction using deep CNN with reinforcement-LSTM model. *International Journal of System Assurance Engineering and Management*, 1-11.

Ashraf, B. N. (2020). Stock markets' reaction to COVID-19: Cases or fatalities? *Research in International Business and Finance*, 54, 101249.

Bai, S., Kolter, J. Z., & Koltun, V. (2018). An empirical evaluation of generic convolutional and recurrent networks for sequence modeling. *arXiv preprint arXiv:1803.01271*.

Barber, B. M., Lee, Y. T., Liu, Y. J., & Odean, T. (2004). Do individual day traders make money? Evidence from taiwan, Working Paper, UC-Davis, Davis, CA.

Bharambe, M. M. P. & Dharmadhikari, S. (2017). Stock market analysis based on artificial neural network with big data. *8th Post Graduate Conference For Information Technology*.

Bitvai, Z. & Cohn, T. (2015). Day trading profit maximization with multi-task learning and technical analysis. *Mach Learn 101,* 187-209. https://doi.org/10.1007/s10994-014-5480-x.

Chen, C. H., Kung, H. Y., & Hwang, F. J. (2019). Deep learning techniques for agronomy applications. *Agronomy*, 9(3), 142. https://www.mdpi.com/2073-4395/9/3/142.

Chen, Y., Fang, R., Liang, T., Sha, Z., Li, S., Yi, Y., Zhou, W., & Song, H. (2021). Stock price forecast based on CNN-BiLSTM-ECA model. *Scientific Programming*, 2021, 2446543. https://doi.org/10.1155/2021/2446543.

Chen, Y., Kang, Y., Chen, Y., & Wang, Z. (2020). Probabilistic forecasting with temporal convolutional neural network. *Neurocomputing*, 399, 491-501.

Chou, J. & Nguyen, T. (2018). Forward forecast of stock price using sliding-window metaheuristic-optimized machine-learning regression. *IEEE Transactions on Industrial Informatics*, 14(7), 3132-3142. https://doi.org/10.1109/TII.2018.2794389.

Deng, S., Zhang, N., Zhang, W., Chen, J., Pan, J. Z., & Chen, H. (2019). Knowledge-driven stock trend prediction and explanation via temporal convolutional network. *Companion Proceedings of The 2019 World Wide Web Conference.*

Devadoss, A. V. & Ligori, T. A. A. (2013). Forecasting of stock prices using multi layer perceptron. *International Journal of Computing Algorithm*, 2(1), 440-449.

Gandhmal, D. P. & Kumar, K. (2019). Systematic analysis and review of stock market prediction techniques. *Computer Science Review*, 34, 100190.

Han, J. J. & Kim, H. j. (2021). Stock price prediction using multiple valuation methods based on artificial neural networks for KOSDAQ IPO companies. *Investment Analysts Journal*, 50(1), 17-31.

Hasan, R. & Mohammad, S. M. (2015). Multifractal analysis of asian markets during 2007-2008 financial crisis. *Physica A: Statistical Mechanics and its Applications*, 419, 746-761. https://doi.org/https://doi.org/10.1016/j.physa.2014.10.030.

Heaton, J. B., Polson, N. G., & Witte, J. H. (2017). Deep learning for finance: deep portfolios. *Applied Stochastic Models in Business and Industry*, 33(1), 3-12.

Hewage, P., Behera, A., Trovati, M., Pereira, E., Ghahremani, M., Palmieri, F., & Liu, Y. (2020). Temporal convolutional neural (TCN) network for an effective weather forecasting using time-series data from the local weather station. *Soft Computing*, 24(21), 16453-16482. https://doi.org/10.1007/s00500-020-04954-0.

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.

Hoseinzade, E. & Haratizadeh, S. (2019). CNNpred: CNN-based stock market prediction using a diverse set of variables. *Expert Systems with Applications*, 129, 273-285.

Hota, H., Handa, R., & Shrivas, A. (2017). Time series data prediction using sliding window based RBF neural network. *International Journal of Computational Intelligence Research*, 13(5), 1145-1156.

Hu, Z., Zhao, Y., & Khushi, M. (2021). A survey of forex and stock price prediction using deep learning. *Applied System Innovation*, 4(1), 9. https://www.mdpi.com/2571-5577/4/1/9.

Huang, Y., Capretz, L. F., & Ho, D. (2021). Machine learning for stock prediction based on fundamental analysis. 2021 *IEEE Symposium Series on Computational Intelligence (SSCI)*, Orlando, USA, 1-10.

Jin, Z., Guo, K., Sun, Y., Lai, L., & Liao, Z. (2020). The industrial asymmetry of the stock price prediction with investor sentiment: Based on the comparison of predictive effects with SVR. *Journal of Forecasting*, 39(7), 1166-1178.

Jin, Z., Yang, Y., & Liu, Y. (2020). Stock closing price prediction based on sentiment analysis and LSTM. *Neural Computing and Applications*, 32(13), 9713-9729.

Kumar, K. & Haider, M. T. U. (2021). Enhanced prediction of intra-day stock market using metaheuristic optimization on RNN-LSTM network. *New Generation Computing*, 39(1), 231-272.

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, 521(7553), 436-444.

Lin, Y., Yan, Y., Xu, J., Liao, Y., & Ma, F. (2021). Forecasting stock index price using the CEEMDAN-LSTM model. *The North American Journal of Economics and Finance*, 57, 101421.

Liwei, T., Li, F., Yu, S., & Yuankai, G. (2021). Forecast of LSTM-XGBoost in Stock Price Based on Bayesian Optimization. *Intelligent Automation \& Soft Computing*, 29(3), 855-868. http://www.techscience.com/iasc/v29n3/43035.

Lu, W., Li, J., Li, Y., Sun, A., & Wang, J. (2020). A CNN-LSTM-Based model to forecast stock prices. *Complexity*, 2020, 6622927. https://doi.org/10.1155/2020/6622927.

Medsker, L. & Jain, L. C. (1999). *Recurrent Neural Networks: Design and Applications*, CRC Press.

Newall, P. W. S. & Weiss-Cohen, L. (2022). The Gamblification of investing: How a new generation of investors is being born to lose. *International Journal of Environmental Research and Public Health,* 19(9), 5391. https://doi.org/10.3390/ijerph19095391.

Pawar, K., Jalem, R. S., & Tiwari, V. (2019). Stock market price prediction using LSTM RNN. *Emerging Trends in Expert Applications and Security*, Springer, 493-503.

Rahimy, E. (2018). Deep learning applications in ophthalmology. *Current opinion in ophthalmology*, 29(3), 254-260.

Rundo, F. (2019). Deep LSTM with reinforcement learning layer for financial trend prediction in FX high frequency trading systems. *Applied Sciences*, 9(20), 4460. https://www.mdpi.com/2076-3417/9/20/4460.

Selvin, S., Vinayakumar, R., Gopalakrishnan, E. A., Menon, V. K., & Soman, K. P. (2017). *Stock price prediction using LSTM, RNN and CNN-sliding window model.* 2017 *International Conference on Advances in Computing, Communications and Informatics (ICACCI).*

Serrano, W. (2018). Fintech model: The random neural network with genetic algorithm. *Procedia Computer Science*, 126, 537-546.

Sezer, O. B., Ozbayoglu, M., & Dogdu, E. (2017). A deep neural-network based stock trading system based on evolutionary optimized technical analysis parameters. *Procedia Computer Science*, 114, 473-480.

Shen, J. & Shafiq, M. O. (2020). Short-term stock market price trend prediction using a comprehensive deep learning system. *Journal of Big Data*, *7*(1), 66. https://doi.org/10.1186/s40537-020-00333-6.

Usmani, S. & Shamsi, J. A. (2021). News sensitive stock market prediction: literature review and suggestions. *PeerJ Computer Science*, 7, e490.

Vidal, A. & Kristjanpoller, W. (2020). Gold volatility prediction using a CNN-LSTM approach. *Expert Systems with Applications*, 157, 113481. https://doi.org/10.1016/j.eswa.2020.113481.

Wang, H., Wang, J., Cao, L., Li, Y., Sun, Q., & Wang, J. (2021). A stock closing price prediction model based on CNN-BiSLSTM. *Complexity*, 2021, 5360828. https://doi.org/10.1155/2021/5360828.

Wang, J., Wang, X., Li, J., & Wang, H. (2021). A prediction model of CNN-TLSTM for USD/CNY exchange rate prediction. *IEEE Access*, 9, 73346-73354. https://doi.org/10.1109/ACCESS.2021.3080459.

Wang, X., Wang, Y., Weng, B., & Vinel, A. (2020). Stock2Vec: A hybrid deep learning framework for stock market prediction with representation learning and temporal convolutional network. *arXiv preprint arXiv:2010.01197*.

Wu, D., Wang, X., Su, J., Tang, B., & Wu, S. (2020). A labeling method for financial time series prediction based on trends. *Entropy*, 22(10), 1162.

Wu, J. M. T., Li, Z., Herencsar, N., Vo, B., & Lin, J. C. W. (2021). A graph-based CNN-LSTM stock price prediction algorithm with leading indicators. *Multimedia Systems*. https://doi.org/10.1007/s00530-021-00758-w.

Wu, Y., Schuster, M., Chen, Z., Le, Q. V., Norouzi, M., Macherey, W., Krikun, M., Cao, Y., Gao, Q., & Macherey, K. (2016). Google's neural machine translation system: Bridging the gap between human and machine translation. *arXiv preprint arXiv:1609.08144*.

Wu, Z., Fan, J., Gao, Y., Shang, H., & Song, H. (2019). Study on prediction model of space-time distribution of air pollutants based on artificial neural network. *Environmental Engineering & Management Journal (EEMJ)*, 18(7), 1575-1590.

Zhang, X., Gu, N., Chang, J., & Ye, H. (2021). Predicting stock price movement using a DBN-RNN. *Applied Artificial Intelligence*, 35(12), 876-892.

Zhao, Z., Rao, R., Tu, S., & Shi, J. (2017). Time-Weighted LSTM model with redefined labeling for stock trend prediction. 2017 *IEEE 29th International Conference on Tools with Artificial Intelligence (ICTAI)*.

Zhou, X., Pan, Z., Hu, G., Tang, S., & Zhao, C. (2018). Stock market prediction on high-frequency data using generative adversarial nets. *Mathematical Problems in Engineering*, 4907423. https://doi.org/10.1155/2018/4907423.

Zhu, R., Liao, W., & Wang, Y. (2020). Short-term prediction for wind power based on temporal convolutional network. *Energy Reports*, 6, 424-429.