

龔千芬、郝沛毅(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。

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

龔千芬

國立高雄科技大學 智慧商務系

郝沛毅*

國立高雄科技大學 智慧商務系

摘要

股價預測是橫跨金融與計算機科學領域的經典預測問題,由於成功預測股價的潛在好處,它吸引一代又一代的學者與投資者從不同的角度、無數的學理、眾多的投資策略和不同的實踐經驗來開發各種預測方法。股價預測的困難癥結點在於影響股票漲跌的因素太多。股市波動通常是由熱門新聞推動的,社群媒體上的推文則是反映了新聞事件的熱度,以及投資者對於該事件的態度。因此,分析社群推文的文字資訊與股價技術指數的數值資料,為能夠幫助於我們預測未來的股價變化。

由於股價受到眾多因素影響,很難通過簡單的模型進行預測。深度學習擁有優異的特徵學習能力,支持向量機則是擁有優異的推理能力,本研究結合兩者的優點。本研究提出一個混和深度模型來自動學習重要的特徵,該混和深度模型是由卷積神經網路(CNN)、雙向長短期記憶(BiLSTM)與注意力機制(AM)組成。CNN用於擷取輸入數據的位置不變特徵,BiLSTM則是提取長時間依賴性的特徵,AM用於捕捉過去不同時間特徵狀態對股票收盤價的影響,以提高預測的正確率。接著,本研究將擷取得到的特徵餵給模糊學生支持向量機來建立最佳的股價預測模型,並且透過轉移學習理論建立嶄新的深層模糊學生支持向量機。本研究在台積電股票的預測正確率最高為76.9667%,友達股票的預測正確率最高為87.0856%,與經典的股價預測模型相比,本研究所提出的方法的預測正確率明顯優於最先進的股價預測模型。

關鍵詞：股價預測、模糊支持向量機、卷積神經網路、雙向長短期記憶體、注意力機制

* 本文通訊作者。電子郵件信箱：haupy@nkust.edu.tw
2022/4/21 投稿；2022/6/18 修訂；2022/7/18 接受

Kung, C.F. & Hao, P.Y (2022). A Stock Closing Price Prediction Model based on Deep Neural Networks and Deep Fuzzy Twin Support Vector Machine. *Journal of Information Management*, 29(4), 303-333.

A Stock Closing Price Prediction Model based on Deep Neural Networks and Deep Fuzzy Twin Support Vector Machine

Chien-Feng Kung

Department of Intelligent Commerce, National Kaohsiung University of Science and
Technology

Pei-Yi Hao*

Department of Intelligent Commerce, National Kaohsiung University of Science and
Technology

Abstract

Stock closing price prediction is an important problem in the intersection of computer science and finance. Due to the potential advantages of stock closing price prediction, it attracts generation after generation of investors as well as scholars to continuously develop various prediction methods from different perspectives, a multitude of investment strategies, different practical experiences, and a myriad of theories. The stock closing price is often affected by hot news, and the tweets related to news reflect the heat of the breaking news, as well as the sentiment of investors towards the breaking news. Consequently, analyzing tweets and historical stock market indices may help us to predict future price changes.

Since the stock price is affected by many factors, it is difficult to predict through a simple model. Deep learning methods have the advantage of learning features. Support vector machines have the advantage of generalizing very well. This paper combines the advantages of both models. Therefore, this paper develops a hybrid deep model to automatically learn important features. This deep model is composed of convolutional neural networks (CNN), bi-directional long short-term memory (BiLSTM), and attention mechanism (AM). CNN is utilized to discover the time-invariant features of the input data. BiLSTM is used to extract time dependency features. AM is used to extract the influence of feature states on the stock closing price at

* Corresponding author. Email: haupy@nkust.edu.tw

2022/4/21 received; 2022/6/18 revised; 2022/7/18 accepted

different times in the past to improve the classification performance. The obtained feature is fed to a deep fuzzy support vector machine to build an optimal stock prediction model. As for this study, the highest forecast accuracy is 76.9667% and 87.0856% for Taiwan Semiconductor Manufacturing Corporation and AU Optronics Corporation, respectively. When compared with previous prediction models, the method proposed in this study is significantly better than the state-of-the-art support vector machine and deep learning models.

Keywords: Stock prediction, Fuzzy support vector machine, Convolutional neural networks, Bi-directional long short-term memory, Attention mechanism

壹、前言

股票市場是資金的避風港，已發行的股票在此進行轉讓、交易和流通。股票價格是以發行市場為基礎，但股票市場的結構和交易活動遠比發行市場本身複雜得多。因此，做出準確的股票價格預測，成為一項複雜而艱鉅的任務。近年來，隨著經濟的快速發展，越來越多的人開始投資股市。股票走勢預測是投資者關注的重點領域，如果能夠成功預測股票走勢，那麼投資者就可以採取更合適的交易策略，從而可以顯著地增加獲利與降低投資風險。由於股票預測的諸多潛在好處，它吸引一代又一代的學者和投資者從不同的角度、無數的理論、眾多的投資策略和不同的實踐經驗，來不斷開發各種預測方法。即便至今，股價預測仍然是一個熱門且充滿挑戰性的任務，仍然有許多關於股價預測的研究能刊登在IEEE Access、IEEE KDE與IEEE trans. on Fuzzy Systems等頂尖的國際期刊上面(Chandra et al. 2021; Cao et al. 2020; Li et al. 2021)，可見得它仍然是主流學術界認可的重要研究主題。造成未來股票價格預測困難的癥結點，在於能夠影響股票漲跌的因素太多。對於上市櫃公司，能夠影響其發行股票價格的因素包含：收益和利潤的新聞發布、未來估計收益、新產品的推出或產品召回、獲得新的大合同、員工裁員與管理階層的重大變動、預期的收購或合併，以及會計錯誤或醜聞。以上這些因素只是公司因素，還有其他因素可能會影響其股價，例如相關行業的獲益表現、投資者情緒、政治與經濟因素等。由於股價變化受到眾多因素影響，因此本研究將要考慮各種不同的數據來源，以建立更有效的股價變動預測模型。

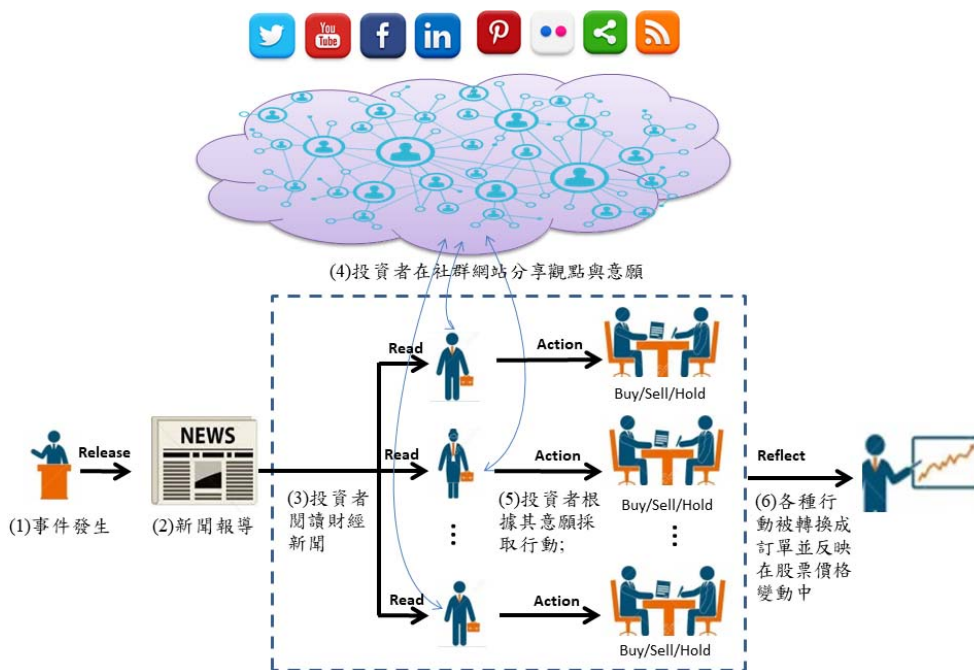


圖 1:社群推文對股價影響的情況

(1) 事件發生; (2) 新聞報導; (3) 投資者閱讀財經新聞; (4) 投資者在社群網站分享觀點與意願; (5) 投資者根據其意願採取行動; 和 (6) 各種行動被轉換成訂單並反映在股票價格變動中。

許多文字數據來源，例如財經新聞、社群推文和年度財務報告等，可以對其進行分析以挖掘重要的訊息。尤其是財經新聞，已經被證實是比數值數據更好的隱藏訊息來源 (Hao et al. 2021a; Li et al. 2021)。例如，一篇關於公司的新聞中包含諸如“辭職”、“違約風險”之類的詞句或短語，有助於投資者預測其公司股價的下跌。事實上，當重大事件發生時（例如，某間全球金融銀行破產），傳統的技術分析預測模型將無法有效地預測股價的變動，因為這些事件首先出現在突發新聞中，而不是過去的價與量的數據中（例如，過去五天的股票價格與成交量）。除此之外，社群媒體上的推文(tweet and post)內容亦是股價預測的重要數據來源，研究顯示，社群媒體上的內容大多數都與焦點新聞相關。此外，社群媒體可以更迅速地傳播訊息，同時確保熱門新聞中包含的重要訊息的高覆蓋率(Zhao et al. 2011; Xu & Cohen 2018)。從這個觀點來看，社群媒體的推文是預測股票市場的良好來源。透過社群推文，我們可以獲得所有在交易日內發生的重大新聞事件，並且從社群推文中，我們可以得到人們對於關注的事件的情緒與評論。圖1顯示社群媒體的文字內容是如何影響股價的行為。因此，本研究將要結合股票價格數據與社群網站上的文字內容來預測股票價格的走勢。

近年來，隨著各種深度神經網路 (Deep Neural Networks, DNN) 技術的出現，各種透過智能計算來解決複雜的股票市場預測問題的深度模型如雨後春筍般冒出 (Li et al. 2021; Chandra et al. 2021)。自從2016年3月AlphaGo以四比一擊敗南韓圍棋高手李世石而聲名大噪後，人工智慧一夕之間成為家喻戶曉的熱門科技，帶動這波AI熱潮的幕後推手便是深度學習(Deep Learning)。「深度學習」可以說是類神經網路的品牌重塑，Hinton & Salakhutdinov (2006)在《Science》期刊上的論文闡述：具有多個隱藏層的人工神經網路擁有優異的特徵學習能力，有利於視覺化或分類；換句話說，深度學習乃是透過組合低層特徵，來形成更加抽象的高層特徵，以得到資料的分層特徵表示。因此，「深度模型」是手段，「特徵學習」才是目的。深度學習的關鍵優勢，是具有自動抽取特徵的能力，它可以取代傳統專家在特徵工程所花費的時間，並且對於圖像、語音這種特徵不明顯的資料類型，能夠在大規模訓練數據上取得優異的效果。由於深度學習的優異特徵學習能力，很適合處理股價時間序列類型的資料，並且從中自動學習關鍵特徵，取代傳統專家根據經驗手動建立的特徵(亦即是，所謂的技術指標)，所以近年來最頂尖的股價預測模型，都是以深度學習為基礎(Li et al. 2021; Chandra et al. 2021)。

深度學習的崛起，也取代了支持向量機(support vector machine, SVM)在機器學習領域霸的地位，但是，還是有一些研究顯示，支持向量機的預測效能是優於深度學習的(Xu & Cohen 2018)。這樣的結果並不讓人意外，因為支持向量機雖然沒有學習特徵的能力，但是如果在給定良好特徵的情況下，基於統計學習理論創建的支持向量機仍然是最佳的分類器。Huang & LeCun (2006)首度發表結合卷積神經網路(convolutional neural network, CNN)與SVM的方法，文章中提到，CNN很適合學習圖像中位置或方向不變性的特徵(invariant feature)，但是它並不是最適合於分類任務的方法，相反的，SVM在給定良好特徵的條件下，它是最佳分類

器，但是SVM無法學習圖像中位置或方向不變性的特徵(invariant feature)。所以，他們使用CNN學習不變性特徵，再將特徵交給SVM建立分類器，獲得正確率明顯地提升。這也是說，人工智慧的大師，CNN的創始人 LeCun也為SVM背書，所以SVM的研究並沒有因為深度學習而落幕，許多學者跟隨著大師的腳步，使用深度學習來擷取特徵，再將該特徵交給SVM來學習分類器，並且應用到各種現實的辨識任務中(Xie & Yu 2021)。因此，支持向量機並不是一個被取代掉的落伍技術，許多關於支持向量機的研究，還是可以發表在頂尖的國際期刊上面(Hao 2021b)，這證明支持向量機還是具有研究的潛力。

因此，本研究利用深度學習模型來擷取關鍵特徵，再將特徵餵給支持向量機來學習預測模型，以完美融合這兩種方法的優點。除此之外，由於簡單的深度學習模型無法有效的分析錯綜複雜的異質股價相關數據，因此，本研究提出一種CNN-BiLSTM-AM混和模型來擷取特徵，該模型由卷積神經網路(convolutional neural networks, CNN)、雙向長短期記憶體(bi-directional long short-term memory, BiLSTM)和注意力機制(attention mechanism, AM)組成。CNN用於提取輸入數據的特定模式，BiLSTM則是使用兩種方向(從過去到未來與從未來到過去)來發掘CNN找到的模式的因果關係，AM用於捕捉在過去不同的時間狀態特徵對於股票收盤價預測的影響，以提高預測精度。這些深度模型擷取得到的關鍵特徵，將要餵給模糊學生支持向量機(fuzzy twin support vector machine)來建立預測模型(Hao et al. 2021a)，模糊理論很適合處理股價預測問題，因為股價的漲跌之間存在一個模糊的邊界，而且新聞文章與社群推文上的文字本質就是模糊的，而本研究提出的模糊學生支持向量機則完美的融合模糊理論與支持向量機的優點。除此之外，雖然具有嚴謹理論基礎的SVM被證明擁有優異的推理能力，研究者仍然不斷的要改善SVM，尤其是，如何透過深度學習的深層網路架構來改良SVM的性能。近幾年，具有深層結構的SVM網路的研究如雨後春筍般冒出(Wang et al. 2019; Pławiak, Abdar, & Acharya 2019)，由此可見，具有深層結構的支持向量機在學術界上具有一定的吸引力。因此，本研究使用轉移學習理論來建構具有深層架構的支持向量機，以提高支持向量機的預測能力。

貳、文獻探討

長短期記憶體(long short-term memory, LSTM)網路已被證明是學習和預測具有長期依賴性的時間數據的非常有用的工具，Yan, Weihan, & Chang (2021b)利用LSTM深度神經網路對上海金融交易數據進行建模和預測，他使用前30天的股價來預測3天後的漲跌趨勢(分為大漲、漲、跌、大跌等類別)。Cherati, Haeri, & Ghannadpour (2021)採用LSTM模型預測BTC/USD每日收盤價方向，他使用貝葉斯優化模型(Bayesian optimization model)來優化LSTM模型的超參數，以提高模型的準確性，他們並評估MACD指標和輸入矩陣維度對預測準確性的影響。卷積神經網路(convolutional neural network, CNN)在圖像辨識領域帶來了重大的突破，它能夠挖掘圖像中與位置無關的特徵，自從導入CNN來分析金融數據以

來，許多研究人員致力於透過將股票市場數據轉換為圖像來預測股票趨勢。Maratkhan et al. (2021)使用 15 個不同期間(6-21 天)來計算 15 個不同的技術指標的數值，得到的 15×15 的數值矩陣可以用來建立一個 15×15 的圖片，並且將這些技術指標得到的 2D 圖片餵入到 CNN 模型當中，來預測隔日的股價漲跌趨勢。此外，股票領先指標(leading indicators)是對影響未來經濟發展的經濟指標的統計，市場分析師經常參考這些指標來分析未來經濟發展及其對未來匯率發展方向的影響。Wu et al. (2020)使用領先指標中的期權和期貨(options/futures)，由這兩個指數和股價構成了每個樣本的特徵，它將目標股價的前 30 天的股票價格的 20 個期權（10 個看漲期權和 10 個看跌期權）來生成期權數據，建立一個 30×20 的 2D 矩陣，並且將此矩陣輸入到 CNN 以預測股價的漲跌。Chen et al. (2021)提出一種使用基於圖卷積特徵的卷積神經網路（graph convolutional feature based convolutional neural network, GC-CNN）模型進行股票趨勢預測，他運用相關分析找出跟目標股票有相關性的股票，把這些股票的技術指標轉換成為 2D 矩陣，然後餵給 2D-CNN(也稱為 graph-based CNN)以預測目標股票的漲跌趨勢。

由於長短期記憶體（long short-term memory, LSTM）與卷積神經網路(convolutional neural network, CNN)很適合處理時間序列的數據，因此廣泛運用於股價預測的研究當中。不過，除了 LSTM 與 CNN 之外，還有研究者調查其他深度模型對於股價預測的能力，例如，Yan & Yang (2021a)提出採用注意力機制的編碼-解碼模型(encoder-decoder model of attention mechanism)，從特徵和時間兩個方面加入注意力機制，編碼器和解碼器都使用 LSTM 神經網路。該方法解決了時間序列預測中的兩個問題；第一個問題是多個輸入特徵對目標序列的影響程度不同，使用特徵注意力機制來處理這個問題，可以得到不同輸入特徵的權重，得到更穩健的特徵關聯關係；第二個問題是序列前後的數據有很強的時間相關性。使用時間注意力機制來處理這個問題，可以得到不同時間點的權重，以獲得強健的時序依賴性特徵。Haq et al. (2021) 提出使用多種特徵選擇方法來選擇重要的特徵，並將這些特徵組合起來作為具有注意力機制(attention mechanism)的深度生成模型(deep generative model)的輸入，他們從四十四個技術指標中，透過獨立訓練邏輯斯迴歸模型、支持向量機和隨機森林來計算這些特徵的重要性，並透過預先指定的閾值，來刪除排名最低的特徵。Yun, Yoon, & Won (2021)提出一種混合 GA-XGBoost 的預測系統，該系統具有增強的特徵工程，並且使用 GA-XGBoost 算法來進行最佳特徵集選擇，以在維數的詛咒(the curse of dimensionality)和維數的祝福(the blessing of dimensionality)之間取得良好的平衡與和諧。此外，過去這些基於深度學習的模型大多屬於監督學習方法，無法處理長期目標。因此，Shi et al. (2021)提出一種基於深度強化學習(deep reinforcement learning based)的股市交易模型，他們精心設計獎勵函數和基於深度學習的策略網路，使模型能夠捕獲股票數據中的隱藏依賴項和潛在動態。由於深度網路模型具有優異的特徵學習能力，而支持向量機(support vector machine, SVM)則是最佳的分類器，因此，Xie & Yu (2021)提出使用卷積自動編碼器(convolutional autoencoder, CAE)來非監督的學習

特徵，並且將學習得到的特徵餵給 SVM 來訓練最佳的股價預測模型。此外，股價趨勢預測屬於模糊分類的問題，在股價{上漲}與{下跌}的邊界之間，存在一個模糊的區域，股價上漲+3%的樣本，很明顯地屬於{上漲}的類別，反之，股價上漲+0.0003%的樣本，則是位於股價{上漲}與{下跌}的邊界之間的模糊地帶，Hao et al. (2021a)提出使用模糊孿生支持向量機(fuzzy twin SVM)來預測股價的漲跌趨勢，他們使用一個模糊超平面(fuzzy hyperplane)來分割{上漲}與{下跌}類別，並且運用情緒辭典來擷取文章中的情緒特徵，以及使用潛在狄利克雷分配(Latent Dirichlet allocation, LDA)來擷取文章中潛在的主題特徵，並將這些文章中情緒與主題特徵餵給模糊孿生支持向量機，以分類隔日股票價格的漲跌變動。

上述研究文獻顯示，由於 CNN 與 LSTM 特別適合處理時序數據，所以它們是股價預測最常使用的學習模型，此外，AM 可以對較重要的因素給予較高的關注，進而提升預測效能。因此，本研究將要使用 CNN、BiLSTM 與 AM 來處理股價預測問題中的時序資料，並將混和深度模型學習得到的特徵餵給 SVM 來建構最佳的股價預測模型。此外，使用模糊超平面的模糊孿生支持向量機，可以有效地處理股價漲跌之間的模糊曖昧的區域，本研究將要透過深度轉移學習策略來進一步的提升模糊孿生支持向量機的預測能力。

參、研究方法

股價預測是艱鉅困難的問題癥結點，在於股價變動受到許多外部因素的影響，因此，為了成功地預測股價的變動，在我們的股價預測模型當中，除了考慮每日的股票價格之數據資料外，我們還將社群網路的群眾對於新聞事件的評論當作預測明日股價漲跌的依據。我們將應用機器學習與文字探勘技術，自動化處理與分析在社群媒體討論區上大量的文章資料，並且搭配每日股價的數據資料，融合這些資訊後，本研究將導入新穎的混和深度網路模型與具有深層結構的模糊支持向量機網路，來預測未來股票價格變動的趨勢，由於影響股價的訊息是異質且錯綜複雜，很難透過簡單的模型進行預測。為更準確地預測股票價格，本研究提出一種基於 CNN-BiLSTM-AM 與深度模糊支持向量機的方法來預測第二天的股票收盤價。透過混和深度模型來擷取異質數據中的關鍵特徵，再將這些特徵餵給深層模糊支持向量機來建構最佳預測模型。本研究用來擷取關鍵特徵的深度模型是由卷積神經網路 (convolutional neural networks, CNN)、雙向長短期記憶體 (bi-directional long short-term memory, BiLSTM) 和注意力機制 (attention mechanism, AM) 組成。CNN 可以從輸入的股票數據中提取特徵。長短期記憶體 (LSTM) 是循環神經網路 (Recurrent neural network, RNN) 的改進，它避免了 RNN 帶來的梯度消失和梯度爆炸問題。BiLSTM 可以充分發現股票時間序列數據的相互依存關係。AM 可以對不同時間狀態特徵進行加權計算，從而提高預測的準確性。詳細流程如圖 2 所示。首先，我們彙總各種不同訊息的數據源來分析，每日匯總得到的異質資料將餵入 CNN 擷取特徵，而 CNN 輸出的每日特徵向量將視為時間序列並且餵給 BiLSTM 網路。

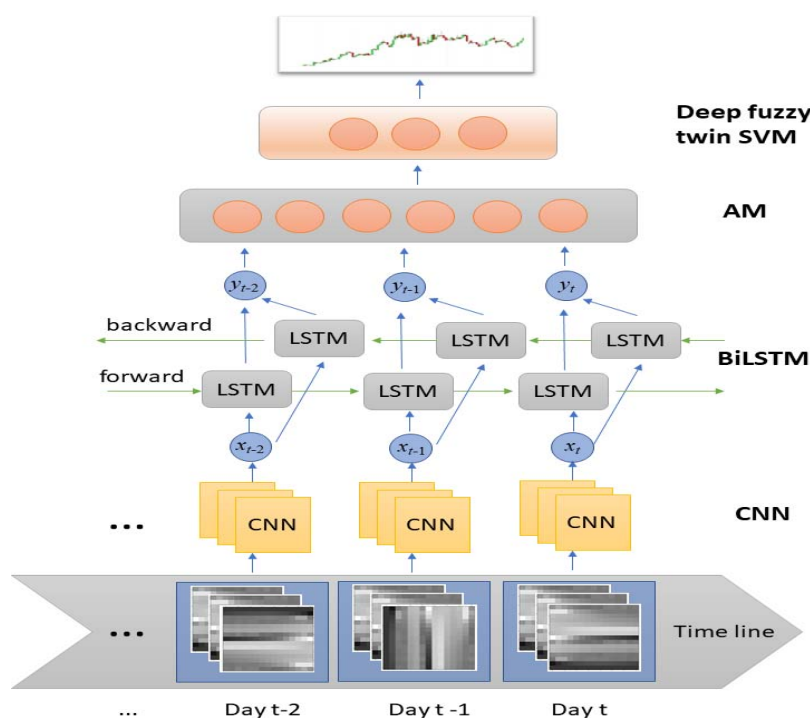


圖 2: 結合 CNN-BiLSTM-AM 與模糊深層支持向量機的股價預測模型

BiLSTM 網路透過前向與反向運算來分析這些時間序列的因果關係，例如指數的脈動、新聞事件的蔓延與情緒的波動，並且使用注意力機制來關注那些重要的因素，最終，重要的因素將餵給最頂層的深度模糊支持向量機，建構最佳推論模型來預測明日的股價。接下來，我們將詳細介紹混和模型中的各項元件的原理與運作流程。

一、透過卷積神經網路學習每日異質資料來源的關鍵特徵

每當深度學習技術又有重大突破時，這些突破大多都與卷積神經網路（Convolutional Neural Networks, CNN）有關，CNN 是目前深度神經網路（deep neural network）的發展主力，在語音與圖片辨別的任務上，它甚至可以做到比人類還更精準的分類正確率。CNN 也可以用來處理其他型態的資料，技巧是將其它型態的資料轉換成類似圖片的格式。在本研究中，我們將使用 CNN 來擷取每日的股價指數的深層技術特徵與社群推文的深度語義特徵，底下我們介紹如何使用 CNN 來分析各種不同類型的資料並擷取關鍵特徵。針對不同類型的資料，我們將使用彼此獨立的 CNN 網路來分析與擷取特徵，此處，CNN 擷取的特徵是指 CNN 最後一層輸出的向量。最後，本研究再將這些針對不同類型資料的各個 CNN 網路的輸出向量串接在一起，然後餵給下一層的雙向長短期記憶網路（BiLSTM）進行分析長時間的依賴性。

(一) 透過 CNN 來擷取每日的股價短期技術指標特徵

自從引入 CNN 來分析金融數據以來，許多研究人員致力於透過將股票市場數據轉換為圖像來預測股票趨勢。在本研究中，受到 Maratkhan et al. (2021) 研究

的啟發，我們使用基於技術分析的方法來創建圖像，技術分析方法是利用過去的財務數據進行數學建模以預測未來的市場趨勢。技術指標是交易者用於技術分析的主要工具。在本研究中，我們使用 15 種不同的技術指標，其值以 15 種不同的時間間隔（6-20 天）生成。指標有簡單移動平均線 (simple moving average, SMA)、加權移動平均線 (weighted moving average, WMA)、指數移動平均線 (exponential moving average, EMA)、赫爾移動平均線(Hull Moving Average, HMA)、三重指數移動平均線(triple exponential moving average, TEMA)、平滑異同移動平均線指標(Moving Average Convergence & Divergence, MACD)、相對強弱指標 (Relative Strength Index, RSI)、威廉指標 (Williams %R, W%R)、商品通道指數 (Commodity Channel Index, CCI)、錢德動量擺動指標 (Chande Momentum Oscillator, CMO)、百分比價格振盪器 (percentage price oscillator, PPO)、價格變動率指標 (Rate of change, ROC)、蔡金資金流量指標 (Chaikin Money Flow Index, CMFI)、定向運動指標 (directional movement indicator, DMI) 和拋物線停損點轉向指標(Parabolic Stop And Reverse, PSI)。我們使用 TA4J (Java 技術分析函數庫) 計算 15 個不同時間間隔的上述指標的值(亦即是，使用前 6 日的股價資料計算技術指標的數值，逐步將時間間隔增加，到使用前 20 日的股價資料計算技術指標的數值)，因為我們的目標是捕捉中短波交易特徵。然後我們創建一個尺寸為 15×15 的圖像，其中行代表技術指標，列代表不同的時間間隔(6-20 天)。該過程產生圖像的範例如圖 3 所示，這些技術指標的計算公式顯示在附錄中。參考(Maratkhan et al. 2021)的研究，我們將技術指標的數值範圍正規化到[0,255]的區間，並且將得到的 15×15 的數值矩陣，視為尺寸為 15×15 的灰階影像，輸入給 CNN 模型分析，學習每日的股價技術指標特徵，並且將 CNN 最後一層輸出的向量，餵給 BiLSTM 模型分析長時間的因果關係。

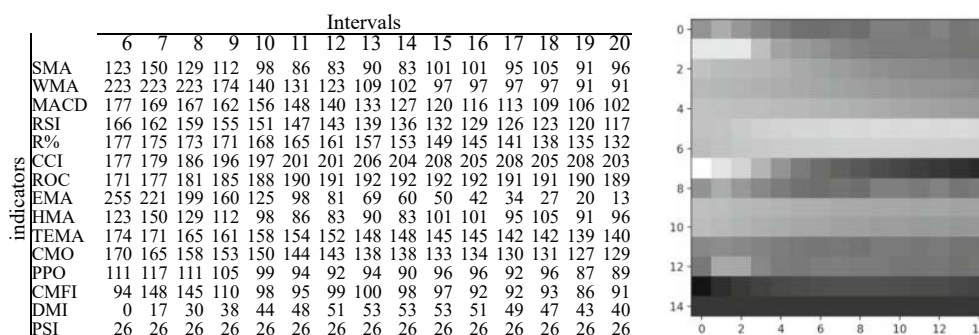


圖 3: CNN 技術指標 2D 輸入圖像的範例

(二)透過CNN來擷取每日的社群推文語義特徵

在本研究中，受到 Kim (2014)研究的啟發，我們使用詞嵌入向量(word embedding vector)與卷積神經網路 (CNN) 構建文字語義分析模型，對財經論壇的推文與帖子進行分析。CNN 經常應用於自然語言處理領域，根據 Kim (2014)的研究指出，應用較少的層數和相對較多的濾波器 (例如，他們的工作中有 128 個)，通常可以在文字分析任務中，實現更高的分類精度。因此，我們在卷積層

上設置了 128 個濾波器，本研究處理文字分析的 CNN 網路包含數據輸入層、一個卷積層、一個最大池化層和一個全連接層。首先，我們將每日發表的論壇帖子彙總在一起，視為單一篇文章，並且進行中文斷詞與去除停用詞，結果如表 1 所示，接者，本研究使用 word2vec 工具¹來生成 64 維的向量代表詞嵌入特徵，word2vec 將根據字詞在不同句子的上下文內容來學習詞嵌入特徵向量，經常出現在相似內容情境的字詞，將被映射成彼此接近的兩個詞嵌入特徵向量，詞嵌入特徵向量可以表示語意特徵(情緒也是屬於語意特徵的一種)。我們將出現在同一個句子的關鍵詞的詞嵌入向量計算其平均值，結果如表 2 所示。

表 1:對句子進行中文斷詞與去除停用詞

類型	範例
句子	公司的第三季財報出來了，我感覺表現不錯。
斷詞結果	公司 的 第三季 財報 出來 了 ， 我 感覺 表現 不錯 。
去除停用字	公司 第三季 財報 出來 感覺 表現 不錯

表 2:使用詞向量得到句子嵌入向量

詞	d_1	d_2	...	d_n
公司	-0.28	0.20	...	-0.63
第三季	0.81	-0.45	...	0.23
財報	-0.18	1.21	...	0.35
出來	0.41	0.68	...	-1.53
感覺	1.16	-0.77	...	0.70
表現	0.75	1.12	...	0.84
不錯	1.13	0.55	...	0.67
平均值	0.30	0.36	...	0.20

對於句子“公司|第三季|財報|出來|感覺|表現|不錯”，我們將這些詞的詞嵌入向量計算出平均值，其向量為(0.30,0.36,...,0.20)，並且將該向量視為該句子的嵌入向量，在 CNN 的輸入層中，給定一篇文章 t 是由 n 個句子組成， $\mathbf{s}_i \in R^k$ 是文章中第 i 個句子的對應的嵌入向量， k 是句子嵌入向量的維度。在此基礎上，文章 t 可以表示為：

$$\mathbf{s}_{1:n} = \mathbf{s}_1 \oplus \mathbf{s}_2 \oplus \dots \oplus \mathbf{s}_n \quad (1)$$

其中 \oplus 表示連接運算符，代表文章是由這些句子的向量串接而成尺寸為 $n \times k$ 的 2D 矩陣。接下來，我們將此 2D 矩陣視為圖像輸入給 CNN 網路，在 CNN 的卷積層，透過對輸入層滑動一個大小為 $h \times k$ 的卷積核(convolution kernel)進行卷積操作，得到特徵圖 \mathbf{c} ，其計算過程如方程式(2)所示：

$$c_j = f(\mathbf{w} \cdot \mathbf{s}_{j:j+h-1} + b) \quad (2)$$

其中， c_j 表示常規處理後的第 j 個特徵值。 f 是非線性的卷積核激發函數， \mathbf{w} 是卷積核權重矩陣 ($\mathbf{w} \in R^{h \times k}$ ， h 表示卷積核的視窗涵蓋的句子數量， k 表示句

¹ word2vec. <https://code.google.com/p/word2vec/>.

子向量的維度)， b 是偏移量(bias)，兩者都會在訓練階段學習。為了提取最重要的特徵，在池化層中使用最大值池化 (max-pooling) 操作，算式如下：

$$\hat{c} = \max(c_j) \quad (3)$$

經過最大值池化後，輸出特徵 $\mathbf{v} = [\hat{c}_1, \hat{c}_2, \dots, \hat{c}_m]$ 將輸入至全連接層，其中， m 是濾波器的數量，最後一層輸出的向量將會餵給 BiLSTM 模型分析長時間的因果關係。

依據過去的研究成果(Kung et al., 2021)，我們將 CNN 最後一個全連接層(full connected layer)的輸出，視為 CNN 擷取到的特徵。更明確的說，假設處理股價技術指標的 CNN 的最後一個全連接層有 m_1 個神經元，我們將這 m_1 個神經元的輸出值，視為一個維度為 m_1 的向量 \mathbf{v}_p ， \mathbf{v}_p 是 CNN 從 15 種技術指標的前 6 到 20 天的數值資料(尺寸為 15×15 的矩陣)所擷取到的特徵向量。假設處理社群推文的 CNN 的最後一個全連接層有 m_2 個神經元，我們將這 m_2 個神經元的輸出值，視為一個維度為 m_2 的向量 \mathbf{v}_t ， \mathbf{v}_t 是 CNN 從社群媒體推文的文字訊息(n 個句子，句子嵌入向量維度 $k=64$)，尺寸為 $n \times k$ 的實數值矩陣)所擷取到的特徵向量。為了將異質資料(股價與文字)整合在一起，我們將 \mathbf{v}_p 與 \mathbf{v}_t 串接在一起，亦即是 $\mathbf{v} = \mathbf{v}_p \oplus \mathbf{v}_t$ ，其中 \oplus 代表串接運算元， \mathbf{v} 是一個維度為 $m_1 + m_2$ 的向量，我們將 \mathbf{v} 視為整合特徵向量，並且將每一日擷取到的整合特徵向量，視為時間序列餵給 BiLSTM，讓 BiLSTM 分別從前向與後向分析他們的因果關係。

二、透過雙向長短期記憶體網路學習時間因果特徵

長短期記憶體(Long short-term memory, LSTM)是一種循環神經網路(Hochreiter & Schmidhuber 1997)，他是專門設計用於避免深度循環網路中梯度消失和梯度爆炸的問題。LSTM 網路能夠將隱藏狀態的先前訊息，保留在內部儲存器。因此，它特別適用於處理事件之間存在長期時間依賴性的任務。LSTM 是由儲存單元狀態(memory cell state)，表示為 C_t ，與以下三個閘門構成：遺忘閘門 $f_t \in [0, 1]$ ，輸入閘門 $i_t \in [0, 1]$ ，以及輸出閘門 $o_t \in [0, 1]$ 。這三個閘門會相互影響，以控制訊息的流動。在訓練過程中，網路會學習記住什麼，以及何時允許讀/寫，以最大程度地減少分類錯誤。更具體地說，「忘記」閘門決定來自先前儲存單元狀態的哪些訊息已過期，並且應該予以刪除；輸入閘門從候選儲存單元狀態 C_t^* 中選擇訊息，以更新單元狀態；輸出閘門會過濾儲存單元中的訊息，因此模型僅考慮與預測任務相關的訊息。每個閘門的數值的計算如下，其中 $W_{[i,f,C,o]}$ 是權重矩陣， $b_{[i,f,C,o]}$ 是偏差向量：

$$i_t = \text{sigmoid}(W_i \cdot [y_{t-1}, X_t] + b_i) \quad (4)$$

$$f_t = \text{sigmoid}(W_f \cdot [y_{t-1}, X_t] + b_f) \quad (5)$$

$$C_t^* = \text{tanh}(W_C \cdot [y_{t-1}, X_t] + b_C) \quad (6)$$

$$o_t = \text{sigmoid}(W_o \cdot [y_{t-1}, X_t] + b_o) \quad (7)$$

LSTM 模組中的儲存單元值 C_t 和輸出標籤 y_t 則是使用以下公式計算

$$C_t = C_{t-1} \cdot f_t + C_t^* \cdot i_t \quad (8)$$

$$y_t = o_t * \tanh(C_t) \quad (9)$$

對於我們的任務，只有在最後一個時間步伐 T 的輸出向量，是我們擷取出來的特徵。LSTM 可以有效地捕獲時間序列數據中的潛在時間結構，因此特別適合對於股票相關數據中的動態訊息進行建模，在股價預測任務中，長時間間隔內事件 (CNN 擷取的特徵) 之間存在很強的統計依賴性。而這種依賴性，是識別股價波動變化的早期跡象的關鍵特徵。

LSTM 雖然可以獲取遠距離的特徵訊息的依賴性，但是，傳統 LSTM 的一個主要缺點是它們只利用先前的狀態來確定未來的狀態，並沒有使用反向訊息。然而，在時間序列預測應用中，能夠充分考慮時間序列數據的前向和後向的訊息規律，可以有效提高預測精度。因此，Graves & Schmidhuber (2005) 提出雙向長短期記憶體 (bi-directional long short-term memory, BiLSTM)，BiLSTM 是由兩個 LSTM 組成，一種是從過去到未來，一種是從未來到過去，這與傳統的 LSTM 網路中的單向狀態傳輸不同，BiLSTM 透過向後運送訊息，可以保留來自未來的狀態訊息。由於 BiLSTM 考慮數據傳輸正向與反向的變化規律，可以利用過去和未來的訊息做出更完整與詳細的決策，圖 2 中的 BiLSTM 的水平方向的箭頭表示模型中時間序列訊息的雙向流動，而圖 2 中的 BiLSTM 的垂直方向的箭頭則是表示數據訊息的流動 (從輸入層到隱藏層到輸出層)。

三、透過注意力機制捕抓重要的時間狀態特徵

注意力機制 (Attention mechanism, AM) 是由 Treisman & Gelade (1980) 提出，AM 的主要思想來自於人類視覺注意的過程，當我們觀察外在事物時，我們通常只會先關注其中的某一部分，每個人的關注焦點都不一樣。例如，當我們觀察一個人時，我們首先注意到的可能是他的臉、身高、體型、衣服等等。得到各個部分的訊息後，就可以對其進行排列組合，最終得到這個人的整體感覺。換句話說，人類視覺可以快速找到關鍵區域，並在關鍵區域添加注意力焦點以獲得所需的詳細訊息。同樣，AM 選擇性地關注一些更重要的訊息，忽略不重要的訊息，並且分配訊息的重要性。注意力機制現在已經成為人工智能結構的重要組件部分，並且在自然語言處理、時間序列預測、語音訊號處理與智能計算等諸多領域都有大量應用。注意力機制一般由兩部分任務組成：確定哪些輸入需要更多的注意力，提取關鍵部分的特徵以獲得重要訊息。換句話說，注意力機制可以對輸入的每一部分賦予不同的權重來過濾主要特徵，這也是注意力機制被如此廣泛使用的原因。

在處理輸入時，注意力值的計算是所有操作中最關鍵的。首先，將所有輸入作為一個整體記錄為 X ，然後選擇一個查詢變量 (query variable) 為 q 。 q 的作用是在整個 X 中尋找與選擇部分訊息。這裡使用 soft attention 機制來選擇所有輸入，對重要的輸入給予更多的關注，對不太重要的輸入給予較少的關注。假設共

有 N 個輸入， q 是查詢變量， s 是注意力計算函數，那麼第 i 個輸入的注意力權重為：

$$a_i = \text{softmax}(s(x_i, q)) = \frac{\exp(s(x_i, q))}{\sum_{j=1}^N s(x_j, q)} \quad (10)$$

注意力計算函數有很多種，下面介紹四種方法：

$$s(x_i, q) = v^T \tanh(Wx_i + Uq) \quad (11)$$

$$s(x_i, q) = x_i^T q \quad (12)$$

$$s(x_i, q) = \frac{x_i^T q}{\sqrt{d}} \quad (13)$$

$$s(x_i, q) = x_i^T Wq \quad (14)$$

這四種方法稱為加法、點積法、縮放點積法和雙線性法，其中 W 、 U 、 v 是需要學習的網路參數， d 是輸入訊息的維度。得到注意力權重後，我們可以透過加權平均計算注意力值：

$$\text{att}(X\Delta q) = \sum_{i=1}^N a_i x_i \quad (15)$$

四、模糊孿生支持向量機作為最頂層的分類器

在本研究中，我們將以模糊孿生支持向量機(Hao et al. 2021a)作為最頂層的分類器，以判別股價的漲跌趨勢。模糊孿生支持向量機(fuzzy twin support vector machine)將模糊集合理論融合至孿生支持向量機(twin SVM)當中，並且完美結合支持向量機與模糊理論的優點。

模糊孿生支持向量機(fuzzy twin SVM)尋找一對非平行模糊超平面來最佳分割正負類別，考慮一組資料樣本集合 $T = \{(\mathbf{x}_1, y_1), (\mathbf{x}_2, y_2), \dots, (\mathbf{x}_N, y_N)\}$ ，其中 $\mathbf{x}_i \in R^n$ 是第 i 筆資料向量， $y_i \in \{-1, 1\}$ 則是第 i 筆資料所對應的類別，而 $i=1, \dots, N$ 。假設屬於正類別的樣本數目有 N_p 筆，這 N_p 筆正類別資料向量可以用矩陣 $\mathbf{A} \in R^{N_p \times n}$ 表示，同樣地，假設屬於負類別的樣本數目有 N_n 筆，這 N_n 筆負類別資料向量可以用矩陣 $\mathbf{B} \in R^{N_n \times n}$ 表示。要延伸到非線性的模糊分類器，我們可以使用核心函數技巧，先將樣本點經由非線性轉換映射到一個高維度特徵空間(feature space)，然後在特徵空間估計最佳的模糊超平面，它對應到原來空間就是最佳的模糊曲面(fuzzy surface)，考慮下列由核心函數定義的模糊曲面。

$$K(\mathbf{x}^T, \mathbf{C}^T) \widetilde{\mathbf{W}}_1 + \widetilde{\mathbf{B}}_1 = \theta \quad \text{與} \quad K(\mathbf{x}^T, \mathbf{C}^T) \widetilde{\mathbf{W}}_2 + \widetilde{\mathbf{B}}_2 = \theta, \quad (16)$$

其中， $\widetilde{\mathbf{W}}_1 = (\mathbf{w}_1, \mathbf{c}_1)$ 與 $\widetilde{\mathbf{W}}_2 = (\mathbf{w}_2, \mathbf{c}_2)$ 分別是這兩個超曲面的模糊權重向量，而 $\widetilde{\mathbf{B}}_1 = (b_1, d_1)$ 與 $\widetilde{\mathbf{B}}_2 = (b_2, d_2)$ 分別是這兩個超曲面的模糊偏移量。 θ 表示“模糊零(fuzzy zero)”，它也是一個三角形模糊數，以零為中心點，而模糊半徑為 O_w ， $\mathbf{C}^T = [\mathbf{A} \mathbf{B}]^T$ 與 K 是適合的核心函數。要估計一對最佳的模糊曲面，其中一個類別對應一個模糊超曲面，使得每一個超曲面能夠與對應的類別的樣本點能夠距離越近越好，同時要與另一個類別的樣本點能夠距離越遠越好，我們可以求解下列最佳化問題：

$$\begin{aligned}
\min & \frac{1}{2} \|K(\mathbf{A}, \mathbf{C}^T) \mathbf{w}_1 + \mathbf{e}_1 b_1\|^2 + \frac{M_1}{2} (\|\mathbf{w}_1\|^2 + \|\mathbf{c}_1\|^2 + b_1^2 + d_1^2) + C_1 \mathbf{s}_2^T (\xi_{11} + \xi_{12}) \\
s.t. & -[(K(\mathbf{B}, \mathbf{C}^T) \mathbf{w}_1 + \mathbf{e}_2 b_1) + (K(\mathbf{B}, \mathbf{C}^T) \mathbf{c}_1 + \mathbf{e}_2 d_1)] + \xi_{11} \geq \mathbf{e}_2 + \mathbf{e}_2 \mathbf{I}_w, \\
& -[(K(\mathbf{B}, \mathbf{C}^T) \mathbf{w}_1 + \mathbf{e}_2 b_1) - (K(\mathbf{B}, \mathbf{C}^T) \mathbf{c}_1 + \mathbf{e}_2 d_1)] + \xi_{11} \geq \mathbf{e}_2 - \mathbf{e}_2 \mathbf{I}_w, \xi_{11}, \xi_{12} \geq 0
\end{aligned} \tag{17}$$

與

$$\begin{aligned}
\min & \frac{1}{2} \|K(\mathbf{B}, \mathbf{C}^T) \mathbf{w}_2 + \mathbf{e}_2 b_2\|^2 + \frac{M_2}{2} (\|\mathbf{w}_2\|^2 + \|\mathbf{c}_2\|^2 + b_2^2 + d_2^2) + C_2 \mathbf{s}_1^T (\xi_{21} + \xi_{22}) \\
s.t. & [(K(\mathbf{A}, \mathbf{C}^T) \mathbf{w}_2 + \mathbf{e}_1 b_2) + (K(\mathbf{A}, \mathbf{C}^T) \mathbf{c}_2 + \mathbf{e}_1 d_2)] + \xi_{21} \geq \mathbf{e}_1 + \mathbf{e}_1 \mathbf{I}_w, \\
& [(K(\mathbf{A}, \mathbf{C}^T) \mathbf{w}_2 + \mathbf{e}_1 b_2) - (K(\mathbf{A}, \mathbf{C}^T) \mathbf{c}_2 + \mathbf{e}_1 d_2)] + \xi_{21} \geq \mathbf{e}_1 - \mathbf{e}_1 \mathbf{I}_w, \xi_{21}, \xi_{22} \geq 0
\end{aligned} \tag{18}$$

QPP(17)與(18)的目標函數的第一項是模糊超平面到該對應類別(例如, 類別 1)的樣本點的距離的總和。因此, 最小化該項等同於確保模糊超平面與對應的類別越接近越好, 最小化 QPP(17)與(18)的目標函數的第二項則是實現結構風險最小化(structure risk minimization)的精神, 它證明同時最小化訓練誤差與模型複雜度(model complexity)等同於最小化推理誤差的機率上界(upper bound), 參數 $M_1 > 0$ 與 $M_2 > 0$ 是用戶設定的調控參數。限制條件則是要求超平面與另一個類別(例如, 類別 2)的樣本點的距離至少為 1, 違反限制條件的誤差將被差額變數(slack variables) $\xi_{11}, \xi_{12}, \xi_{21}$ 與 ξ_{22} 量測, 而參數 C_1 與 C_2 是用戶設定的懲罰參數, 並在 QPP(17)與(18)的目標函數的第三項懲罰這些誤差(被差額變數所量測), 最小化 QPP(17)與(18)的目標函數的第三項等同於最小化訓練樣本的分類錯誤率。每一個訓練樣本都指派一個模糊歸屬度, 代表該筆訓練樣本屬於對應類別的強度, 向量 \mathbf{s}_1 與 \mathbf{s}_2 代表正類別與負類別的樣本點的模糊歸屬度所組成的向量, 並在 QPP(17)與(18)的目標函數的第三項調控對於誤差的懲罰的強度, 對於較為不確定的樣本點(例如, 雜訊), 則指派較低的模糊歸屬程度, 來減少它被分類錯誤時的懲罰強度, 如此可以避免過度學習(overfitting)那些雜訊點, 並且增加推理能力。使用 Lagrangian 理論, 我們得到下列對偶最佳化問題:

$$\begin{aligned}
\max & \begin{cases} \frac{-1}{2} (\boldsymbol{\alpha}_{11} + \boldsymbol{\alpha}_{12})^T \mathbf{R} (\mathbf{S}^T \mathbf{S} + M_1 \mathbf{I})^{-1} \mathbf{R}^T (\boldsymbol{\alpha}_{11} + \boldsymbol{\alpha}_{12}) \\ \frac{-1}{2M_1} (\boldsymbol{\alpha}_{11} - \boldsymbol{\alpha}_{12})^T \mathbf{R} \mathbf{R}^T (\boldsymbol{\alpha}_{11} - \boldsymbol{\alpha}_{12}) + (1 + \mathbf{I}_w) \boldsymbol{\alpha}_{11}^T \mathbf{e}_2 + (1 - \mathbf{I}_w) \boldsymbol{\alpha}_{12}^T \mathbf{e}_2 \end{cases} \\
s.t. & \mathbf{0} \leq \boldsymbol{\alpha}_{11}, \boldsymbol{\alpha}_{12} \leq C_1 \mathbf{s}_2
\end{aligned} \tag{19}$$

與

$$\begin{aligned}
\max & \begin{cases} \frac{-1}{2} (\boldsymbol{\alpha}_{21} + \boldsymbol{\alpha}_{22})^T \mathbf{S} (\mathbf{R}^T \mathbf{R} + M_1 \mathbf{I})^{-1} \mathbf{S}^T (\boldsymbol{\alpha}_{21} + \boldsymbol{\alpha}_{22}) \\ \frac{-1}{2M_1} (\boldsymbol{\alpha}_{21} - \boldsymbol{\alpha}_{22})^T \mathbf{S} \mathbf{S}^T (\boldsymbol{\alpha}_{21} - \boldsymbol{\alpha}_{22}) + (1 + \mathbf{I}_w) \boldsymbol{\alpha}_{21}^T \mathbf{e}_1 + (1 - \mathbf{I}_w) \boldsymbol{\alpha}_{22}^T \mathbf{e}_1 \end{cases} \\
s.t. & \mathbf{0} \leq \boldsymbol{\alpha}_{21}, \boldsymbol{\alpha}_{22} \leq C_1 \mathbf{s}_2
\end{aligned} \tag{20}$$

$$\text{其中 } \mathbf{S} = [K(\mathbf{A}, \mathbf{C}^T) \quad \mathbf{e}_1], \mathbf{R} = [K(\mathbf{B}, \mathbf{C}^T) \quad \mathbf{e}_2] \quad (21)$$

$$\mathbf{u}_1 = [\mathbf{w}_1^T \quad b_1]^T = -(\mathbf{S}^T \mathbf{S} + M_1 \mathbf{I})^{-1} \mathbf{R}^T (\mathbf{a}_{11} + \mathbf{a}_{12}) \text{ 與 } \mathbf{v}_1 = [\mathbf{c}_1^T \quad d_1]^T = \frac{-1}{M_1} \mathbf{R}^T (\mathbf{a}_{11} - \mathbf{a}_{12}) \quad (22)$$

$$\mathbf{u}_2 = [\mathbf{w}_2^T \quad b_2]^T = (\mathbf{R}^T \mathbf{R} + M_2 \mathbf{I})^{-1} \mathbf{S}^T (\mathbf{a}_{21} + \mathbf{a}_{22}) \text{ 與 } \mathbf{v}_2 = [\mathbf{c}_2^T \quad d_2]^T = \frac{1}{M_2} \mathbf{S}^T (\mathbf{a}_{21} - \mathbf{a}_{22}) \quad (23)$$

求解 QPPs (19)與(20)，我們得到模糊權重向量 $\widetilde{\mathbf{W}}_1 = (\mathbf{w}_1, \mathbf{c}_1)$ 與 $\widetilde{\mathbf{W}}_2 = (\mathbf{w}_2, \mathbf{c}_2)$ ，模糊偏移量 $\widetilde{\mathbf{B}}_1 = (b_1, d_1)$ 與 $\widetilde{\mathbf{B}}_2 = (b_2, d_2)$ ，一旦得到模糊曲面，我們根據新進樣本點 $\mathbf{x} \in R^n$ 與哪一個模糊曲面最接近，來決定樣本點 $\mathbf{x} \in R^n$ 是屬於哪一個類別，對於任意樣本點 \mathbf{x}_i ，它與模糊曲面 $K(\mathbf{x}_i^T, \mathbf{C}^T) \widetilde{\mathbf{W}}_1 + \widetilde{\mathbf{B}}_1 = \theta$ 的距離也是一個對稱三角形模糊數，其中心點為 $|K(\mathbf{x}_i^T, \mathbf{C}^T) \mathbf{w}_1 + b_1|$ 與模糊半徑為 $K(\mathbf{x}_i^T, \mathbf{C}^T) \mathbf{c}_1 + d_1$ 。樣本點 \mathbf{x}_i 與模糊曲面 $K(\mathbf{x}_i^T, \mathbf{C}^T) \widetilde{\mathbf{W}}_2 + \widetilde{\mathbf{B}}_2 = \theta$ 的距離也是一個對稱三角形模糊數，其中心點為 $|K(\mathbf{x}_i^T, \mathbf{C}^T) \mathbf{w}_2 + b_2|$ 與模糊半徑為 $K(\mathbf{x}_i^T, \mathbf{C}^T) \mathbf{c}_2 + d_2$ 。要決定新進樣本點 \mathbf{x}_i 是屬於哪一個類別，我們必須要判別它跟哪一個模糊曲面最接近，亦即是說，我們必須比較兩個對稱三角形模糊數的大小。對於任意兩個對稱三角形模糊數 $A = (m_A, c_A)$ 與 $B = (m_B, c_B)$ ，模糊數 A 大於模糊數 B 的程度可以使用下列歸屬函數來計算：

$$R_{\geq B}(A) = R(A, B) = \begin{cases} 1 & \text{if } \alpha > 0 \text{ and } \beta > 0 \\ 0 & \text{if } \alpha < 0 \text{ and } \beta < 0 \\ 0.5 \left(1 + \frac{\alpha + \beta}{\max(|\alpha|, |\beta|)} \right) & \text{o.w.} \end{cases}, \quad (24)$$

其中 $\alpha = (m_A + c_A) - (m_B + c_B)$ 與 $\beta = (m_A - c_A) - (m_B - c_B)$ ，令對稱三角形模糊數 $\widetilde{D}_1 = (|K(\mathbf{x}_i^T, \mathbf{C}^T) \mathbf{w}_1 + b_1|, K(\mathbf{x}_i^T, \mathbf{C}^T) \mathbf{c}_1 + d_1)$ 代表模糊曲面 $K(\mathbf{x}_i^T, \mathbf{C}^T) \widetilde{\mathbf{W}}_1 + \widetilde{\mathbf{B}}_1 = \theta$ 與樣本點 \mathbf{x}_i 的模糊距離。令對稱三角形模糊數 $\widetilde{D}_2 = (|K(\mathbf{x}_i^T, \mathbf{C}^T) \mathbf{w}_2 + b_2|, K(\mathbf{x}_i^T, \mathbf{C}^T) \mathbf{c}_2 + d_2)$ 代表模糊曲面 $K(\mathbf{x}_i^T, \mathbf{C}^T) \widetilde{\mathbf{W}}_2 + \widetilde{\mathbf{B}}_2 = \theta$ 與樣本點 \mathbf{x}_i 的模糊距離，我們的模糊支持向量機的決策函數為：

$$f(\mathbf{x}) = R_{\geq \widetilde{D}_1}(\widetilde{D}_2) = R(\widetilde{D}_2, \widetilde{D}_1). \quad (25)$$

此模糊決策函數傳回值是界於0與1之間，代表樣本點 \mathbf{x} 屬於正類別的模糊歸屬程度。在正類別與負類別之間存在一個模糊的邊界，此模糊邊界更能夠處理現實中不精確與模糊的特性。值得注意的是，模糊決策函數 $f(\mathbf{x})$ 也提供了對於預測結果的信心程度， $f(\mathbf{x})$ 的傳回值越大，代表 \mathbf{x} 屬於正類別的可能性越強，對於預測結果的信心程度也就越大。對於決策制定任務，能夠提供預測結果的信心程度是很重要的優點(Hao 2021b)。

五、透過轉移學習建構的深層模糊支持向量機網路

受到 Wang et al. (2019) 發表的深層支持向量機所啟發，在本研究中，我們導入轉移學習(transfer learning)理論來建立創新的具有深層結構的模糊支持向量機。在給定源域(source domain, DS)及其學習任務(source task, TS)，目標域(target domain, DT)及其學習任務(target task, TT)，轉移學習的目標是當 $DS \neq DT$ 或 TS

≠ TT 時，透過利用 DS 中的知識，來幫助改進 DT 中的學習過程。在我們的研究中，我們透過堆疊方式建構深層網路結構，堆疊結構中的上一層的模糊支持向量機網路模組視為源域 DS，下一層的相鄰模糊支持向量機網路模組視為目標域 DT，相鄰的模糊支持向量機網路模型具有相同的原始特徵空間，並且將上一層模糊支持向量機網路模組的預測結果，視為附加特徵，將它與原始特徵串接在一起，以反映有關原始特徵空間中的股價估計訊息。亦即，DS 和 DT 之間存在一定程度的一致性，透過轉移學習，我們可以使用先前從源域 DS 中學習到的知識，來幫助目標域 DT 中的學習過程。此策略屬於轉移學習中的模型/參數轉移方法 (model/parameter-transfer approach)，它假設源域和目標域在一定程度上共享參數或預測模型的先驗分佈 (Wang et al. 2019)。

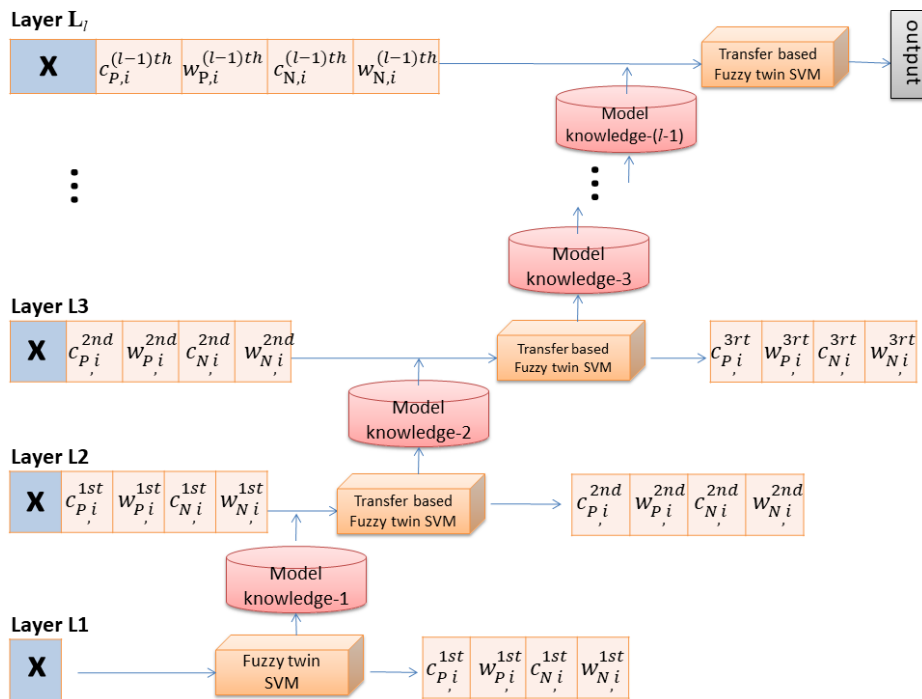


圖 4: 基於轉移學習的深層模糊支持向量機

詳細作法如下，給定訓練資料集 $\{(x_i, y_i)\}, i=1, \dots, N$ ，其中 $x_i \in R^n$ 代表第 i 筆資料向量， y_i 為其對應的訓練目標，我們先使用公式 (19)與(20)的二次最佳化問題求解出第一層的模糊支持向量機，在建立好第一層的模糊支持向量機後，我們將第一層模糊支持向量機的預測結果當作附加向量，跟原始的資料特徵向量串接在一起。舉例來說，假設在第一層中，對於第 i 筆資料點 $x_i = [x_{1i}, x_{2i}, \dots, x_{ni}]^T$ 的預測輸出為 $C_{P,i}^{1st}, W_{P,i}^{1st}, C_{N,i}^{1st}, W_{N,i}^{1st}$ ，其中， $C_{P,i}^{1st}$ 與 $W_{P,i}^{1st}$ 分別代表第 i 筆資料點 x_i 與正類別模糊曲面的模糊距離的中心點與寬度， $C_{N,i}^{1st}$ 與 $W_{N,i}^{1st}$ 分別代表第 i 筆資料點 x_i 與負類別模糊曲面的模糊距離的中心點與寬度，然後，我們將預測結果與原來的資料特徵向量串接在一起，亦即是 $x_i^{2nd} = [x_{1i}, \dots, x_{ni} | C_{P,i}^{1st}, W_{P,i}^{1st}, C_{N,i}^{1st}, W_{N,i}^{1st}]$ ，接著，對所有訓練樣本執行上述擴充輸入特徵向量的動作，我們將得到 N 筆新的訓練資料 $x_1^{2nd}, x_2^{2nd}, \dots, x_N^{2nd}$ ，他們對應的訓練目標 y_1, y_2, \dots, y_N 皆維持不變。這新的 N 筆訓練資料將餵給模糊支持向量機訓練，堆疊(stack)建構第二層的網路結構，同

時，為了將上一層學習的知識轉移到下一層，所以第二層以後的模糊支持向量機(命名為 transfer-based fuzzy SVM)的最佳化問題改寫如下：

$$\begin{aligned} \min \frac{1}{2} & \left\| K(\mathbf{A}^{lth}, \mathbf{C}^{lthT}) \mathbf{w}_{1,T} + \mathbf{e}_1 b_{1,T} \right\|^2 + \frac{M_1}{2} \left(\|\mathbf{w}_{1,T}\|^2 + \|\mathbf{c}_{1,T}\|^2 + b_{1,T}^2 + d_{1,T}^2 \right) \\ & + C_1 \mathbf{s}_2^T (\xi_{11} + \xi_{12}) + \|\mathbf{w}_{1,T} - \mathbf{w}_{1,S(l-1)}\|^2 + \|b_{1,T} - b_{1,S(l-1)}\|^2 \\ & + \|\mathbf{c}_{1,T} - \mathbf{c}_{1,S(l-1)}\|^2 + \|d_{1,T} - d_{1,S(l-1)}\|^2 \\ \text{subject to} & \end{aligned} \quad (26)$$

$$\begin{aligned} - & \left[\left(K(\mathbf{B}^{lth}, \mathbf{C}^{lthT}) \mathbf{w}_{1,T} + \mathbf{e}_2 b_{1,T} \right) + \left(K(\mathbf{B}^{lth}, \mathbf{C}^{lthT}) \mathbf{c}_{1,T} + \mathbf{e}_2 d_{1,T} \right) \right] + \xi_{11} \geq \mathbf{e}_2 + \mathbf{e}_2 \mathbf{I}_w \\ - & \left[\left(K(\mathbf{B}^{lth}, \mathbf{C}^{lthT}) \mathbf{w}_{1,T} + \mathbf{e}_2 b_{1,T} \right) - \left(K(\mathbf{B}^{lth}, \mathbf{C}^{lthT}) \mathbf{c}_{1,T} + \mathbf{e}_2 d_{1,T} \right) \right] + \xi_{11} \geq \mathbf{e}_2 - \mathbf{e}_2 \mathbf{I}_w, \\ & \xi_{11}, \xi_{12} \geq 0 \end{aligned}$$

與

$$\begin{aligned} \min \frac{1}{2} & \left\| K(\mathbf{B}^{lth}, \mathbf{C}^{lthT}) \mathbf{w}_{2,T} + \mathbf{e}_2 b_{2,T} \right\|^2 + \frac{M_2}{2} \left(\|\mathbf{w}_{2,T}\|^2 + \|\mathbf{c}_{2,T}\|^2 + b_{2,T}^2 + d_{2,T}^2 \right) \\ & + C_2 \mathbf{s}_1^T (\xi_{21} + \xi_{22}) + \|\mathbf{w}_{2,T} - \mathbf{w}_{2,S(l-1)}\|^2 + \|b_{2,T} - b_{2,S(l-1)}\|^2 \\ & + \|\mathbf{c}_{2,T} - \mathbf{c}_{2,S(l-1)}\|^2 + \|d_{2,T} - d_{2,S(l-1)}\|^2 \\ \text{subject to} & \end{aligned} \quad (27)$$

$$\begin{aligned} & \left[\left(K(\mathbf{A}^{lth}, \mathbf{C}^{lthT}) \mathbf{w}_{2,T} + \mathbf{e}_1 b_{2,T} \right) + \left(K(\mathbf{A}^{lth}, \mathbf{C}^{lthT}) \mathbf{c}_{2,T} + \mathbf{e}_1 d_{2,T} \right) \right] + \xi_{21} \geq \mathbf{e}_1 + \mathbf{e}_1 \mathbf{I}_w, \\ & \left[\left(K(\mathbf{A}^{lth}, \mathbf{C}^{lthT}) \mathbf{w}_{2,T} + \mathbf{e}_1 b_{2,T} \right) - \left(K(\mathbf{A}^{lth}, \mathbf{C}^{lthT}) \mathbf{c}_{2,T} + \mathbf{e}_1 d_{2,T} \right) \right] + \xi_{22} \geq \mathbf{e}_1 - \mathbf{e}_1 \mathbf{I}_w \\ & \xi_{21}, \xi_{22} \geq 0 \end{aligned}$$

其中， $\mathbf{w}_{k,T}$, $\mathbf{c}_{k,T}$, $b_{k,T}$ 與 $d_{k,T}$, $k=1,2$ 代表目標域(也就是第 l 層的模糊支持向量機網路模組)的模糊權重參數，而 $\mathbf{w}_{k,S(l-1)}$, $\mathbf{c}_{k,S(l-1)}$, $b_{k,S(l-1)}$ 與 $d_{k,S(l-1)}$, $k=1,2$ 代表源域(也就是第 $l-1$ 層的模糊支持向量機網路模組)的模糊權重參數， $\mathbf{A}^{lth} = [\mathbf{x}_1^{lth}, \mathbf{x}_2^{lth}, \dots, \mathbf{x}_N^{lth}]$ 則是要餵給第 l 層的網路模組的訓練資料，它是將原始資料特徵向量與第 $l-1$ 層的預測結果串接在一起而得到的，亦即是 $\mathbf{x}_i^{lth} = [x_{1i}, \dots, x_{ni} | c_{P,i}^{(l-1)th}, w_{P,i}^{(l-1)th}, c_{N,i}^{(l-1)th}, w_{N,i}^{(l-1)th}]$ ，而在公式(26)與(27)二次最佳化問題中的目標函數內最小化 $\|\mathbf{w}_{k,T} - \mathbf{w}_{k,S(l-1)}\|^2$, $\|\mathbf{c}_{k,T} - \mathbf{c}_{k,S(l-1)}\|^2$, $\|b_{k,T} - b_{k,S(l-1)}\|^2$ 與 $\|d_{k,T} - d_{k,S(l-1)}\|^2$, $k=1,2$ ，便是實踐模型參數轉移學習的策略，它將會把源域(也就是第 $l-1$ 層的網路模組)學習得到的模型參數的知識 $\mathbf{w}_{k,S(l-1)}$, $\mathbf{c}_{k,S(l-1)}$, $b_{k,S(l-1)}$ 與 $d_{k,S(l-1)}$, $k=1,2$ ，轉移到目標域(也就是第 l 層的網路模組)，來幫助它

學習 $w_{k,T}$, $c_{k,T}$, $b_{k,T}$ 與 $d_{k,T}$, $k=1,2$ 。透過這種方式，逐層將知識轉移學習到下一層，在最後一層，我們使用新的 N 筆訓練資料 $x_1^{final}, x_2^{final}, \dots, x_N^{final}$ ，交給模糊支持向量機學習最終的股價預測模型，詳細的流程圖顯示在圖 4。

肆、實驗結果

本研究試圖通過文字挖掘和機器學習技術來預測台灣公司的股價變化趨勢。為了建構能夠預測股票價格漲跌變動的自動分類器，必須要準備一個事先用類別標籤標記的訓練數據集。但是，用戶在社群網路上創建的推文和帖子，並不包含用於訓練用途的類別標籤。為了推測用戶的推文和帖子對股票價格的影響，我們整合文字信息和市場的股票價格，並考慮用戶發布的時間。我們將用戶的推文和帖子對股票價格的影響，定義為用戶發布討論特定公司的推文和帖子後，該公司股票の開盤價和收盤價之間的差額。股價變動趨勢預測是一個經典的分類問題，可分成兩派作法，其中一派作法將預測結果分成{上漲、下跌}兩個類別，另一派作法將預測結果分成{上漲、平盤、下跌}，在本研究中，參考(Hao et al. 2018; Hao et al. 2021a)等研究，我們將預測結果分成{上漲、下跌}兩個類別。

另外，Xu & Cohen(2018)的研究表明，用戶會傾向於熱烈討論股市交易量大的公司股票，相反地，股市交易量較小的股票，則用戶討論的社群帖子數量則較少。因此，本論文針對股市交易量高的股票：台積電與友達(根據台灣證卷交易所²資料顯示，台積電在民國 108 年度的股票成交金額為 2,084,935,959,207，成交筆數為 2,821,457；友達在民國 108 年度的股票成交金額為 117,500,587,969，成交筆數為 2,085,296)，本研究蒐集在 Mobile01 網站(<https://www.mobile01.com/>)的台積電與友達の討論區的推文與帖子，資料蒐集時間從民國 108 年 1 月 2 日至 109 年 1 月 2 日，並且把同一個交易天的所有推文與帖子，合併成為每日用戶社群推文語料檔，表 3 描述本論文蒐集到的社群推文的摘要。

表 3: 資料集描述

統計量	台積電	友達
交易天數	235	235
社群推文數量	7936	7408
上漲比率	54.81%	50.73%
每日社群推文平均數目	33.77	31.52

股價變動趨勢預測任務是一個模糊的分類問題，對於股票價格上漲+3%與上漲+0.0003%的股票樣本，它們屬於{上漲}類別的隸屬程度並不一樣，在訓練過程中，它們扮演的重要程度也應該不同。股票價格上漲+3%的樣本，它屬於{上漲}類別的隸屬程度比較高，所以在訓練過程中，它比較重要，也較不允許分類錯誤。反之，上漲+0.0003%的樣本，則是位於{上漲}和{下跌}類別之間的模糊灰色地帶，它比較可能是雜訊，因此可以適度地允許被分類錯誤。在本論文中，我們運用以下 S 隸屬度函數為每個數據樣本分配模糊隸屬度。

² <https://www.twse.com.tw/zh/page/trading/exchange/FMNPTK.html>

$$\mu(\Delta p_i, a, b) = \begin{cases} 0 & |\Delta p_i| \leq a \\ 2 \left(\frac{|\Delta p_i| - a}{b - a} \right)^2 & a \leq |\Delta p_i| \leq \frac{a+b}{2} \\ 1 - 2 \left(\frac{|\Delta p_i| - b}{b - a} \right)^2 & \frac{a+b}{2} \leq |\Delta p_i| \leq b \\ 1 & |\Delta p_i| \geq b \end{cases} \quad (28)$$

其中 Δp_i 是與第 i 個數據樣本相對應的公司的股價變化，此外，我們設定參數 $a=0$ 和 $b=\text{median}(|\Delta p_i|)$ ， $i=1, \dots, N$ ，亦即是， b 是 $|\Delta p_i|$ 的中位數，根據公式(28)，股票價格波動越大，屬於上升（或下降）類別的隸屬度越大，這些數據樣本在分類器訓練階段所起的作用就越重要。相反，價格波動越小，屬於上升（或下降）類別的隸屬程度就越小。在這種情況下，這些數據的樣本容易是雜訊。因此，在分類器訓練階段，可以適當地忽略這些數據樣本的預測誤差。

在第一個實驗中，本論文比較我們提出的方法與最先進的股票價格漲跌變動分類模型的預測效能，我們比較的對象有：

- Day & Lee (2016)的方法運用深層類神經網路(deep neural networks)為基礎，並且透過文章中的情緒特徵來分類股票價格漲跌的變動。
- Pinheiro & Dras (2017)的方法採用字元層級的語言模型(character-level language model)與循環神經網路(recurrent neural networks, RNN)，並且透過文章中出現的詞彙來分類股票價格漲跌的變動。
- Yun, Sim, & Seok (2019)的方法套用捲積神經網路(convolutional neural network, CNN)為基礎，他們先學習文章中每一個關鍵字的詞嵌入向量，並且將文章中的關鍵字的詞嵌入向量逐一並排，整合成為一個 n 乘 m 的矩陣，其中， n 是文章中的關鍵字數目， m 是詞嵌入向量的維度，並且將此2D文章矩陣視為圖像，餵入在圖像辨識領域獲得重大成功的CNN網路，來分類隔日的股票價格漲跌變動。
- Chong, Han, & Park. (2017)的方法採用深層類神經網路(deep neural network)為基礎，並且透過自動編碼器(autoencoder)來自動學習金融時間序列的關鍵特徵，以分類隔日股票價格的漲跌變動。
- Fischer & Krauss (2018)的方法運用長短期記憶體(Long-Short Term Memory, LSTM)來擷取股價指數時間序列中的長時間依賴性，再運用LSTM來分類隔日的股票價格的漲跌變動。
- Maratkhani et al. (2021)的方法運用2D捲積神經網路(convolutional neural network, CNN)為基礎，為了將股價指數的時間序列資料轉換成2D的圖像資料，他們使用15個技術指標，並且計算它們在15個不同的時間區間的技術指標值，然後，整合這些資料成為一個15乘15的2D圖像，然後餵入至CNN網路，以分類隔日股票價格的漲跌變動。

- Hao et al. (2021a)的方法運用情緒辭典擷取文章中的情緒特徵，並使用潛在狄利克雷分配(Latent Dirichlet allocation, LDA)來擷取文章中潛在的主題特徵，並且將這些文章中情緒與主題特徵餵給模糊學生支持向量機，以分類隔日股票價格的漲跌變動。

表 4 顯示我們的方法與其他深度網路模型的實驗比較的結果，這裡，我們使用經典的十折交叉驗證(ten-folds cross validation)來評估分類的正確率。其中，Day & Lee (2016)、Yun et al. (2019)、Hao et al. (2021a)與 Pinheiro & Dras (2017)的預測模型，均依據社群網站內的文章訊息，來預測隔天股票價格的漲跌變動，此外，Chong et al. (2017)、Maratkhan et al. (2021)與 Fischer & Krauss (2018)的深度模型，均由歷史股票價格指數的金融時間序列，來預測隔天的股票價格的漲跌變動。表格 4 的倒數第二列顯示各個預測模型對於台積電與友達公司股票價格漲跌變動分類正確率的平均值，排名第一的是本論文使用的混和 CNN, BiLSTM, AM 與深度模糊 SVM 的預測模型，其分類平均正確率是 82.1407%，比排名第二的 Maratkhan et al. (2021)的模型高 5.1988%，表格 4 中最後一列是此八種預測模型，對於台積電與友達公司股票價格分類正確率的個別排名的平均值，本論文提出的預測模型對於台積電與友達公司股價漲跌變動的分類正確率皆是排名第 1 名，兩者排名的平均值是 1，Maratkhan et al. (2021)的模型對於台積電公司股價漲跌變動的分類正確率排名是第 2 名，對於友達公司股價漲跌變動的分類正確率排名為第 3 名，兩者的排名平均值是 2.5。

表 4: 使用不同深度網路模型的股價漲跌預測正確率

	Day & Lee (2016)	Pinheiro & Dras (2017)	Yun et al. (2019)	Chong et al. (2017)	Fischer & Krauss (2018)	Maratkhan et al. (2021)	Hao et al. (2021a)	本研究的 模型
台積電	62.9167	66.6667	67.6000	69.1667	70.1250	70.4167	71.2500	76.9667
友達	61.4285	65.4286	63.5238	81.2381	81.5418	83.2381	81.9048	87.0856
正確率平均	62.1726	66.0476	65.5619	75.2024	75.8334	76.8274	76.5774	82.0262
排名平均	8	6.5	6.5	5	4	2.5	2.5	1

如表格 4 所示，本論文使用的模型在正確率平均值與排名平均值，均比其他最先進的股價預測模型表現的更優異，顯示本論文提出的股價漲跌預測模型的有效性。實驗結果顯示，由於股價資料的複雜性，Day & Lee (2016)、Pinheiro & Dras (2017)、Yun et al. (2019)、Chong et al. (2017)、Fischer & Krauss (2018)、Maratkhan et al. (2021) 與 Hao et al. (2021a)提出的預測方法僅使用單一個機器學習模型，是難以獲得令人滿意的預測結果。相反地，本論文提出的方法混和了 CNN、BLSTM、AM 與深層模糊學生 SVM，其中 CNN 可以擷取出股價資料中的「非時變(time-invariant)」特徵，而 BiLSTM 可以發掘這些「非時變」特徵的前向與後向因果關係，另外，AM 可以聚焦在重要的特徵，最終，深層模糊 SVM 則同時具有優異的推理能力與處理雜訊資料的能力，因此，本研究提出的混和模型可以獲得更加優異的效果。此外，股價漲跌受到許多因素的影響，包含新聞事件、社會大眾的情緒與過去歷史股價資訊等等，Day & Lee(2016)、Pinheiro & Dras (2017)、Yun et al. (2019)、Chong et al. (2017)、Fischer & Krauss (2018)、Maratkhan et al. (2021) 與 Hao et al. (2021a)提出的預測方法皆僅考慮單一個訊息來源(社群推文的文字資訊

或者歷史股價的數值資訊)，因此難以獲得令人滿意的成果。相反地，本研究提出的方法同時考慮社群推文的文字資訊與歷史股價的數值資訊，所以可以得到更佳優異的預測正確率。此外，那些以深度學習網路演算法為基礎的預測模型，並沒有考慮到現實世界的模糊性，股價漲跌變動本質上是一個模糊分類任務，漲 3% 與漲 0.03% 的樣本雖然皆屬於{上漲}的類別，但是它們屬於{上漲}的隸屬程度並不相同，所以在建立分類器來描述{上漲}的類別時，它們的影響力應該並不相同。而本論文使用的深層模糊學生 SVM 因為能夠考慮訓練樣本的模糊性質，並且允許不同模糊度的訓練樣本在建立預測模型時擁有不同程度的影響力，所以可以比那些深度學習網路模型擁有更好的分類正確率。實驗結果顯示，導入模糊超平面的模糊學生 SVM 很適合處理現實世界充滿雜訊資料的應用問題，而且比原始 SVM 更能夠避免雜訊的不良影響，因此可以獲得更優異的預測性能。

表 5:針對台積電與友達資料集的消融分析測試

	台積電	友達
移除 CNN	71.1667 ↓	81.5238 ↓
移除 BiLSTM	70.9167 ↓	82.6824
移除 AM	75.0326	85.2529
移除 twin FSVM	73.1250	84.8231
移除深層轉移學習架構	75.6667	85.2337

註：其中‘↓’代表預測效能明顯下降(超過 5%)。

本研究提出一個混和 CNN、BiLSTM、AM 與深層模糊學生支持向量機的股價預測模型，並且融合各項模組的優點，為了進一步驗證本研究使用的每個模組的重要性，在第二個實驗當中，我們使用消融分析(Ablation Analysis)來驗證各項關鍵模組對於股價預測的貢獻。消融分析是機器學習中常見的技巧，它通過刪除某個關鍵模組並且評估僅使用剩餘組件的預測效能，以了解該被刪除的關鍵模組對於整個系統預測的貢獻程度，簡單地說，如果刪除某個關鍵模組之後，預測準確率出現明顯的下降，代表該被刪除的組件對於預測效能有很大的貢獻。為了驗證 CNN、BiLSTM、AM、模糊學生 SVM 與深層轉移學習架構對於預測效能的貢獻，接下來，我們分別刪除其中一個模組，並且測驗使用剩餘模組的預測準確率，以得到被刪除模組的貢獻程度，實驗結果顯示在表 5。

- **移除 CNN:** 如表 5 所示，移除卷積神經網路(CNN)模組後，對於台積電股價的預測效能下降了 5.8%，對於友達股價的預測效能下降了 5.5618%，由此可見，CNN 模組對於預測準確率有顯著的貢獻，主要原因應該是 CNN 可以對每日的社群推文與股價指數擷取更緊密的「位置無關」特徵，以便後續的模組能夠有效的分析時間序列的前向/後向因果關係，注意重要的特徵並且準確地預測股價漲跌。
- **移除 BiLSTM:** 如表 5 所示，移除雙向長短期記憶體(BiLSTM)模組後，對於台積電股價的預測效能下降了 6.05%，對於友達股價的預測效能下降了 4.4032%，由此可見，BiLSTM 模組對於預測準確率有顯著的貢獻，主要原因應該是 BiLSTM 可以分析時間序列的前向/後向因果關係，而這些隱藏在時間序列中的因果關係，正是決定股價漲跌的關鍵因素。

- **移除 AM:** 如表 5 所示，移除注意力機制(AM)模組後，對於台積電股價的預測效能下降了 1.9341%，對於友達股價的預測效能下降了 1.8327%，AM 模組對於預測準確率的貢獻並不顯著，但是 AM 確實可以些微的提升預測準確率，主要原因應該是 AM 可以對於影響股價漲跌的關鍵因素給予更多的注意力，而適當的忽略較不重要的因素。
- **移除 twin FSVM:** 在此處，我們將原來模型中的模糊學生支持向量機(twin FSVM)刪除，並且使用傳統的單層全連結神經網路(single-layer full-connected neural network)模組取代，如表 5 所示，移除模糊學生支持向量機(twin FSVM)模組後，對於台積電股價的預測效能下降了 3.8417%，對於友達股價的預測效能下降了 2.2625%，由此可見，twin FSVM 模組對於預測準確率有顯著的貢獻，主要原因應該是 CNN 與 BiLSTM 等深度網路模型雖然很擅長從時間序列中找到關鍵的特徵，但是他們並不適最佳的分類器，相反地，支持向量機雖然沒有自動學習特徵的能力，但是在給定良好特徵的條件下，基於統計學習理論的支持向量機被證明是最佳分類器。除此之外，模糊學生支持向量機融合了模糊集合理論，它使用一個模糊超平面來更有效的減少雜訊的不良影響，並且允許不同的訓練樣本在建立分類器時擁有不同的影響力，因此特別適用於股價預測的模糊分類問題。
- **移除深層轉移學習架構:** 在此處，我們刪除透過轉移學習建立深層模糊學生支持向量機的架構，而僅使用單一層的模糊學生支持向量機，如表 5 所示，移除深層轉移學習架構後，對於台積電股價的預測效能下降了 1.3%，對於友達股價的預測效能下降了 1.8519%，由此可見，透過轉移學習策略建立的深層模糊學生支持向量機架構對於預測準確率的貢獻並不顯著，但是該深層架構確實可以些微的提升預測準確率，主要原因應該是因為透過逐層堆疊，將上一層學習得到的知識(預測模型中的學習參數)轉移到下一層的預測模型，可以更加精煉的學習到能夠幫忙預測股價漲跌的關鍵知識。

從上述消融分析測試結果顯示，本研究提出的每一種關鍵模組，都對於預測股價漲跌擁有一定程度的貢獻。

在最後一個實驗中，本研究比較僅採用社群推文的文字資訊、僅採用歷史股價指數的技術指標特徵、以及混和兩種類型的特徵，在針對台積電與友達公司股票價格漲跌趨勢的分類正確率，表 6 顯示實驗比較的結果。從表 6 結果可以發現，對於台積電公司的股票價格預測，根據社群推文捕獲的文字特徵，比股價指數特徵擁有較高的分類正確率；相反地，對於友達公司的股票價格預測，則是使用股價指數為基礎的技術指標特徵，比社群文字特徵擁有較高的分類正確率，這意味在本實驗當中，以社群文字為基礎的高階語義特徵，與以股價指數為基礎的技術指標特徵，兩者之間各有其擅長之處，並沒有誰佔有絕對的優勢，而且彼此可以互補長短，融合股價技術指標特徵與社群推文的文字特徵，其分類正確率是高於單純僅使用其中任一種特徵的正確率，這意味著融合股價的歷史指數訊息與社群推文的文字訊息之後，能夠進一步提升分類準確率。對於台積電公司的股票

價格預測實驗，結合高階文字特徵與股價指數特徵的正確率，比僅使用股價指數特徵的準確率提高 6.7167%，對於友達公司的股票價格預測實驗，結合高階文字特徵與股價指數特徵的正確率，比僅使用股價指數特徵的準確率亦提高 3.1506%，這表示社群網站上推文的文字內容，提供了預測股價行為的重要訊息。

表 6：使用股價指數特徵、社群文字特徵、混合特徵(股價指數+社群文字特徵)的預測正確率

	台積電	友達
股價指數特徵	70.2500	83.9352
社群文字特徵	71.667	72.7619
社群文字+股價指數特徵	76.9667	87.0856

伍、研究限制

股價預測的困難在於影響股價漲跌的因素太多，例如國際政治情勢、相關行業的獲益表現、股票基金與外資投信的買賣行為、國家/中央銀行調升或降低基準利率、國際匯率與投資者情緒等等，本研究僅聚焦於考慮社群推文與歷史股價，這是因為股市波動通常是由熱門新聞推動的，社群媒體上的推文則是反映了新聞事件的熱度，以及投資者對於該事件的態度。因此，分析社群推文的文字資訊與股價技術指數的數值資料，為能夠幫助於我們預測未來的股價變化。然而，本研究沒有考慮其他可能影響股價的因素，例如股票基金、外資與投信的買賣行為、國際匯率與基準利率等，這是本研究的研究限制。

本論文提出的預測模型必須挖掘社群網站上的用戶推文與帖子，所以如果某間公司沒有用戶對它在社群網站上進行討論，則本研究的預測模型就會沒有文字資料挖掘而無法進行預測。這也是本研究的研究限制，因此，本論文僅針對在社群網站上有用戶熱烈討論的兩間公司(台積電與友達)來進行挖掘與預測。此外，受限於網路爬蟲剖析文字的能力，本研究侷限在分析 Mobile01 網站的股票討論區的推文與帖子。在未來，我們預期將針對更多的社群網站，蒐集更多間的公司與關於這些公司的大量用戶留言，來進一步驗證社群網站對於股票市場漲跌行為的預測能力。

陸、結論與未來建議

社群網站內豐富的大眾情緒數據的挖掘任務已經獲得學術研究的重視，社群網站內的推文與帖子儼然成為預測隔日股價漲跌變動的熱門來源。本論文目的是使用台灣社群網站內的用戶推文與帖子內容，來預測台積電與友達公司股票價格隔日漲跌趨勢。實驗結果驗證社群網站內容對於的台灣股市漲跌變動的預測能力，雖然社群網站的文字內容並沒有辦法完全取代以股價指數為基礎的技術指標，但是兩者可以截長補短，若是僅挖掘股價指數的歷史數據，則無法得知影響股價漲跌變動的重大事件的發生，舉例來說，一間科技廠商的重大訂單即將取消的消息，並無法從過去幾日的歷史股價指數上面得知，但是卻會在社群網站的金融討論區當中獲得大量的回響與貼文，然而，若是僅挖掘社群網站的用戶推文，則無法查

覺到股票市場交易的操作資訊，例如交易量或融資量的暴跌或暴增。因此，唯有同時挖掘股價指數的歷史時間序列與社群網站的文字訊息，才能夠融合彼此的優點，更進一步的提升股市漲跌分類正確率。根據本研究的實驗，對於台積電公司股票價格漲跌的預測任務當中，結合社群網站的高階文字特徵與股價指數的技術指標特徵的預測準確率是 76.9667%，比僅使用技術指標特徵來預測的準確率增加 6.7167%，對於友達公司的股票價格漲跌的預測任務當中，結合社群網站的高階文字特徵與股價指數的技術指標特徵的預測準確率是 87.0856%，比僅使用技術指標特徵來預測的準確率增加 3.1506%，這表明社群推文與股價歷史指數可以互補，考慮社群網站的文字資訊可以進一步提升股價漲跌預測的效能。

除此之外，實驗結果也顯示股價漲跌變動是一個模糊分類的任務，漲+3%與漲 0.03%的樣本皆屬於{上漲}的類別，但是屬於漲的隸屬程度並不相同，所以在建立分類器來描述{上漲}的類別時，它們的貢獻程度應該不一樣。而本論文提出的模糊學生 SVM 由於能夠考慮訓練樣本的模糊性質，並且使用模糊超平面來處理漲跌類別之間隱諱不明的性質，所以能夠更強健地避免股市趨勢預測中大量雜訊資料的不良影響，並且獲得更好的股價漲跌變動分類正確率。

在社群網站當中，許多用戶會運用大量的網路術語或表情符號，而這些網路術語與表情符號並沒有相對應的詞嵌入向量，這會限制本研究的 CNN 網路模型分析文字資訊的效能，進而影響股價漲跌預測的結果。在未來，我們將會開發網路術語/表情符號的詞嵌入向量學習的系統，使得股價漲跌變動的類別正確率可以再一步提升。

參考文獻

- Cao, B., Zhao, J., Lv, Z., Gu, Y., Yang, P., & Halgamuge, S. K. (2020). Multiobjective evolution of fuzzy rough neural network via distributed parallelism for stock prediction, *IEEE Transactions on Fuzzy Systems*, 28(5), 939-952.
- Chandra, R., Goyal, S., & Gupta, R. (2021). Evaluation of deep learning models for multi-step ahead time series prediction, *IEEE Access*, 9, 83105-83123.
- Chen, W., Jiang, M., Zhang, W.-G., & Chen, Z. (2021). A novel graph convolutional feature based convolutional neural network for stock trend prediction, *Information Sciences*, 556, 67-94.
- Cherati, M. R., Haeri, A., & Ghannadpour, S. F. (2021). Cryptocurrency direction forecasting using deep learning algorithms, *Journal of Statistical Computation and Simulation*, 91(12), 2475-2489.
- Chong, E., Han, C., & Park, F. C. (2017). Deep learning networks for stock market analysis and prediction: Methodology, data representations, and case studies, *Expert Systems with Applications*, 83, 187-205.
- Day, M.-Y. & Lee, C.-C. (2016). Deep learning for financial sentiment analysis on finance news providers, *2016 IEEE/ACM International Conference on Advances*

- in Social Networks Analysis and Mining (ASONAM)*, 1127-1134.
- Fischer, T. & Krauss, C. (2018). Deep learning with long short-term memory networks for financial market predictions, *European Journal of Operational Research*, 270, 654-669.
- Graves, A. & Schmidhuber, J. (2005). Framewise phoneme classification with bidirectional LSTM and other neural network architectures, *Neural Networks*, 18(5), 602-610.
- Hao, P.-Y., Ou, J.-B., Huang, T.-S., Lin, Z.-Y., & Wu, J.-S. (2018). Sentiment and topic analysis on financial news for stock movement prediction by using fuzzy support vector machine, *Journal of Information Management*, 25(4), 363-396.
- Hao, P.-Y., Kung, C.-F., Chang, C.-Y., & Ou, J.-B. (2021a). Predicting stock price trends based on financial news articles and using a novel twin support vector machine with fuzzy hyperplane, *Applied Soft Computing*, 98, 106806.
- Hao, P.-Y. (2021b). Asymmetric possibility and necessity regression by twin support vector networks, *IEEE Trans. on Fuzzy Systems*, 29(10), 3028-3042.
- Haq, A. U., Zeb, A., Lei, Z., & Zhang, D. (2021). Forecasting daily stock trend using multi-filter feature selection and deep learning, *Expert Systems with Applications*, 168, 114444.
- Hinton, E. & Salakhutdinov, R. (2006). Reducing the dimensionality of data with neural networks, *Science*, 313 (5786), 504-507.
- Hochreiter, S. & Schmidhuber, J. (1997). Long short-term memory, *MIT Press*, 9(8), 1735-1780.
- Huang, F. J. & LeCun, Y. (2006). Large-scale learning with SVM and convolutional for generic object categorization, 2006 *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06)*, New York, NY, USA, 284-291.
- Kim, Y. (2014). Convolutional neural networks for sentence classification, *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, Doha, Qatar, 1746-1751.
- Kung, C.-F., Kung, C.-T., Su, C.-M., Hao, P.-Y., & Lin, Y.-J. (2021). Early detection of sepsis utilizing deep learning and fuzzy support vector machine, *Journal of Information Management*, 28(4), 447-478.
- Li, Q., Tan, J., Wang, J., & Chen, H. (2021). A multimodal event-driven LSTM model for stock prediction using online news, *IEEE Transactions on Knowledge and Data Engineering*, 33(10), 3323-3337.
- Maratkhon, A., Ilyassov, I., Aitzhanov, M., Demirci, M. F., & Ozbayoglu, A. M. (2021). Deep learning-based investment strategy: technical indicator clustering and residual blocks, *Soft Computing*, 25, 5151-5161.

- Pinheiro, L. D. S. & Dras, M. (2017). Stock market prediction with deep learning: A character-based neural language model for event-based trading, *Proceedings of Australasian Language Technology Association Workshop*, 6-15.
- Plawiak, P., Abdar, M., & Acharya, U. R. (2019). Application of new deep genetic cascade ensemble of SVM classifiers to predict the Australian credit scoring, *Applied Soft Computing*, 84, 105740.
- Shi, Y., Li, W., Zhu, L., Guo, K., & Cambria, E. (2021). Stock trading rule discovery with double deep Q-network,” *Applied Soft Computing*, 107, 107320.
- Treisman, A. & Gelade, G. (1980). A feature-integration theory of attention, *Cognitive Psychology*, 12(1), 97-146.
- Wang, G., Zhang, G., Choi, K., & Lu, J. (2019). Deep additive least squares support vector machines for classification with model transfer, *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 49(7), 1527-1540.
- Wu, J. M.-T., Li, Z., Srivastava, G., Tasi, M.-H., & Lin, J. C.-W. (2020). A graph-based convolutional neural network stock price prediction with leading indicators, *Software: Practice and Experience*, 51(3), 628-644.
- Xie, L. & Yu, S. (2021). Unsupervised feature extraction with convolutional autoencoder with application to daily stock market prediction, *Concurrency and Computation: Practice and Experience*, 33(16), e6282.
- Xu, Y. & Cohen, S. B. (2018). Stock movement prediction from tweets and historical prices, *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics*, 1970-1979.
- Yan, Y. & Yang, D. (2021a). A stock trend forecast algorithm based on deep neural networks, *Scientific Programming*, 2021(7510641), 1-7.
- Yan, X., Weihan, W., & Chang, M. (2021b). Research on financial assets transaction prediction model based on LSTM neural network, *Neural Computing and Applications*, 33, 257-270.
- Yun, H., Sim, G., & Seok, J. (2019). Stock prices prediction using the title of newspaper articles with Korean natural language processing, *International Conference on Artificial Intelligence in Information and Communication (ICAIIIC)*, 019-021.
- Yun, K. K., Yoon, S. W., & Won, D. (2021). Prediction of stock price direction using a hybrid GA-XGBoost algorithm with a three-stage feature engineering process, *Expert Systems with Applications*, 186, 115716.
- Zhao, W. X., Jiang, J., Weng, J., He, J., Lim, E.-P., Yan, H., & Li, X. (2011). Comparing twitter and traditional media using topic models, *European Conference on Information Retrieval*, 338-349.

附錄

簡單移動平均線 (Simple Moving Average, SMA) 為 N 天的收盤價的平均數值，計算公式如下

$$SMA(N)_t = (c_t + c_{t-1} + \dots + c_{t-N})/N \quad (29)$$

其中 c_t 代表第 t 天的收盤價，SMA 的值域為 $(0, \infty)$ ，以附表 1 為例，第 5 天的 5 日期間的 SMA 是 $\frac{106.54+106.74+106.67+106.97+107.96}{5} = 106.97$ 。

加權移動平均線 (Weighted Moving Average, WMA) 在計算平均價格時，越近期(新)的價格權重會大於較遠期(舊)的價格權重，設定好時間週期(N)後，WMA 計算公式如下：

$$WMA(N)_t = \frac{Nc_t + (N-1)c_{t-1} + \dots + c_{t-N}}{N + (N-1) + \dots + 1} \quad (30)$$

其中 c_t 代表第 t 天的收盤價，WMA 的值域為 $(0, \infty)$ ，以附表 1 為例，第 5 天的 5 日期間的 WMA 是 $\frac{106.54+2 \times 106.74+3 \times 106.67+4 \times 106.97+5 \times 107.96}{1+2+3+4+5} = 107.1806$ 。

指數移動平均線 (Exponential Moving Average, EMA) 的初始值計算方法與簡單移動平均相同，也就是將 N 日間的收盤價的合計除以 N 算出。然後從第 2 日起，以前一日的 EMA+平滑化常數 $\alpha \times$ (當日收盤價-前日 EMA) 算出，平滑化常數為 $\alpha = 2 \div (N+1)$ 。使用此平滑化常數計算，比重就會呈現指數函數性的衰減，第 t 日的 EMA 計算公式如下。

$$EMA(N)_t = [(c_t - EMA(N)_{t-1}) \times \frac{2}{1+N}] + EMA(N)_{t-1} \quad (31)$$

EMA 的值域為 $(0, \infty)$ ，以附表 1 為例，第 5 天的 5 日期間的 EMA 值為 106.97，第 6 天的 5 日期間的 EMA 值為 $[(108.36 - 106.97) \times \frac{2}{1+5}] + 106.97 = 107.43$ 。

赫爾移動平均線 (Hull Moving Average, HMA) 是為了降低移動平均線會延遲的弱點而開發。 N 週期的赫爾移動平均線是以「 $N \div 2$ 日 WMA 乘以 2 倍，然後減掉 N 日 WMA，將所得結果以 N 的平方根日 WMA 平均」計算，詳細公式如下：

$$HMA = WMA((2 * WMA(N/2) - WMA(N)), \text{round}(\text{sqrt}(N))) \quad (32)$$

HMA 的值域為 $(0, \infty)$ 。

三重指數移動平均線 (triple exponential moving average, TEMA) 它是混合單倍、雙倍和三倍指數平滑平均線，它可以提供更小的延遲，詳細計算公式如下：

$$TEMA = (3 \times EMA1) - (3 \times EMA2) + EMA3 \quad (33)$$

其中， $EMA1 =$ 指數移動平均 (EMA)， $EMA2$ 是 $EMA1$ 的指數移動平均， $EMA3$ 是 $EMA2$ 的指數移動平均。 $TEMA$ 的值域為 $(0, \infty)$ 。

平滑異同移動平均線指標 (Moving Average Convergence & Divergence, MACD) 先計算出快線 (n 日 EMA)、慢線 (m 日 EMA)，EMA 是指數移動平均，最常

用的是 12 及 26 日 EMA，接著將長短天期 EMA 的相減，所得出差額就是「差離值 (DIF)」，代表短期 EMA 偏離長期 EMA 的情形。然後，再將 DIF 線再作一次指數移動平均，就會得出「MACD 線」，詳細計算公式如下：

$$MACD(N)_{t+1} = MACD(N)_t + \frac{2}{N+1}(DIFF_t - MACD(N)_{t-1}) \quad (34)$$

其中 $DIFF_t = EMA(12) - EMA(26)$ ， $MACD$ 的值域為 $(-\infty, \infty)$ 。

相對強弱指標 (Relative Strength Index, RSI) 是一種用來評估「買賣盤雙方力道強弱」情況的技術指標，RSI 把價格上漲日視為買方的力量(漲勢)，把價格下跌日視為賣方的力量(跌勢)，並且以價格的漲跌值作為力量的強弱度值，詳細計算公式如下

$$RSI(N)_t = 100 - \frac{100}{1 + (\sum_{i=0}^{N-1} up_{i-1}/N)/(\sum_{i=0}^{N-1} dw_{i-1}/N)} \quad (35)$$

其中， up_t 是第 t 天上漲的價格變化與 dw_t 是下跌的價格變化，RSI 的範圍值在 0 到 100 之間，若 RSI 值超過 50 時表示漲勢的力量大於跌勢的力量，反之則跌勢的力量大於漲勢的力量。以附表 1 為例，第 6 天的 5 日期

間的 RSI 值為 $100 - \frac{100}{1 + ((0.2+0.3+0.99+0.4)/5)/(0.07/5)} = 96.02272$ 。

錢德動量擺動指標 (Chande Momentum Oscillator, CMO) 在計算公式的分子中採用上漲日和下跌日的數據，詳細公式如下

$$CMO = (Su - Sd)/(Su + Sd) \times 100 \quad (36)$$

Su 是在指定時段內的上漲日中的當前收盤價與前天收盤價之差異的總和， Sd 是在指定時段內的下落日中的當前收盤價與前天收盤價之差異絕對值的總和，CMO 震盪的範圍介於 -100 和 +100 之間，以附表 1 為例，第 6 天的 5 日期間的

CMO 值為 $\frac{(0.2+0.3+0.99+0.4)-0.07}{(0.2+0.3+0.99+0.4)+0.07} \times 100 = 92.04545$ 。

威廉指標 (Williams %R, W%R) 依股價的擺動點來度量股票/指數是否處於超買或超賣的現象，詳細計算公式如下：

$$W\%R = \frac{H_N - C_t}{H_N - L_N} \times 100 \quad (37)$$

其中 C_t 代表當日收盤價， H_N 代表過去 N 日間的最高價， L_N 代表過去 N 日間的最低價 $W\%R$ 的值在 -100~0% 的範圍內。以附表 1 為例，第 5 天的 5 日期間的 $W\%R$

值為 $\frac{107.96-107.96}{107.96-106.54} \times 100 = 0$ 。

百分比價格振盪器 (percentage price oscillator, PPO) 是一種動量振盪器，用於衡量快速平均線(m Day EMA)與慢速平均線(n Day EMA)之間的差異，並以百分比的形式表現二者的差異，其計算公式如下：

$$PPO = \frac{m \text{ Day EMA} - n \text{ Day EMA}}{n \text{ Day EMA}} \times 100 \quad (38)$$

其中 EMA 代表指數平均線， PPO 的範圍在 $(-\infty, \infty)$ ，當較短均線位於較長均線以上時， PPO 為正值。當較短均線位於較長均線以下時， PPO 為負值。

商品通道指數(Commodity Channel Index, CCI) 衡量商品價格與其平均統計價格的偏差值，其計算公式如下

$$CCI = \frac{M_t - SM_t}{0.015D_t} \quad (39)$$

其中 $M_t = (C_t + L_t + H_t)/3$ ， $SM_t = \sum_{i=1}^N M_{t-i+1}/N$ 與 $D_t = \sum_{i=1}^N |M_{t-i+1} - SM_{t-i+1}|/N$ ， CCI 指標的波動範圍在 $(-\infty, \infty)$ ，該指數越高表明價格比之均值較高，越低說明價格比之均值過低。以附表 1 為例，第 6 天的 3 日期間的 CCI 值為 $\frac{108.2733 - 107.6089}{0.015 \times 0.44444} = 99.6667$ 。

蔡金資金流量指標 (Chaikin Money Flow Index, CMFI) 是用於在一段時間內衡量資金流量的技術分析指標，其計算公式如下：

$$CMFI = \frac{N \text{ period sum of money flow volume}}{N \text{ period sum of volume}} \quad (40)$$

其中，資金流量乘數(Money flow multiplier) = $\{(Close - low) - (high - close)\} / (high - low)$ 與 資金流量(Money flow volume) = 資金流量乘數(Money flow multiplier) × 該期間的交易量(volume for the period)， $CMFI$ 的值在 1 和 -1 之間。以附表 1 為例，第 6 天的 3 日期間的 $CMFI$ 值為 $\frac{0.265306 \times 75 + 0.980392 \times 77 + 0.317073 \times 74}{75 + 77 + 74} =$

0.525891902。

價格變動率指標 (Rate of change, ROC) 使用當日收盤價與 N 日前收盤價這兩個不同時點之間的價格差異，來反映股票市場變動的快慢程度， ROC 的計算方式如下

$$ROC = \frac{C_t - C_{t-N}}{C_{t-N}} \times 100 \quad (41)$$

ROC 的範圍是 $(-\infty, \infty)$ ，它可以找出股價變動的「成長率」軌跡。以附表 1 為例，第 5 天的 5 日期間的 ROC 值為 $\frac{107.96 - 106.54}{106.54} \times 100 = 1.332$ 。

拋物線停損點轉向指標(Parabolic Stop And Reverse, Parabolic SAR) 提供投資人做「停止並反轉」的參考，其計算公式如下：

$$SAR_{n+1} = SAR_n + \alpha (EP - SAR_n) \quad (42)$$

其中， SAR_n 代表當前周期的 SAR 值， EP 表示上升趨勢中的最高價和下降趨勢中的最低價， α 是加速因子，以 0.02 為起始值，只要價格創新高，則 α 值每次增加 0.02，最高上限值為 0.2，當 α 至升到 0.2 後價格再創新高時， α 值仍以 0.2 代入計算，反之亦然， SAR 值範圍在 $(0, \infty)$ 。

附表 1: 股價資料範例

日期	開盤價 (Open)	最高價 (High)	最低價 (Low)	收盤價 (Close)	交易量 (Volume)	較前一天 漲跌
第 1 天	106.15	106.69	105.84	106.54	76	
第 2 天	106.48	106.88	106.36	106.74	70	+0.2
第 3 天	106.75	106.95	106.65	106.67	72	-0.07
第 4 天	106.67	107.15	106.66	106.97	75	+0.3
第 5 天	106.97	107.97	106.95	107.96	77	+0.99
第 6 天	107.96	108.64	107.82	108.36	74	+0.4

