國立政治大學經濟學系研究所

碩士學位論文

以深度學習模型預測台灣 ETF 價格走勢

國立政治大學
National Chengchi University

指導教授：蔡炎龍 博士

蕭明福 博士

研究生：吳凱華 撰

中華民國一百零八年七月

DOI:10.6814/NCCU201900589

致謝詞

能夠完成此篇論文要感謝的人實在太多了。首先要感謝兩位指導老師蔡炎龍老師與蕭明福老師的耐心指導，除了給予我許多論文上的意見，也讓我有許多發揮的空間。

碩士兩年生活一轉眼就結束，雖然時間快速流淌，但生活也沒有過得潦草，無論未來生活如何，我始終相信人生的道路是不會白走。感謝我的好友們黃之灝、陳逸倫與黃騵禾，在碩士生活中增添許多色彩，感謝你們在學業上提供的許多幫助與資源，讓我能順利的將課程修完，雖然畢業後各奔東西，但也沒忘了在貓空山上喝茶，聊未來談人生的時光。感謝劉緯綸、林亦軒、王瑾毅在生活中提供許多歡樂，也感謝陳財丁給予我許多學習上的意見，讓我不會迷路。

感謝古詩青多年的陪伴，在人生低潮時刻從不缺席，總是給予我鼓勵與依靠，未來日子還需要多多關照。還有我的家人，給予許多的支持鼓勵，兩年前支持我就讀研究所的決定，兩年的支持讓我完成學業。最後要再次感謝蔡炎龍老師與蕭明福老師的用心指導，還有在口試給我許多寶貴意見的口試委員殷壽鏞老師。

緣分太深，只能怪罪於時光太短。兩個夏天也許短暫，但也是人生中求學階段充滿歡笑聲的歲月。

吳凱華 2019/07

國立政治大學 National Chengchi University

DOI:10.6814/NCCU201900589

摘要

交易所買賣基金(Exchange Traded Funds, ETF)有別於個股投資，具有分散風險的特性，是一種追蹤特定股價指數的投資商品，也就是一種將股票指數商品化並長期持有的金融商品。

持有金融商品的目的就是獲利，因此價格或趨勢的預測準確率就變得相當的重要。文獻上實證發現類神經網路較傳統時間序列方法的預測能力高，加上近年機器學習快速發展，本文以類神經網路長短期記憶模型與生成對抗網路為研究方法，建立一個能廣泛運用在台灣非金融類交易所買賣基金的價格與走勢預測。變數除了有收盤價與成交量之外，交易所買賣基金屬於長期持有的商品，產業與總體的變化也是影響行情走勢的重要因素，因此加入匯豐台灣製造業採購經理人指數做為總體變數。此外，為了捕捉總體變數造成的價格影響，加入二十日與四十五日的收盤價移動平均捕捉價格趨勢。

實證結果發現，使用長短期記憶模型具有預測波動較大的台灣非金融類交易所買賣基金之收盤價格能力，而生成對抗網路具有較高的預測漲跌能力，且行情確實為牛市的時候，生成對抗網路也有較高的能力夠捕捉此趨勢。

關鍵字：深度學習、類神經網路、交易所買賣基金

DOI:10.6814/NCCU201900589

目錄

致謝詞 I
摘要 II
目錄 III
表目錄 IV
圖目錄 V
第一章 緒論 1
第一節 研究動機 1
第二節 研究架構 3
第二章 文獻回顧 4
第一節 深度學習時間序列資料應用文獻 4
第二節 深度學習金融價格預測應用文獻 6
第三章 研究方法 8
第一節 模型方法 8
一 長短期記憶模型 8
二 生成對抗網路 10
三 卷積神經網路 12
第二節 活化函數 13
第三節 損失函數 15
第四節 梯度下降法 16
第四章 資料處理 18
第一節 資料與特徵選取 18
第二節 訓練方式 20
第三節 模型參數設定 22
第五章 實證結果 24
第一節 長短期記憶模型實證結果 24
第二節 生成對抗網路實證結果 26
第三節 各模型結果比較 28
一 長短期記憶模型與生成對抗網路預測富邦台灣摩根結果 28
二 長短期記憶模型與生成對抗網路預測富邦台灣采吉 50 基金結果 35
第四節 小結 42
第六章 結論與建議 44
第一節 結論 44
第二節 後續研究與建議 45
參考文獻 46

DOI:10.6814/NCCU201900589

表目錄

表 4-1ETF 詳細資訊.....19
表 4-2 長短期記憶模型架構.....23
表 4-3 生成對抗網路之鑑別網路架構.....23
表 5-1 長短期記憶模型測試資料之均方誤差與五日收盤價漲跌準確率.....24
表 5-2 生成對抗網路測試資料之均方誤差與五日收盤價漲跌準確率.....26
表 5-3 長短期記憶模型與生成對抗模型預測富邦台灣摩根(0057)之均方誤差與五日收盤價漲跌準確率.....28
表 5-4 長短期記憶模型與生成對抗模型預測富邦台灣采吉 50 基金(006208)之均方誤差與五日收盤價漲跌準確率.....35
表 5-5 富邦台灣摩根(0057)與富邦台灣采吉 50 基金(006208)之平均每日成交量、每日成交量標準差、每日價格變動率、每日價格變動率標準差....43

國立政治大學
National Chengchi University

DOI:10.6814/NCCU201900589

圖目錄

圖 3-1 長短期記憶模型單元.....9
圖 3-2 生成對抗網路架構.....11
圖 3-3 Conv1D 神經網路架構.....13
圖 3-4 S 型函數(Sigmoid function).....14
圖 3-5 Hyperbolic tangent (tanh) 14
圖 3-6 Leaky Rectified linear unit (Leaky ReLU) 15
圖 4-1 長短期記憶模型架構.....21
圖 4-2 生成對抗網路架構.....22
圖 5-1 長短期記憶模型預測富邦台灣摩根(0057)價格預測走勢圖.....25
圖 5-2 長短期記憶模型預測富邦台灣采吉 50 基金(006208)價格預測走勢圖.25
圖 5-3 生成對抗網路預測富邦台灣摩根(0057)價格預測走勢圖.....27
圖 5-4 生成對抗網路富邦台灣采吉 50 基金(006208)價格預測走勢圖.....27
圖 5-5 長短期記憶、生成對抗網路與 0057 真實收價格走勢.....29
圖 5-6 長短期記憶、生成對抗網路與 0057 真實收價格走勢之第一區間.....30
圖 5-7 長短期記憶模型與生成對抗網路預測 0057 之收盤價與 0057 真實資料差距之第一區間.....31
圖 5-8 長短期記憶、生成對抗網路與 0057 真實收價格走勢之第二區間.....32
圖 5-9 長短期記憶模型與生常對抗網路預測 0057 之收盤價與 0057 真實資料差距之第二區間.....32
圖 5-10 長短期記憶、生成對抗網路與 0057 真實收價格走勢之第三區間....33
圖 5-11 長短期記憶模型與生成對抗網路預測 0057 之收盤價與 0057 真實資料差距之第三區間.....34
圖 5-12 長短期記憶、生成對抗網路與 006208 真實收價格走勢.....36
圖 5-13 長短期記憶、生成對抗網路與 006208 真實收價格走勢之第一區間..37
圖 5-14 長短期記憶模型與生成對抗網路預測 006208 之收盤價與 006208 真實資料差距之第一區間.....38
圖 5-15 長短期記憶、生成對抗網路與 006208 真實收價格走勢之第二區間..39
圖 5-16 長短期記憶模型與生成對抗網路預測 006208 之收盤價與 006208 真實資料差距之第二區間.....39
圖 5-17 長短期記憶、生成對抗網路與 006208 真實收價格走勢之第三區間..40
圖 5-18 長短期記憶模型與生成對抗網路預測 006208 之收盤價與 006208 真實資料差距之第三區間.....41

DOI:10.6814/NCCU201900589

國立政治大學
National Chengchi University

DOI:10.6814/NCCU201900589

第一章 緒論

第一節 研究動機

股票指數不只是反映股票市場整體好壞，更是各個國家甚至是全球的經濟景氣的領先指標，因此金融商品的價格預測在金融領域或投資中始終是重要的目標之一。

交易所買賣基金(Exchange Traded Funds, ETF)是一種有別於一般股票投資購買特定公司的股票，具有分散風險的特性，可以避免個股因為負面消息而導致投資人的損失，是一種追蹤特定股價指數的投資商品，也就是一種將股票指數商品化並長期持有的金融商品。

我們持有金融商品，最大的目的就是獲利，因此價格或趨勢的預測準確率就變得相當的重要。傳統上，自我迴歸模型(autoregressive model)、移動平均模型(moving-average model)、自迴歸移動平均模型(autoregressive moving average model)都是常被應用在財務分析的時間序列方法。然而，上述的方法都受同質變異的假設拘束，自我迴歸條件異質變異模型(autoregressive conditional heteroskedasticity model, ARCH)、一般化自我迴歸條件異質變模型(generalized autoregressive conditional heteroskedasticity, GARCH)與其衍生的模型放寬了同質變異數的假設，使變異數並非都是同質。

然而，傳統時間序列方法要求資料序列皆為定態且多為線性模型等等限制也可能限制了模型的預測能力。近年來，深度學習發展快速，類神經網路具有強

DOI:10.6814/NCCU201900589

大的資料提取特徵能力。有許多自然語言與語音辨識等處理時間序列資料的類神經網路模型，也被使用在財務與投資的領域。回顧深度學習在財務經濟領域的應用文獻後，發現前人的貢獻分別是建立預測單一金融商品的模型或使用單一金融商品建立多種模型再比較模型之間的預測能力。因此，本文目的，是建立一個可以廣泛運用在台灣非金融類的交易所買賣基金之價格預測的模型，並參考 Xingyu Zhou et al.(2018)中的模型種類,探討長短期記憶模型(Long Short-Term Memory, LSTM)以及使用長短期記憶模型做為生成網路(generator network)與卷積神經網路(convolutional neural network)做為鑑別網路(discriminator network)的生成對抗網路(General Adversarial Network, GAN)預測台灣非金融業 ETF 收盤價的結論是否如同參考論文一樣，生成對抗網路的預測能力優於長短期記憶模型。

台灣產業以製造業為主,因此製造產業的興衰對國內的景氣有重大的影響。因此，本文使用台灣非金融類的 10 檔 ETF 資料訓練以捕捉各個 ETF 的特徵。訓練的資料包括由二十個交易日的收盤價、成交量、二十天移動平均收盤價、四十五天移動平均收盤價及匯豐台灣製造業經理人採購指數，用來預測第二十一個交易日的收盤價。再以均方誤差、每五個交易日的漲跌方向準確率做為比較兩模型預測能力的基準，最後再觀察兩模型的預測結果與真實資料進行比較。也就是說，本文不針對任何一個 ETF 商品建模，而是建立一個可以廣泛運用在台灣非金融類 ETF 的深度學習模型，以進行台灣非金融類 ETF 預測收盤價與五日漲跌方向。

DOI:10.6814/NCCU201900589

第二節 研究架構

本文共分為六個章節，第一章為緒論，說明本文的研究動機與研究架構，第二章為文獻回顧，首先回顧深度學習處理時間序列資料的文獻，再回顧深度學習在金融資料上的應用。第三章為研究方法，說明本文使用的深度學習模型。第四章為資料處理，說明變數選擇、訓練方法以及模型的架構。第五章為實證結果，比較兩個模型預測能力的差異。第六章為結論與建議。

國立政治大學
National Chengchi University

DOI:10.6814/NCCU201900589

第二章 文獻回顧

第一節 深度學習時間序列資料應用文獻

基於科技進步，讓深度學習再度崛起，發展許多語音辨識與自然語言等處理時間序列資料的深度學習模型(John Gamboa,2017)。最著名處理時間序列資料的類神經網路模型之一為Hochreiter and Schmidhuber (1997)為改善普通遞迴神經網路提出長短期記憶模型(Long Short-Term Memory, LSTM)。

然而，循環神經網路需要估計的參數過於龐大，即使有獨立顯示卡運算，訓練模型所需要的計算參數仍相當的多，導致訓練成本過高。因此許多學者致力於降低循環神經網路的計算成本且不減少其預測能力。Cho et al.(2014)提出Gate Recurrent Unit (GRU)減少長短期記憶模型的閘門機制，使得計算成本縮減又不失其表現能力。另外，Victor Campos et al.(2017)認為並非所有的訓練資料對於目標都有幫助，因此在成本函數加上「懲罰項」，鼓勵模型跳過不重要的資訊，即可減少計算成本，也就是減少對結果無意義的輸入，以降低訓練成本。

除了降低訓練時的運算成本之外，隨著序列長度的增加，增加時間序列模型的「記憶性」也是相當棘手的重要問題。許多處理語言模型的文獻以循環神經網路模型與編碼-解碼為模型架構(Minh-Thang Luong et al., 2015)，而專注力機制(attention mechanism)也是現今許多序列模型中相當普遍採用的作法(Bahdanau et al., 2014)，因為它能夠解決Seq2Seq模型(Sutskever et al., 2014)序列長度過長而導致記憶力下降的問題。在處理語音與翻譯問題時，專注力機制的原理在不同時間點給予不同詞彙不同的權重，使得模型在學習訓練階

DOI:10.6814/NCCU201900589

段更具有合理性，也隱含模型具有推理(relational reasoning)的能力。Alex Graves et al. (2014)藉由建立記憶區塊(memory block)模擬范紐曼電腦架構透過專注力機制以輸入、輸出、控制、記憶與邏輯運算。

即使專注力機制可以讓 Seq2Seq 模型的表現能力提升，但循環神經網路的學習成本仍是令許多人所詬病，Ashish Vaswani et al. (2017)擺脫有別於傳統專注力機制需要與循環神經網路搭配訓練的方法，傳統專注力機制需要用到循環神經網路中解碼網路的狀態(states)，而 Ashish Vaswani et al.的專注力機制則是將輸入資料轉換後的高維度向量對自己做專注力，也就是自我專注力機制(self-attention)。此外，Adam Santoro et al.(2018)將自我專注機制結合 Neural Turing Machines 的記憶區塊中，使記憶區塊中記憶之間互相了解彼此的關係，以增加模型推斷的能力。

增強記憶能力也有另一個作法，Trieu H. Trinh1 et al. (2018) 將序列切成許多子序列並在每一個子序列最後一個時間點(timestep)加上「輔助損失」做模型的預處理，所謂的輔助損失其實就是一個解碼網路(decoder network)，要求每個輔助損失還原前面子序列的所有資訊，以確保模型真的有「記住」前面所有的資訊；也就是說，將序列切成很多子序列，每一子序列的最後一個點外接一個網路做解碼，可以將一個序列視為多個自編碼器(autoencoder)的結合，用來減少訓練時所需要花費的時間，且增加模型的記憶能力。

DOI:10.6814/NCCU201900589

第二節 深度學習金融價格預測應用文獻

即使傳統時間序列模型已發展許多年，財務資料的雜訊與波動等特性使得預測金融商品的價格走勢是相當具有挑戰性的任務。深度學習具有高度提取資料特徵的能力，因此近年來也被使用在財務預測上。早期將類神經網路模型運用在金融領域價格預測上，可以回朔至 Takashi Kimoto et al. (1990)預測東證指數。V.V.Kondratenko and Yu.A Kuperin(2003)利用基礎的遞迴神經網路針對不同的外匯商品分別訓練基本循環神經網路模型找出每種貨幣的趨勢。

隨著深度學習領域發展，有越來越多將類神經網路在財務與經濟領域上預測價格走勢。此外，有些文獻實證發現，長短期記憶模型的預測效果較傳統 ARMA 模型表現更好 (Sima Siami Namini et al., 2018)，而 S.E. Yi (2018)等人證實長短期記憶模型加上傳統注意力機制，比起單純長短期記憶模型的表現更加突出。Amin Hedayati Moghaddam et al. (2015)利用多層感知(multi-layer perceptron, MLP)類神經網路架構預估 NASDAQ 指數趨勢。Masaya Abe et al.(2018)發現在日本股票市場使用深層神經網路的效果比淺層神經網路(shallow neural network)表現優異。Xin-Yao Qian(2017)使用多層感知器、捲積神經網路、多層長短期記憶模型和專注力機制長短期記憶模型預測韓國綜合股價指數。Wei Bao et al. (2017)與 Thomas R.(2017)分別驗證了長短期記憶模型加上自編碼器網路對於單存長短期記憶模型在股票上與總體數據上有更好的表現。Xiao Ding et al.(2015)與 Takashi MATSUBARA et al.(2018)將財金經濟新聞當作輸入資料的特徵用以估計股市價格。Mikołaj Bi´nkowski et al. (2018)以傳統時間序列的自我回歸模型，搭配卷積神經網路估計係數的方式建模，處理非同步的時間序列資料。

DOI:10.6814/NCCU201900589

Kang Zhang et al.(2019)與Xingyu Zhou et al.(2018)將近年興起新的演算法生成對抗網路(general adversarial network,GAN)(Ian Goodfellow et al.,2014)使用在金融領域的商品價格預測當中，實證發現生成對抗網路比長短期記憶模型與無記憶性的類神經網路有更好的預測能力。

國立政治大學
National Chengchi University

DOI:10.6814/NCCU201900589

第三章 研究方法

第一節 模型方法

一 長短期記憶模型

當人們在閱讀文章時，我們會根據前文的文字語意來理解當下詞句的真正意義，因此先前發生事物的記憶對於人類做任何事情而言是相當重要。Sepp Hochreiter and Jurgen Schmidhuber(1997)提出了長短期記憶模型(Long Short-Term Memory)具有模擬人類記憶行為的能力。如今，長短期記憶模型已是相當熱門的遞迴神經網路模型之一，被廣泛運用在語音辨識、情感分析等領域。因為長短期記憶模型具有記憶的特性，而金融商品的價格會受到過去的波動所影響，因此長短期記憶模型常被使用在金融預測上。

相對於標準的遞迴神經網路模型，長短期記憶模型的核心在於每一個單元(cell)透過閘門(gate)機制使此模型具有增加或移除前面所得到訊息的能力，並透過單元狀態(cell state)的傳遞將篩選過的訊息傳遞到下一個單元，藉以讓每一個單元都獲得前面的資訊。

長短期記憶模型架構下圖如，取自 colah’s blog：

DOI:10.6814/NCCU201900589

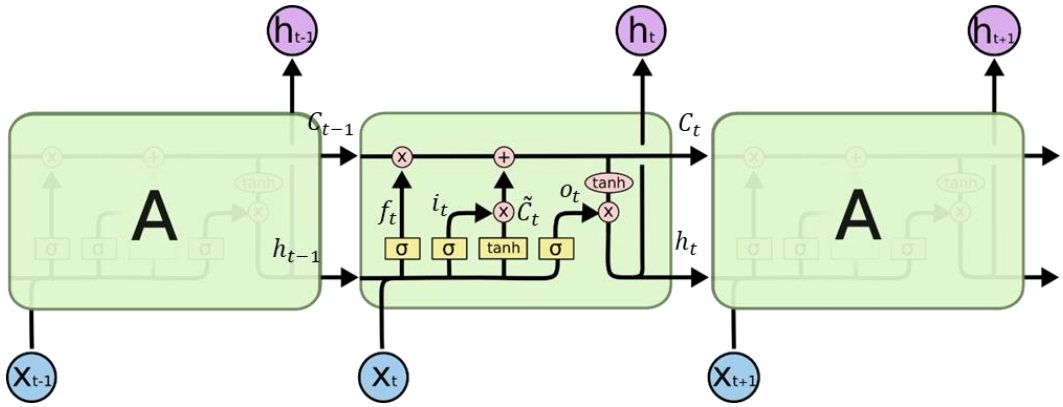

圖 3-1 長短期記憶模型單元

而模型數學式表達如下：

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \quad (1)$$

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \quad (2)$$

$$\tilde{C}_t = \tanh(W_C \cdot [h_{t-1}, x_t] + b_C) \quad (3)$$

$$C_t = f_t \cdot C_{t-1} + i_t \cdot \tilde{C}_t \quad (4)$$

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \quad (5)$$

$$h_t = o_t \cdot \tanh(C_t) \quad (6)$$

其中·代表逐點乘積、W代表模型需要學習的參數、[ ]代表兩個向量的串聯、b代表截距、σ(•)與tanh(•)表示以 S 型函數(Sigmoid function)與 Tangent hyperbolic 作為活化函數。

式子(1)為遺忘閘門，將上一層的隱藏狀態 h 與輸入 x 串聯透過神經層訓練後透過活化函數得出資訊，活化函數使用 S 型函數的目的是為了讓該式子得出的值介於 0 與 1 之間，當作先前資訊的遺忘程度大小並與上一個時間點的單元

DOI:10.6814/NCCU201900589

狀態逐點乘積。下一步要決定輸入的資料訊息，式子(2)是輸入閘，控制此時間點上要輸入的資料，式子(3)則是透過輸入的資料算出一個向量當作新的記憶體單元(cell candidate)，再將式子(2)與式子(3)逐點乘積產生新的單元狀態，並透過式子(4)更新原本的單元狀態。接下來就是透過輸出閘門以輸出資訊，先用式子(5)透過 S 型函數層決定單元狀態的哪一個部分要被輸出，接著單元狀態透過式子(6)透過 Tangent hyperbolic 層讓單元狀態保持在-1 與 1 之間，再與式子(5)算出的向量逐點乘積後輸出要輸出的東西。

標準遞迴神經網路與長短期記憶模型的差別在於，標準神經網路每個時間點的神經元產出的資訊會被放到記憶當中，但到下一個時間點的神經元產出會直接覆蓋上一個時間點的記憶資訊，但是在長短期記憶模型當中，原本的記憶資訊透過閘門機制的處理會保留先前所擁有的資訊，因此可以解決標準遞迴神經網路梯度消失的問題。

二 生成對抗網路

生成對抗網路(generative adversarial network, GAN)(Ian Goodfellow, 2014)是一種無監督式學習的類神經網路，在圖片的生成應用十分廣泛，甚至可以產生肉眼分不出真偽的人像圖片。生成對抗網路由兩個子神經網路組成，一是生成網路(generator)負責產生虛假的資料，二是鑑別網路(discriminator)負責分辨生成模型所產出的資料是否為真，經過反覆的疊代訓練，目的是要讓生成網路能產出以假亂真的資料來騙過鑑別網路。生成網路如同學生，鑑別網路像是老師，監督並修正學生所犯的錯誤直到最佳的情況發生為止。

生成對抗網路每次訓練分成兩個階段：第一階段，固定生成網路，更新鑑別

DOI:10.6814/NCCU201900589

網路的參數；將生成網路產出的假資料與資料集中的資料放入鑑別網路，以判別哪些資料是真是偽。第二階段，固定鑑別網路，更新生成網路參數；生成網路的目的是為了產出以假亂真的資料以混淆鑑別網路，所以在第二階段生成網路學習產出更真實的資料來騙過鑑別網路。

下圖生成對抗網路的模型架構：

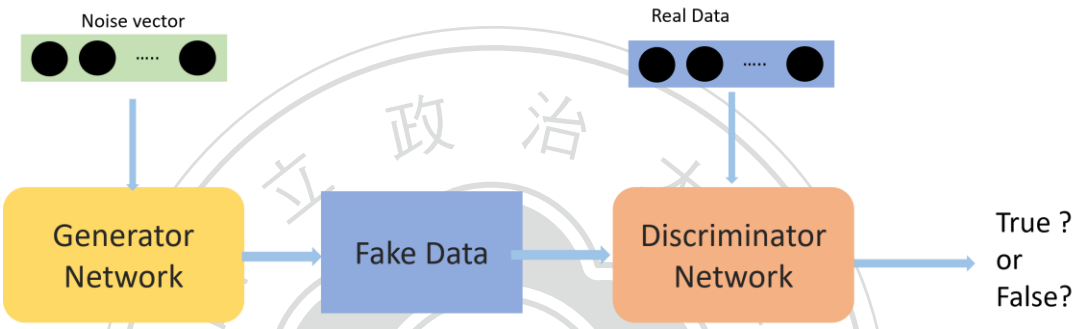

圖 3-2 生成對抗網路架構

生成網路與鑑別網路可以是任何一種類神經網路模型，本文使用的生成對抗模型是參考 Xingyu Zhou et al.(2018)的文章。該文章預測中國四十二家公司的收盤價與漲跌方向，實證發現以長短期記憶模型做為生成網路、卷積神經網路當作鑑別網路的生成對抗網路，比起長短期記憶模型更具有價格與漲跌預測的能力。因此，本文想探討是否長短期記憶網路加上了鑑別模型的監督，並建立廣泛使用於台灣非金融類 ETF 的價格預測上的模型，與長短期記憶模型比較，是否會與上述文章有相同的結果。

DOI:10.6814/NCCU201900589

三 卷積神經網路

卷積神經網路(convolutional neural network, CNN)是一種前饋類神經網路(feed forward network)，在圖像辨識上的應用非常廣泛，卷積神經網路利用核(kernel)進行卷積，模仿人類視覺上依據局部特徵做辨識，同時共享權值也減少了模型要計算的參數數量。核是一個指定尺寸的視窗，如同人類眼鏡的視線範圍，透過核在圖片左而右、上而下移動掃描該區塊圖片的特徵，所謂的卷積就是核對資料提取特徵，並透過逐點乘積後相加所計算的值。

本文的生成對抗網路之鑑別模型採用卷積神經網路，由於金融商品資料具有時間相依性，因此使用 Conv1D。Conv1D 的核只會對一個方向進行卷積，而非由左而右、由上而下的進行特徵提取，核的大小取決於一次要看多少時間點的資料，透過此方式將真實資料判斷為真，將生成網路生成的資料判斷為偽。

圖 3-3 以每時間點維度為 3 的時間序列資料共 6 期、大小為 3 的核為例，透過核從左而右進行掃描，紅色方框與核進行卷積運算：$1\times(-1)+2\times2+3\times(-3)+9\times2+10\times3+11\times1+17\times0+18\times(-1)+19\times(-1)=16$ ，然後核再往右移動至藍色方框進行卷積運算：$2\times(-1)+3\times2+4\times(-3)+10\times2+11\times3+12\times1+18\times0+19\times(-1)+20\times(-1)=18$，以此類推直到掃完所有資料。

DOI:10.6814/NCCU201900589

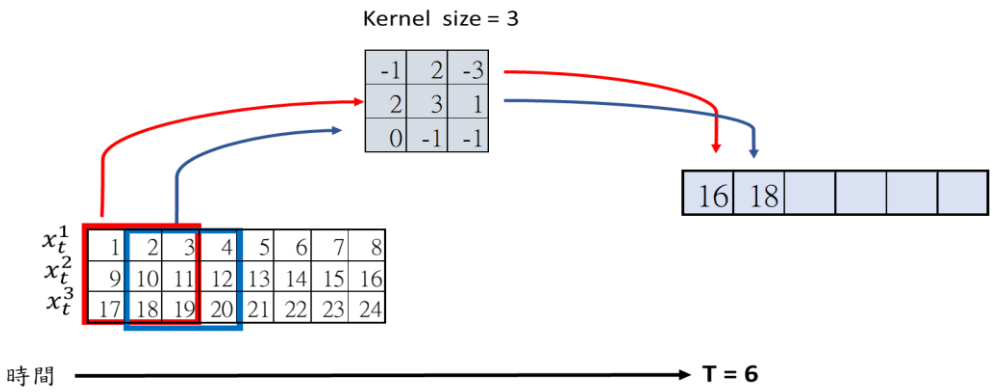

圖 3-3 Conv1D 神經網路架構

第二節 活化函數

傳統時間序列方法的模型大多建立於變數之間的線性關係，而深度學習為了讓類神經網路模型脫離線性關係，我們可以使用活化函數(Activation Function)讓輸出與輸入變成非線性關係，本文使用的活化函數有：S 型函數(Sigmoid function)、Hyperbolic tangent(tanh)、Leaky Rectified linear unit(Leaky ReLU)。

國立政治大學 National Chengchi University

DOI:10.6814/NCCU201900589

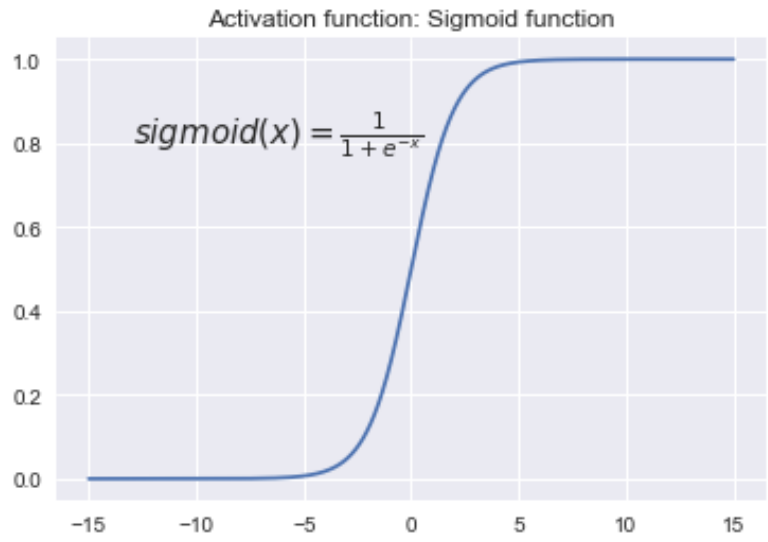

圖 3-4 S 型函數(Sigmoid function)

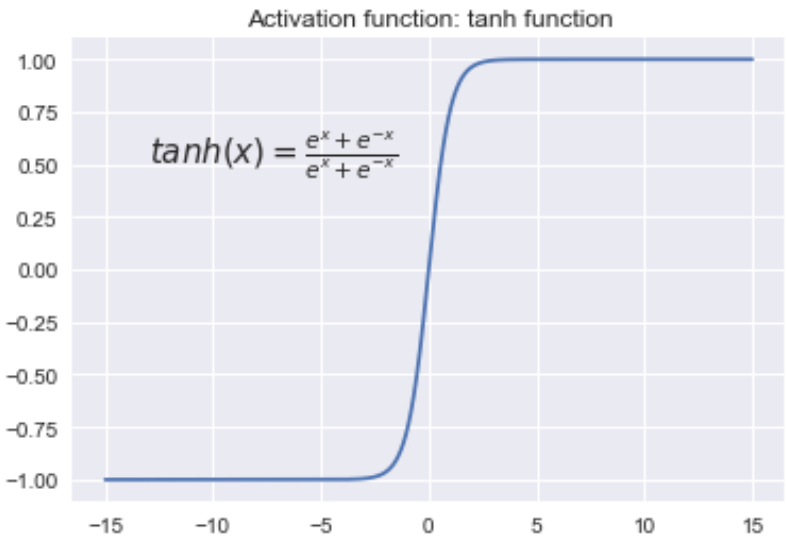

圖 3-5 Hyperbolic tangent (tanh)

DOI:10.6814/NCCU201900589

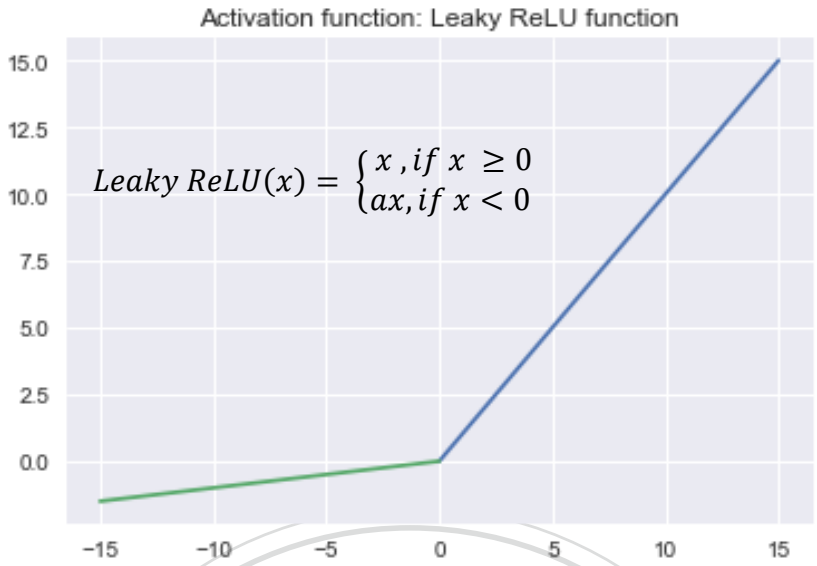

圖 3-6 Leaky Rectified linear unit (Leaky ReLU)

第三節 損失函數

機器學習本質上就是在大量數據中找出一條函數$F_\theta^*(X)$，而我們無法真正找出一條將每一筆輸入資料放入模型中準確得出輸出的函數，因此，我們需要一個損失函數來衡量該模型預估值與實際值的誤差程度。也就是說，給定一組參數θ，當模型得出的預估值與實際值相差程度越小，損失函數值越小，而相差程度越大，損失函數值就越大，我們的目標就是找出最佳化的全局或地區損失最小值。

假設我們透過訓練資料$(x_1, y_1), (x_2, y_2), (x_3, y_3) \dots (x_n, y_n)$得出一條函數$F_\theta(x_i)$，$\theta = (w_1, w_2, \dots w_k, b_h)$，損失函數以$L: R^n \rightarrow R$表示。本文使用的損失函數如下：

DOI:10.6814/NCCU201900589

1. 均方誤差(Mean squared error, MSE):

$$\mathrm{L}(\theta) = \frac{1}{n}\sum_{i=1}^{n}(\mathrm{F}_{\theta}(x_i) - y_i)^2$$

2. 二元交叉熵(Binary cross-entropy):

$$\mathrm{L}(\theta) = -\frac{1}{n}\sum_{i=1}^{n}[y_i \log\big(\mathrm{F}_{\theta}(x_i)\big) + (1 - y_i)\log(1 - \mathrm{F}_{\theta}(x_i))]$$

根據研究目標的不同，模型所設定的損失函數種類也不盡相同。也就是說，我們需要依據研究目的的種類來選擇損失函數的類型。均方誤差主要用在迴歸模型，二元交叉熵使用在二元分類當中。本文的長短期記憶模型預測收盤價格，因此使用均方誤差當作損失函數，生成對抗網路的鑑別網路分辨資料真偽，因此使用二元交叉熵為損失函數。

第四節 梯度下降法

在現實生活中無論從事什麼事情，我們總是追求最佳化的目標，最佳化的目的在於我們希望能夠得到最好的結果。在深度學習中，也不例外要追求最佳化的結果，梯度下降法(gradient descent)是深度學習中最常用來最佳化訓練模型的方法，透過損失函數L求出梯度∇L，以達到最小化損失函數的同時，也得到最好的訓練參數的方法。

下條式子為參數更新的方法：

DOI:10.6814/NCCU201900589

$$W_{t+1} = W_t - \eta * \nabla \mathrm{L}(W_t)$$

其中 $W_t = \begin{pmatrix} w_t^1 \\ w_t^2 \\ \vdots \\ w_t^n \end{pmatrix}$ 為要訓練的參數、$\eta \in [0,1]$ 是學習率、$\nabla \mathrm{L} = \begin{pmatrix} \frac{\partial L}{\partial w_t^1} \\ \frac{\partial L}{\partial w_t^2} \\ \vdots \\ \frac{\partial L}{\partial w_t^n} \end{pmatrix}$ 。

學習率是使用者設定的，介於 0 至 1 之間，學習率的用意在於更新訓練參數的過程中，更新量的大小。若學習率設定太小，可能會使模型無法跳出區域最小值或是需要花更多時間才能找到損失最小值，而設定太大會造成一次更新參數太多，無法找到最小損失。

我們將損失函數最大的值想像成山頂的最高峰，而山底下地勢最低的聚落比喻成損失函數的最小值，我們可以利用山頂附近「向下」的地勢當作下山的路徑，藉由這個路徑我們非常有可能會抵達山下最低點—聚落，而所謂的學習率就是指定我們往地勢向下的方位前進時，每一步要走的距離，所以當學習率過大，表示跨太大步，我們很有可能錯過通往聚落的路徑；當學習率過小，表示步伐過小，需要更多的時間才能抵達目的地。

國立政治大學 National Chengchi University

DOI:10.6814/NCCU201900589

第四章 資料處理

第一節 資料與特徵選取

交易所交易基金(Exchange Traded Funds, ETF)是一種可以在交易所交易的指數股票型基金，此金融商品有追蹤指數、股票型交易與基金的管理模式等特色。台灣以製造產業為主，所以產業的好壞也時常被拿來代表台灣經濟的好壞，本文選擇的 ETF 成分股大多都是由半導體產業與其他二級產業組成，而選擇十檔 ETF 商品的性質相似，某個程度上來說，也就是捕捉整個產業的行情變化。本文選取自台灣交易所交易的十檔 ETF 的收盤價、交易量，選取的交易所交易基金,分別是元大寶來台灣卓越 50(0050)、元大寶來台灣中型 100 (0051)、富邦台灣科技(0052)、元大寶來台灣 ETF 傘型證券投資信託基金之電子科技(0053)、元大寶來台灣高股息(0056)、富邦台灣 ETF 傘型證券投資信託基金之台灣摩根指數(0057)、富邦台灣 ETF 傘型證券投資信託基金之台灣發達指數(0058)、永豐台灣加權(006204)、富邦台灣采吉 50(006208)、兆豐臺灣藍籌 30(00690)，由於各個 ETF 的發售日期不盡相同，因此資料區間最早從 2009/10/1~2019/4/3。

由於本文所選的 ETF 的成分股大都由製造業組成，加上 ETF 屬於長期持有的金融商品，價格走勢會受到產業或總體影響，因此總體因素對於市場仍有重大的影響力，Wei Bao et al.(2017)也將總體變數當作重要特徵之一，因此除了收盤價格與成交量之外，增加匯豐台灣製造業採購經理人指數，以捕捉因總體環境因素對產業發展進而對 ETF 所帶來的價格影響。

另外，技術指標是判斷行情走勢的一種指標，V.V.Kondratenko, Yu. A

DOI:10.6814/NCCU201900589

Kuperin(2003)、Wei Bao et al.(2017)與 Xingyu Zhou et al.(2018)也都將技術指標當作輸入資料的重要特徵。本文加入收盤價的二十天移動平均以捕捉總體指標公布後一個月對價格的影響趨勢。當我們觀察總體指標的時候，並不會只關心當期的公布的數據狀況，還會比較前一期與當期的數據是否有好轉或是變差，而匯豐台灣製造業採購經理人指數並非每月固定日期公布而是在每月的第一個禮拜五公布，造成每次總體指數公布並非相隔二十個交易日，因此加入四十五日的移動平均收盤價以捕捉兩個月匯豐台灣製造業採購經理人指數公布對於 ETF 價格走勢的影響。

本文的資料有五個特徵，分別是收盤價、成交量、收盤價二十日移動平均、收盤價四十五日移動平均與匯豐台灣製造業採購經理人指數，以每二十個交易日的五種變數預測第二十一個交易日的收盤價格。也就是說，每筆訓練資料是一個20 ×5 的矩陣，是由二十個交易日的收盤價、交易量、收盤價二十天移動平均、收盤價四十五天移動平均與匯豐台灣製造業經理人採購指數五個維度所組成，訓練資料的標籤是一個純量，為第二十一個交易日的收盤價格。每個交易日都有五個維度的特徵，然而任何的預測結果都會受限於不同的特徵的數量與選取。

下表格是關於輸入資料的商品代碼、資料區間與追蹤的指數。

表 4-1ETF 詳細資訊

商品代碼	資料區間	追蹤指數
0050	2009/10/01~2019/04/03	臺灣 50 指數
0051	2009/10/01~2019/04/03	臺灣中型 100 指數

DOI:10.6814/NCCU201900589

0052	2009/10/01~2019/04/03	臺灣資訊科技指數
0053	2009/10/01~2019/04/03	電子類加權股價指數
0056	2009/10/01~2019/04/03	臺灣高股息指數
0057	2009/10/01~2019/04/03	MSCI®臺灣指數
0058	2009/10/01~2019/04/03	臺灣發達指數
006204	2011/10/05~2019/04/02	臺灣證券交易所發行量加權股價指數
006208	2012/07/24~2019/04/03	臺灣 50 指數
00690	2017/07/11~2019/04/03	藍籌 30 指數

第二節 訓練方式

類神經網路可以利用大量的訓練樣本，藉由電腦的效能進行大量的運算，從資料之間推斷出隱含的關係。本文欲以類神經網路根據過去二十個交易日的收盤價、成交量、二十日移動平均收盤價、四十五日移動平均收盤價與匯豐台灣製造業採購經理人指數，預測出下一個交易日的收盤價格。

我們隨機將這八筆 ETF 取其中 0050、0051、0052、0053、0056、0058、006204、00690 當作訓練資料集，以捕捉各種 ETF 商品的資料特徵，另外兩筆 0057、006208 當作模型沒看過的測試資料來預測收盤價與五日漲跌方向。也就是說，即使該模型從來沒有看 0057、006208 的資料，也要有預測 0057、006208 收盤價格的能力。

DOI:10.6814/NCCU201900589

本文一筆樣本為二十個交易日的資料預測第二十一天的收盤價格，因此不採用移動窗口法，而是從 15262 筆訓練資料隨機抽取資料進行訓練，也就是隨機梯度下降法(stochastic gradient descent)。為了讓模型更快收斂，將每一筆訓練資料都進行正規化，讓數據維持在 0 與 1 之間，正規化方法如下:

$$X' = \frac{X - \min X}{\max X - \min X}$$

將每一筆訓練資料$X \in R^{20\times5}$的訓練資料放入長短期記憶模型中，並輸出預測值$\hat{Y} \in R^1$，透過每次訓練讓模型產生的預估值$\hat{Y}$更接近於Y。生成對抗網路的生成網路則採用與上述長短期記憶模型相同的放置訓練資料方式，每筆訓練資料放入模型後得到一個收盤價純量。鑑別網路要負責判別生成網路產生的收盤價是否為真實資料，我們將放入生成網路的樣本資料中取出所有的收盤價並與生成網路產出的第二十一個收盤價串接起來形成一個由二十一個收盤價所組成的序列，再將該序列放入鑑別網路訓練並給予標籤為 0 以表示為假資料，接著再放入由二十一個真實收盤價所組成的序列，同樣再放入鑑別網路訓練並給予標籤為 1 以表示為真實資料。

下兩圖是長短期記憶模型與生成對抗網路的架構：

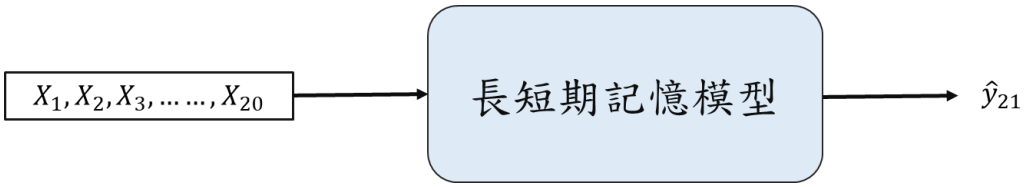

圖 4-1 長短期記憶模型架構

DOI:10.6814/NCCU201900589

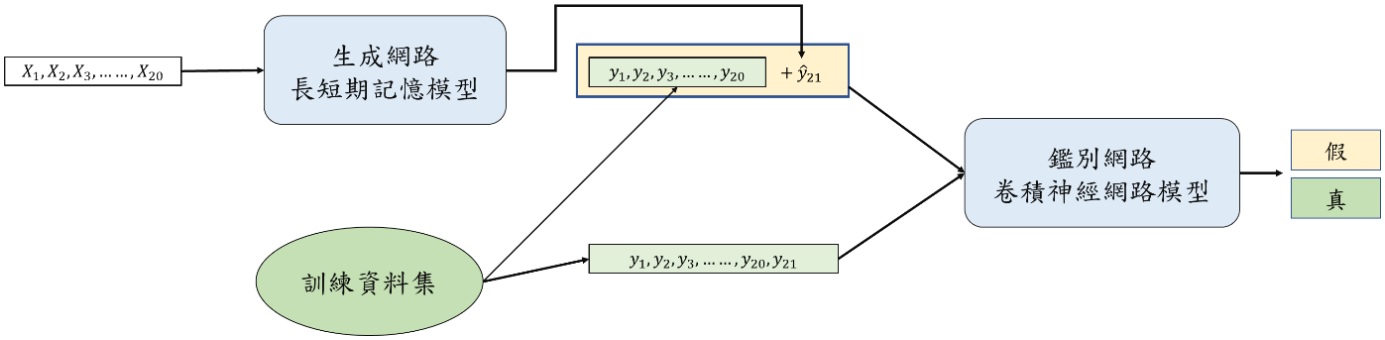

圖 4-2 生成對抗網路架構

第三節 模型參數設定

本研究參考 Xingyu Zhou et al(2018)的模型架構再加以調整。長短期記憶模型採用 121 個神經元，Dropout 為 0.3，通過 S 型函數(Sigmoid function)產生[0,1]之間的值，損失函數為均方誤差，使用 RMSprop 優化，128 筆樣本為一批次，訓練 100 期。

生成對抗網路之生成網路採用與長短期記憶模型一樣的架構，鑑別網路使用一個方向的卷積神經網路，採用三層卷積層，分別為 32, 64, 128 個 filter。核的大小會影響模型進行卷積時所看到的資料長度，本文嘗試大小 2 至 24 的核，最終發現大小為 4 的核表現最好。每個卷積層都會透過活化函數 LeakyReLU 輸出至下一層，最後連結到 128 個神經元再判斷輸入資料是否為真，損失函數使用二元交叉熵，使用 Adam 優化。

DOI:10.6814/NCCU201900589

下表為兩模型的參數設定：

表 4-2 長短期記憶模型架構

層	架構
長短期記憶層	121 神經元、RMSprop、Dropout 0.3
輸出層	1 個神經元、S 型函數(sigmoid)

表 4-3 生成對抗網路之鑑別網路架構

層	架構
卷積層	32 Filters, kernel 大小 4,步伐 1,padding:same, ReakyReLU
卷積層	64 Filters, kernel 大小 4,步伐 1,padding:same, ReakyReLU
卷積層	128 Filters, kernel 大小 4,步伐 1,padding:same, ReakyReLU
全連接層	128 神經元, ReakyReLU
輸出層	2 個神經元、softmax

國立政治大學 National Chengchi University

DOI:10.6814/NCCU201900589

第五章 實證結果

本文將富邦台灣摩根(0057)與富邦台灣采吉 50 基金(006208)作為測試資料，除了使用均方誤差衡量兩個模型在兩測試資料的表現之外，我們再用兩模型得到的預測收盤價格，計算每五日的收盤價格漲跌方向，並與真實資料進行準確率的比較。最後，觀察兩模型預測的收盤價格與真實收盤價格的差異。

第一節 長短期記憶模型實證結果

經過長短期記憶模型訓練 100 期之後，利用長短期記憶模型算出的預測收盤價與真實收盤價計算出均方誤差，再將模型算出的預測收盤價計算每五天收盤價格方向漲跌準確率。富邦台灣摩根(0057)的均方誤差為 0.17439，五日收盤價漲跌準確率為 79.10%，富邦台灣采吉 50 基金(006208)的均方誤差則是 0.18926，五日收盤價漲跌準確率為 79.01%：

表 5-1 長短期記憶模型測試資料之均方誤差與五日收盤價漲跌準確率

測試資料	均方誤差	五日收盤價漲跌準確率
富邦台灣摩根(0057)	0.17439	79.10%
富邦台灣采吉 50 基金(006208)	0.18926	79.01%

DOI:10.6814/NCCU201900589

圖 5-1 長短期記憶模型預測富邦台灣摩根(0057)價格預測走勢圖

圖 5-2 長短期記憶模型預測富邦台灣采吉 50 基金(006208)價格預測走勢圖

DOI:10.6814/NCCU201900589

第二節 生成對抗網路實證結果

生成對抗網路經過 587 期訓練之後，利用生成網路算出的預測收盤價與真實收盤價計算均方誤差，再將生成網路算出的預測收盤價計算每五天收盤價格方向走勢準確率。預測富邦台灣摩根(0057)的均方誤差為 0.17877，五日收盤價漲跌準確率為 79.92%，富邦台灣采吉 50 基金(006208)的均方誤差則是 0.18284，五日收盤價漲跌準確率為 79.86%：

表 5-2 生成對抗網路測試資料之均方誤差與五日收盤價漲跌準確率

測試資料	均方誤差	五日收盤價漲跌準確率
富邦台灣摩根(0057)	0.17877	79.92%
富邦台灣采吉 50 基金(006208)	0.18284	79.86%

國立政治大學 National Chengchi University

DOI:10.6814/NCCU201900589

圖 5-3 生成對抗網路預測富邦台灣摩根(0057)價格預測走勢圖

圖 5-4 生成對抗網路富邦台灣采吉 50 基金(006208)價格預測走勢圖

DOI:10.6814/NCCU201900589

第三節 各模型結果比較

一 長短期記憶模型與生成對抗網路預測富邦台灣摩根結果

本節先根據相同商品進行不同模型之間的均方誤差與五日收盤價方向漲跌準確率進行比較，再透過預測值與真實值來探討兩模型的表現能力。我們將之整理至下表 5-3。在富邦台灣摩根(0057)這檔商品中，長短期記憶模型的均方誤差比生成對抗模型低 0.00447，在五日收盤價格漲跌準確率，長短期記憶模型為 79.10%，生成對抗網路比長短期記憶模型準確度高 0.82%為 79.92%。

表 5-3 長短期記憶模型與生成對抗模型預測富邦台灣摩根(0057)之均方誤差與五日收盤價漲跌準確率

模型	均方誤差	五日收盤價漲跌準確率
長短期記憶模型	0.17430	79.10%
生成對抗網路	0.17877	79.92%

我們進一步計算，當富邦台灣摩根(0057)五日收盤價確實為漲的情況下，長短期記憶模型預測為漲的機率為 80.81%，較生成對抗網路準確率 81.89%低。意味著，當行情真的為牛市的時候，生成對抗網路比長短期記憶模型有較高的機率抓到此波上漲的趨勢。

下圖 5-5 是比較 2009 年 10 月 1 日至 2019 年 4 月 3 日，長短期記憶、生成對抗網路與富邦台灣摩根(0057)真實收價格走勢。

DOI:10.6814/NCCU201900589

圖 5-5 長短期記憶、生成對抗網路與 0057 真實收價格走勢

DOI:10.6814/NCCU201900589

為了清楚呈現兩模型預測值與真實資料之間的差異，我們將測試資料分成三個等區段來探討。下圖 5-6 是長短期記憶、生成對抗網路與富邦台灣摩根(0057)真實收價格走勢之第一區間 2009 年 10 月 1 日至 2012 年 10 月 26 日的比較，可以明顯發現，每當收盤價格因金融事件所造成的巨幅下跌後反彈，生成對抗網路預測反彈後的價格有高估的情況，如 2011 年 3 月至 5 月之間受到日本地震核災與中國調升基準利率抑制通貨膨脹、2011 年 7 月底至 9 月因全球性股災、歐債危機造成的行情巨幅波動。當行情沒有明顯波動的時候，生成對抗網路的預測比長短期記憶模型的預測值更接近原始資料，如 2010 年 10 月左右、2010 年 12 月左右等。圖 5-7 是兩模型預測收盤價與真實收盤價的差距。

圖 5-6 長短期記憶、生成對抗網路與 0057 真實收價格走勢之第一區間

DOI:10.6814/NCCU201900589

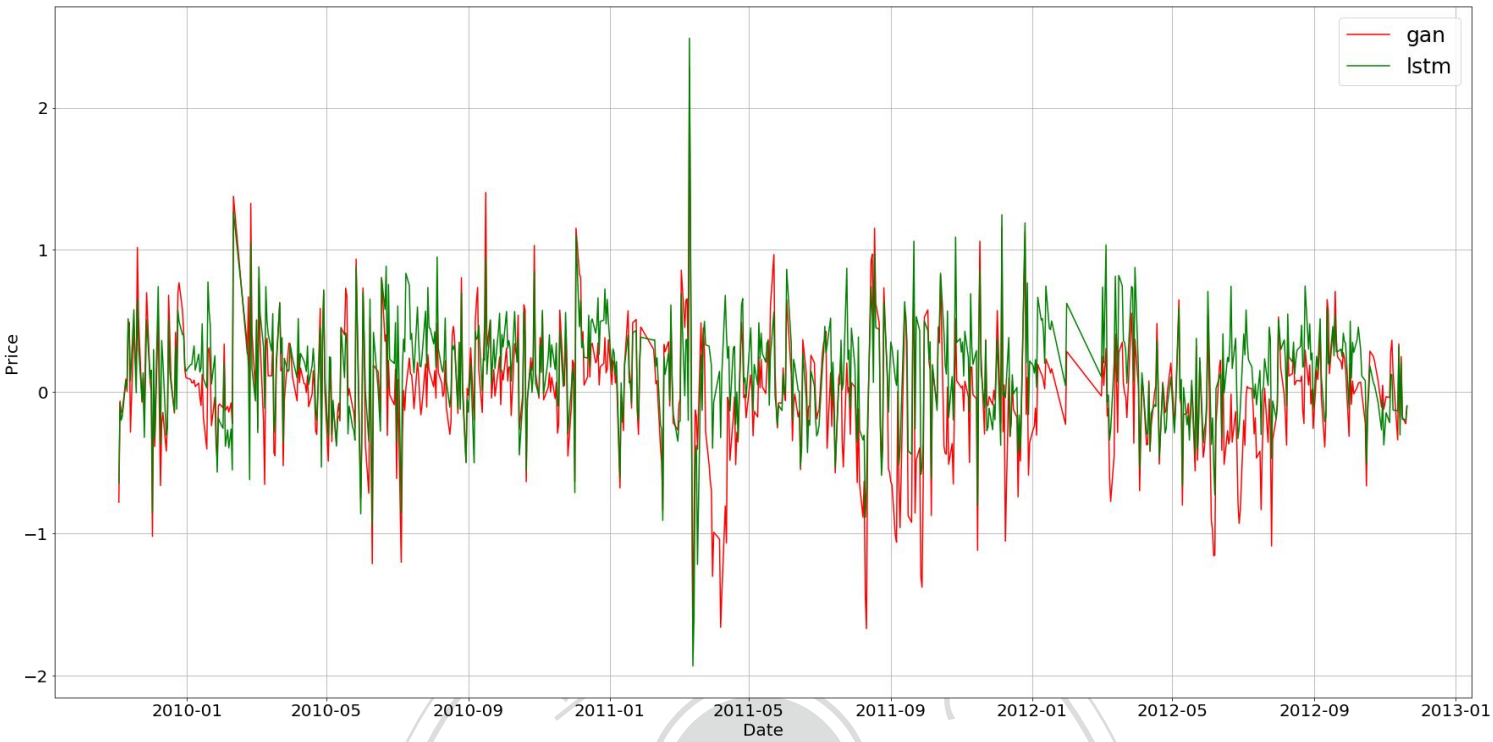

圖 5-7 長短期記憶模型與生成對抗網路預測 0057 之收盤價與 0057 真實資料差距之第一區間

下圖 5-8 長短期記憶、生成對抗網路與富邦台灣摩根(0057)真實收價格走勢之第二區間 2012 年 10 月 26 日至 2015 年 12 月 10 日的比較。可以發現與第一區間的情況，如 2015 年 6 月底至 8 月底因受到中國股市大跌、人民幣貶值、歐美日股市也相繼大跌的影響，富邦台灣摩根(0057)被受牽連下跌後反彈，生成對抗網路的估計值比真實值高。在此區間可以觀察到，在行情較平穩的狀態時，生成對抗網路依舊比長短期記憶模型更接近真實資料，如 2013 年 2 月、2014 年 4 月至 2014 年 5 月、2015 年 1 月。圖 5-9 是兩模型預測收盤價與真實收盤價的差距。

DOI:10.6814/NCCU201900589

圖 5-8 長短期記憶、生成對抗網路與 0057 真實收價格走勢之第二區間

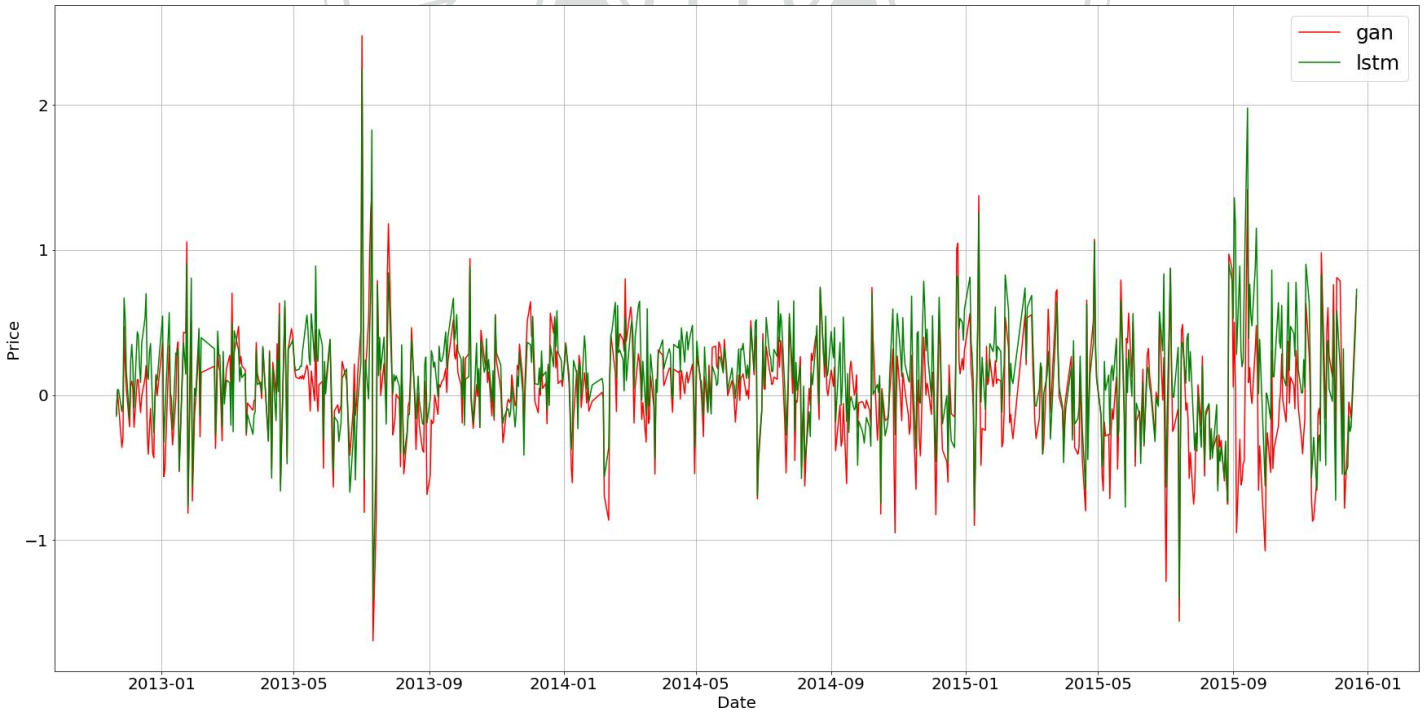

圖 5-9 長短期記憶模型與生常對抗網路預測 0057 之收盤價與 0057 真實資料差

DOI:10.6814/NCCU201900589

距之第二區間

下圖 5-10 長短期記憶、生成對抗網路與 0057 真實收價格走勢之第三區間 2015 年 12 月 10 日至 2019 年 4 月 3 日的比較。2016 年 1 月總統選舉、2018 年 10 月至 12 月初受中美貿易、美國股市與地方選舉影響，下跌後反彈，生成對抗網路價格預估也有偏高的現象。在行情較平穩的狀態時，生成對抗網路依舊比長短期記憶模型更接近真實資料。圖 5-11 是兩模型預測收盤價與真實收盤價的差距。

圖 5-10 長短期記憶、生成對抗網路與 0057 真實收價格走勢之第三區間

DOI:10.6814/NCCU201900589

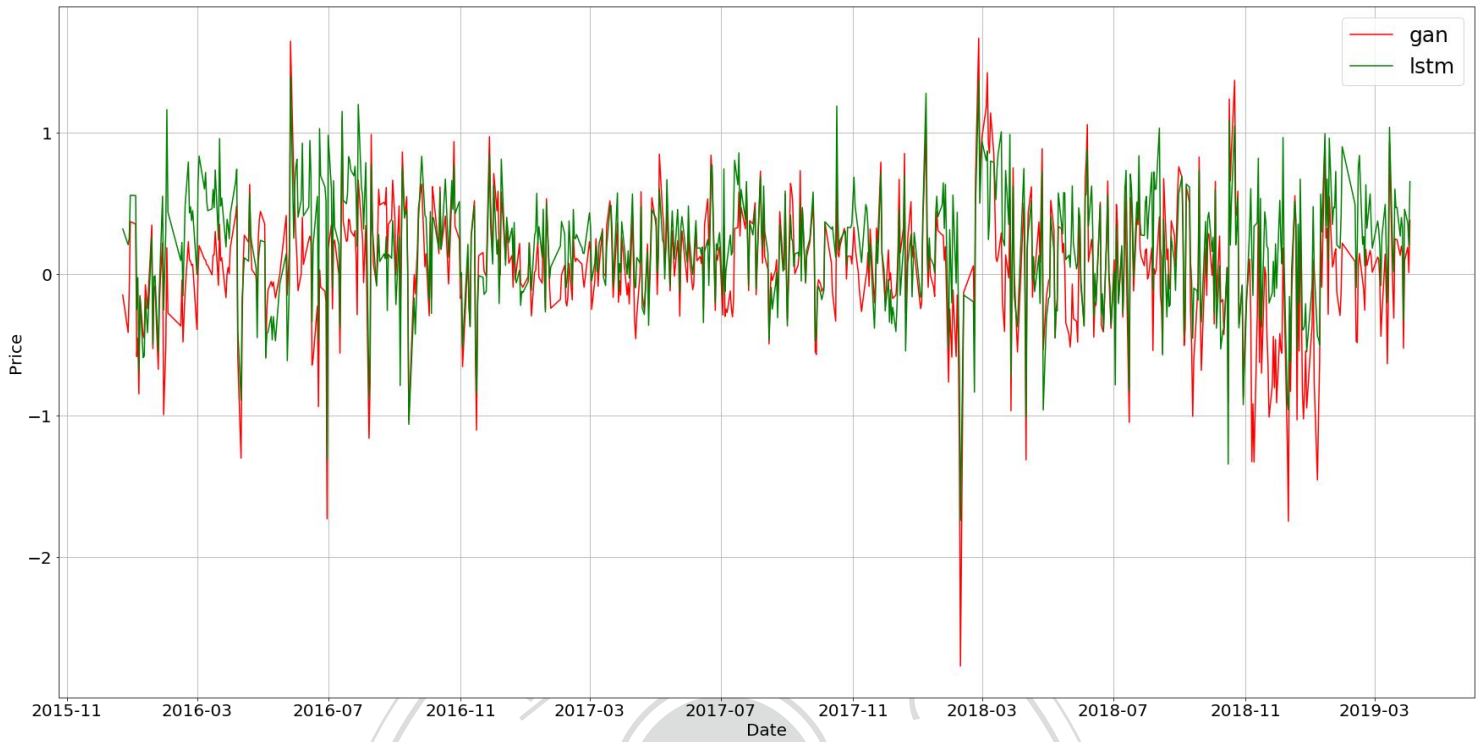

圖 5-11 長短期記憶模型與生成對抗網路預測 0057 之收盤價與 0057 真實資料差距之第三區間

經由富邦台灣摩根(0057)的結果可以得知，在行情沒有巨幅波動的時候，生成對抗網路則較長短期記憶模型更接近真實資料。生成對抗網路易受到大事件所造成的大波動影響，大跌反彈後的價格有高估真實資料的情況。然而長短期記憶模型遇到大波動時表現則較平穩，因此也造成生成對抗網路之均方誤差比長短期記憶模型高出 0.00447 的現象。我們將兩模型經測試資料富邦台灣摩根預測出的收盤價格，分別計算每五日的收盤價漲跌方向準確率，長短期記憶模型預測富邦台灣摩根(0057)每五日漲跌準確率為 79.10%，生成對抗網路則是 79.92%，生成對抗網路有較好的能力預估五日漲跌方向。

DOI:10.6814/NCCU201900589

二 長短期記憶模型與生成對抗網路預測富邦台灣采吉 50 基金結果

我們將整理至下表 5-4，預測富邦台灣采吉 50 基金(006208)這檔商品中，長短期記憶模型的均方誤差比生成對抗模型高 0.00642，在五日收盤價方向的準確度，長短期記憶模型為 79.01%，生成對抗網路比長短期記憶模型準確度高 0.85%為 79.86%。

表 5-4 長短期記憶模型與生成對抗模型預測富邦台灣采吉 50 基金(006208)之均方誤差與五日收盤價漲跌準確率

模型	均方誤差	五日收盤價漲跌準確率
長短期記憶模型	0.18926	79.01%
生成對抗網路	0.18284	79.86%

我們進一步計算，當富邦台灣采吉 50 基金(006208)五日收盤價確實為漲的情況下，長短期記憶模型預測為漲的機率為 82.78%，較生成對抗網路準確率 83.16%低。意味著，當行情真的為牛市的時候，生成對抗網路比長短期記憶模型有較高的機率抓到此波上漲的趨勢。

下圖 5-12 是比較 2012 年 9 月 18 日至 2019 年 4 月 3 日，長短期記憶、生成對抗網路與富邦台灣采吉 50 基金(006208)真實收價格走勢。

DOI:10.6814/NCCU201900589

圖 5-12 長短期記憶、生成對抗網路與 006208 真實收價格走勢

DOI:10.6814/NCCU201900589

為了清楚呈現兩模型預測值與真實資料之間的差異，我們將測試資料分成三個等區段來探討。下圖 5-13 是長短期記憶、生成對抗網路與 006208 真實收價格走勢之第一區間 2012 年 9 月 18 日至 2014 年 9 月 30 日的比較，可以觀察到，大部分生成對抗模型的預測值都比起長短期記憶模型更接近真實資料。圖 5-14 是兩模型預測收盤價與真實收盤價的差距。

圖 5-13 長短期記憶、生成對抗網路與 006208 真實收價格走勢之第一區間

DOI:10.6814/NCCU201900589

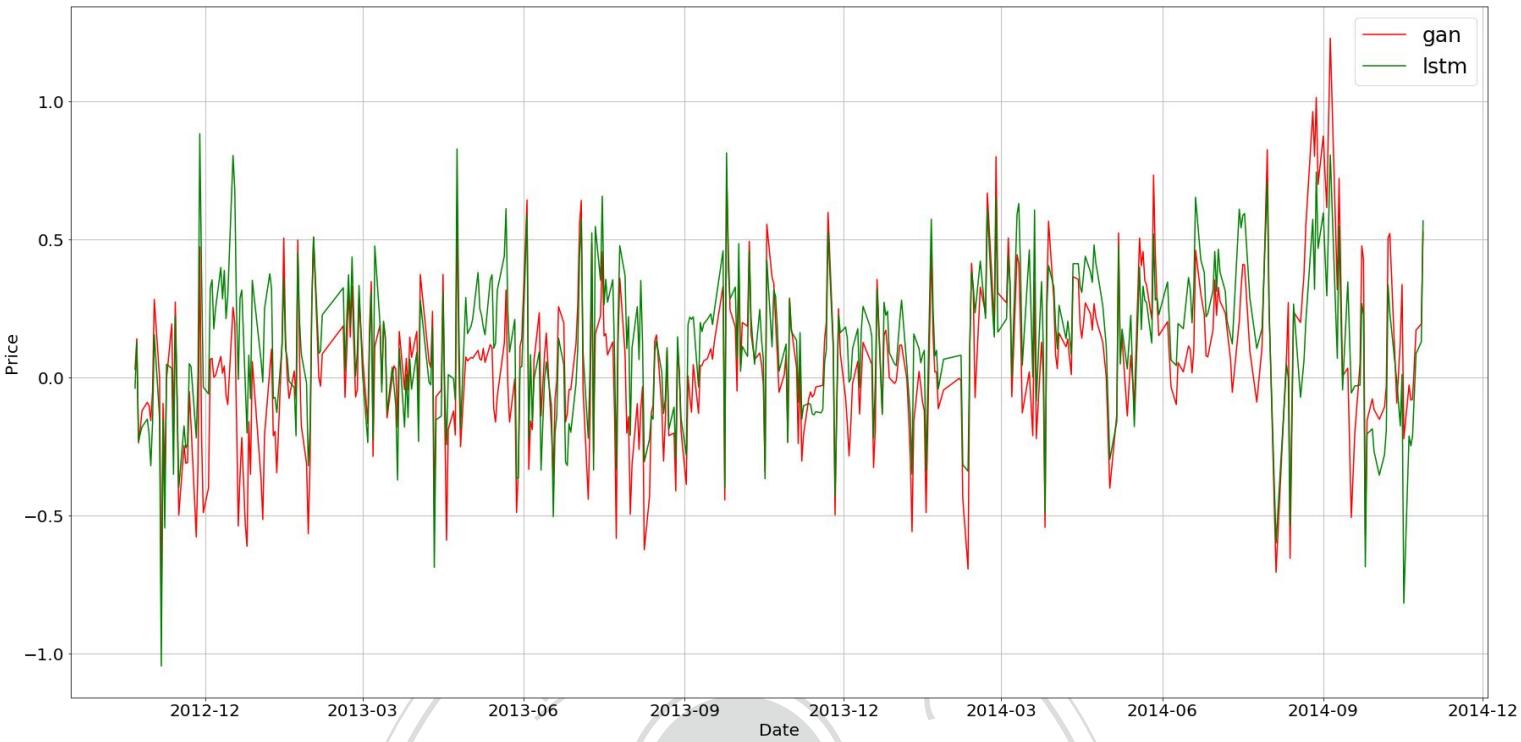

圖 5-14 長短期記憶模型與生成對抗網路預測 006208 之收盤價與 006208 真實資料差距之第一區間

下圖 5-15 長短期記憶、生成對抗網路與 006208 真實收價格走勢之第二區間 2014 年 10 月 3 日至 2017 年 3 月 6 日的比較。在此區間觀察到長短期記憶模型所預測的收盤價格有需多處較生成對抗模型混亂，如：2014 年 12 月左右、2016 年 3 月至 2016 年 5 月。生成對抗網路下跌反彈後的價格高估的現象出現次數不少，如 2015 年 6 月底至 8 月底因受到中國股市大跌、人民幣貶值、歐美日股市也相繼大跌、2016 年 1 月總統選舉。在小波動的時，生成對抗模型的預測值都比起長短期記憶模型更接近真實資料。圖 5-16 是兩模型預測收盤價與真實收盤價的差距。

DOI:10.6814/NCCU201900589

圖 5-15 長短期記憶、生成對抗網路與 006208 真實收價格走勢之第二區間

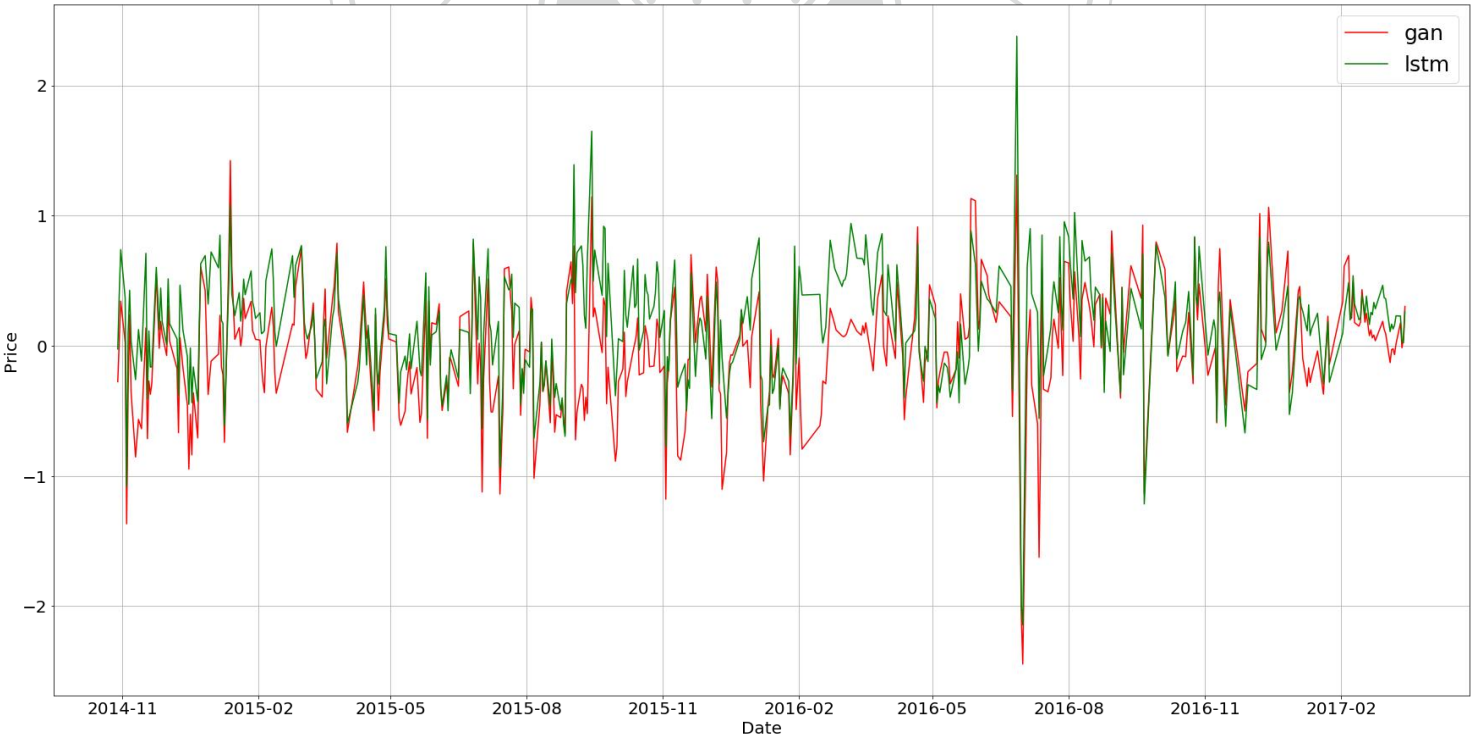

圖 5-16 長短期記憶模型與生成對抗網路預測 006208 之收盤價與 006208 真實資料差距之第二區間

DOI:10.6814/NCCU201900589

下圖 5-17 長短期記憶、生成對抗網路與 006208 真實收價格走勢之第三區間 2017 年 3 月 6 日至 2019 年 4 月 3 日的比較。2017 年 8 月份大跌反彈後出現長短期記憶模型收盤價格高於生成對抗網路的收盤價格。2018 年 2 月美國股市大跌後反彈生成對抗網路沒有出現高估的情況，但後面幾天卻有高估的情形。在此區間可以觀察到比較凌亂的地方，如 2018 年底至 2019 年。圖 5-18 是兩模型預測收盤價與真實收盤價的差距。

圖 5-17 長短期記憶、生成對抗網路與 006208 真實收價格走勢之第三區間

DOI:10.6814/NCCU201900589

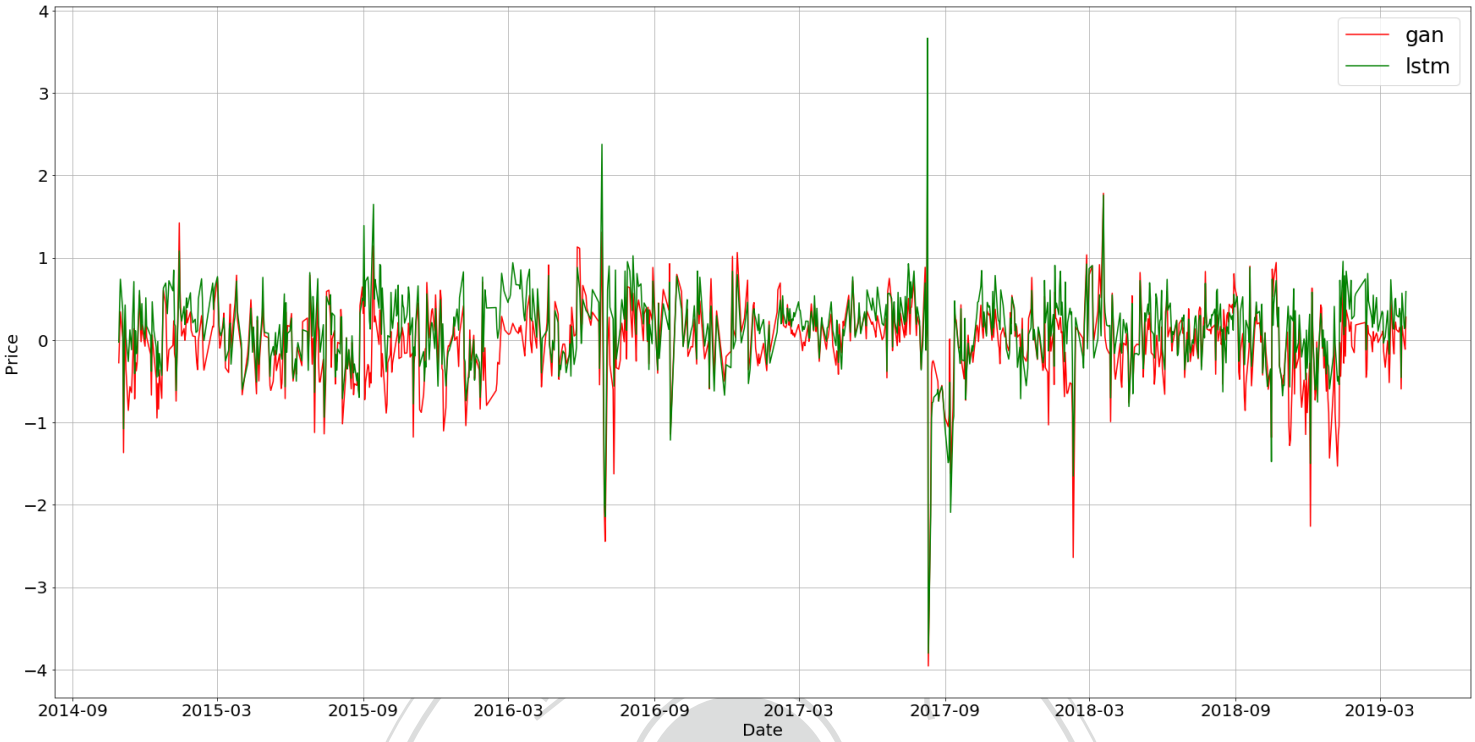

圖 5-18 長短期記憶模型與生成對抗網路預測 006208 之收盤價與 006208 真實資料差距之第三區間

經由將富邦台灣采吉 50 基金(006208)的結果分析可以得知，在行情沒有巨幅波動的時候，生成對抗網路則仍較長短期記憶模型更接近真實資料。生成對抗網路易受到大事件所造成的大波動影響，在大跌反彈後的價格都有高估收盤價的情況，此現象與在測試資料富邦台灣摩根(0057)所觀察到的現象一致。然而在此測試資料中，長短期記憶模型遇到大波動時表現則較卻比富邦台灣摩根(0057)所觀察到的更混亂，加上在小波動區段生成對抗網路更貼近真實資料，因而造成生成對抗網路預測富邦台灣采吉 50 基金(006208)的均方誤差結果比長短期記憶模型小 0.00642。

我們將兩模型經測試資料富邦台灣采吉 50 基金(006208)預測出的收盤價格，分別計算每五日的成交價格漲跌方向的準確率，長短期記憶模型預測富邦

DOI:10.6814/NCCU201900589

台灣采吉 50 基金(006208)五日收盤價漲跌準確率為 79.01%，生成對抗網路則是 79.86%，生成對抗網路有較好的能力預估五日漲跌方向。

第四節 小結

透過富邦台灣摩根(0057)與富邦台灣采吉 50 基金(006208)兩測試資料可以得出長短期記憶模型與生成對抗網路有相同的觀察結果。然而，相同的觀察結果卻出現長短期記憶模型預測富邦台灣摩根(0057)的收盤價能力較好的現象，而加上卷積神經網路做為鑑別模型的長短期記憶模型卻有較差的預測能力。

於是本文探討兩測試資料的性質。富邦台灣采吉 50 基金(006208)追蹤台灣 50 指數，由台灣上市公司前 50 大市值的公司組成，富邦台灣摩根(0057)則是追蹤 MSCI®臺灣指數，成分股散佈局至 18 種產業，不僅上市公司，包含上櫃公司更能反映台灣整體的股票市場。依據經濟學供需原理，正常財的價格高低與數量多寡是密不可分的關係，表 6-1 比較富邦台灣摩根(0057)與富邦台灣采吉 50 基金(006208)每日成交量和價格變動率的平均與標準差發現，台灣摩根(0057)每日平均的成交量與成交量標準差皆比富邦台灣采吉 50 基金(006208)大，在平均每日價格變動率與其標準差方面，台灣摩根(0057)也比富邦台灣采吉 50 基金(006208)大；富邦台灣摩根(0057)的每日的成交量散佈非常廣。也代表著，富邦台灣摩根(0057)價格變動率與成交量的變化程度比富邦台灣采吉 50 基金(006208)更大，其影響價格背後隱含的因素相當多，較難預測。

DOI:10.6814/NCCU201900589

表 5-5 富邦台灣摩根(0057)與富邦台灣采吉 50 基金(006208)之平均每日成交量、每日成交量標準差、每日價格變動率、每日價格變動率標準差

	富邦台灣摩根(0057)	富邦台灣采吉 50 基金(006208)
每日平均成交量	324,970	61,910
每日成交量標準差	5,405,530	183,070
每日價格變動率	0.03%	0.03%
每日價格變動率標準差	1.15%	1.02%

因此，這也代表著長短期記憶模型具有捕捉較為複雜的序列資料特徵的能力，也是造成相同的觀察結果卻出現長短期記憶模型預測富邦台灣摩根(0057)的收盤價能力較好的現象。然而，使用卷積神經網路之鑑別網路監督長短期記憶模型學習的生成對抗網路在波動較大的資料中表現較差，因此鑑別網路的選擇也是相當重要的問題。

國立政治大學 National Chengchi University

DOI:10.6814/NCCU201900589

第六章 結論與建議

第一節 結論

1. 長短期記憶模型具有較強的能力預測波動性較大的商品。本文的觀察與 Thomas Fischer 與 Christopher Krauss(2017)具有相同的結果，該文章利用長短期記憶模型預測與其他無記憶性的神經網路使用 1992 年至 2015 年間 S&P500 成分股的資料建構交易策略，發現使用長短期記憶模型所建構的交易策略之報酬率高過其他無記憶性類神經網路，也代表著長短期記憶模型可以萃取出高波動的股票特徵、強烈短期反轉訊號的能力。

2. 生成對抗網路有較高的預測台灣非金融 ETF 價格走勢能力。無論是邦台灣摩根(0057)還是富邦台灣采吉 50 基金(006208)，使用鑑別網路監督長短期記憶模型訓練的模型都有較高的價格走勢預測能力，且當行情真實為牛市時，藉由鑑別網路輔助長短期記憶模型學習的模型比一般的長短期記憶模型有更高的機率抓住此波上漲的趨勢

3. 本文不針對任何一個 ETF 商品建模，而是透過樣本資料訓練出一個可以廣泛運用在台灣非金融類 ETF 預測的深度學習模型。即使訓練期間，模型沒有看過的測試資料，仍可以預測該測試資料的收盤價格，並且五日價格走勢準確率高達 79.92%。然而，與 Xingyu Zhou et al.(2018)的文章的結果不一樣，在台灣非金融類 ETF 的價格與漲跌的預測中，當我們建立可以廣泛應用的類神經網路模型，生成對抗網路並非都比長短期記憶模型的預測能力來得更好，因此生成對抗網路中的鑑別模型之選擇也是重要的議題。

DOI:10.6814/NCCU201900589

第二節 後續研究與建議

台灣股市很容易受到美國的股市影響，在第五章第三節比較了兩測試資料富邦台灣摩根(0057)與富邦台灣采吉 50 基金(006208)，其中大有許多行情大跌的原因都來自美國股市的影響，可以加入標準普爾 500 或是 NASDAQ 指數進行預測。這也表示本文的最後結果會受到特徵的選擇與數量。

另外，雖然有生成對抗網路有捕捉到下跌的趨勢，但下跌後反彈的價格都有高估的現象。根據 Thomas Fischer and Christopher Krauss(2017)與本文的結果，可以將本文的生成對抗網路之鑑別網路換成長短期記憶模型，以此監督生成網路，捕捉更多高波動的特徵訊號。

國立政治大學 National Chengchi University

DOI:10.6814/NCCU201900589

參考文獻

中文文獻

[1] 胡依淳(2018),「深度卷積神經網路中卷積層之分析及比較」,國立暨南國際大學電機工程學系碩士論文。

[2] 陳全溢(2018),「結合類神經網路預測與投資策略於台灣 50 指數股票型基金之操作」,國立中興大學資訊管理學系碩士論文。

[3] 陳俊諺(2018),「運用類神經網路與田口法預測台灣 ETF 指數」,中原大學,資訊管理系碩士論文。

[4] 黃焜烽(2018),「利用深度類神經網路模型預測台灣股價指數走勢」,國立臺北大學經濟系碩士論文。

[5] 楊國良(2017),「運用倒傳遞類神經網路預測台灣 50 指數 ETF 股價走勢」,國立金門大學理工學院工程科技碩士在職專班資訊系統組碩士論文。

國立政治大學
National Chengchi University

DOI:10.6814/NCCU201900589

英文文獻

[1] Adam Santoro, Ryan Faulkner, David Raposo, Jack Rae, Mike Chrzanowski, Théophane Weber, Daan Wierstra, Oriol Vinyals, Razvan Pascanu, Timothy Lillicrap. (2018). Relational recurrent neural networks. NeurIPS 2018.

[2] Alex Graves, Greg Wayne, Ivo Danihelka. (2014). Neural Turing Machines. arXiv:1410.5401v2

[3] Amin Hedayati Moghaddama, Moein Hedayati Moghaddamb, Morteza Esfandyaric. (2016). Stock market index prediction using artificial neural network. Journal of Economics, Finance and Administrative Science Volume 21, Issue 41, 89－93

[4] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser, Illia Polosukhin. (2017). Attention Is All You Need. arXiv:1706.03762v5.

[5] Dzmitry Bahdanau, Kyunghyun Cho, Yoshua Bengio. (2016). Neural Machine Translation by Jointly Learning to Align and Translate. arXiv:1409.0473v7.

[6] Ian J. Goodfellow, Jean Pouget-Abadiey, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozairz, Aaron Courville, Yoshua Bengio. (2014). Generative Adversarial Nets. NIPS 2014.

[7] Ilya Sutskever, Oriol Vinyals, Quoc V. Le. (2014). Sequence to Sequence Learning with Neural Networks. NIPS 2014.

[8] John Gamboa (2017). Deep Learning for Time-Series Analysis. arXiv:1701.01887v1.

[9] KangZhang, GuoqiangZhong, JunyuDong, ShengkeWang, YongWang.

國立政治大學 National Chengchi University

DOI:10.6814/NCCU201900589

(2019). Stock Market Prediction Based on Generative Adversarial Network. Procedia Computer Science, Volume 147, 2019, Pages 400-406.

[10]Kyunghyun Cho, Bart van Merrienboer, Caglar Gulcehre, Dzmitry Bahdanau, Fethi Bougares, Holger Schwenk, Yoshua Bengio. (2014). Learning Phrase Representations using RNN Encoder-Decoder for Statistical Machine Translation. arXiv:1406.1078v3.

[11]Masaya Abe1 and Hideki Nakayama2. (2018). Deep Learning for Forecasting Stock Returns in the Cross-Section. arXiv:1801.01777v4.

[12]Mikołaj Bi´nkowski, Gautier Marti, Philippe Donnat. (2018). Autoregressive Convolutional Neural Networks for Asynchronous Time Series. arXiv:1703.04122v4.

[13]Minh-Thang Luong Hieu Pham Christopher D. Manning. (2015). Effective Approaches to Attention-based Neural Machine Translation. arXiv:1508.04025v5.

[14]S.E. Yi, A. Viscardi, T. Hollis. (2018). A Comparison of LSTMs and Attention Mechanisms for Forecasting Financial Time Series. arXiv:1812.07699v1

[15]Sepp Hochreiter and Jurgen Schmidhuber. (1997). LONG SHORT-TERM MEMORY. Neural Computation 9(8):1735-1780.

[16]SIMA SIAMI NAMIN1, AKBAR SIAMI NAMIN2. (2018). FORECASTING ECONOMIC AND FINANCIAL TIME SERIES: ARIMA VS. LSTM. arXiv:1803.06386v1.

[17]T. Kimoto ; K. Asakawa ; M. Yoda ; M. Takeoka. (1990). Stock market prediction system with modular neural networks.IEEE

DOI:10.6814/NCCU201900589

10.1109/IJCNN.1990.137535.

[18]Takashi MATSUBARA, Member, Ryo AKITA, Nonmember, and Kuniaki UEHARA. (2018). Stock Price Prediction by Deep Neural Generative Model of News Articles. IEICE Transactions on Information and Systems, 2018 Volume E101.D Issue 4, Pages 901-908.

[19]Thomas R. Cook and Aaron Smalter Hall. (2017). Macroeconomic Indicator Forecasting with Deep Neural Networks. Federal Reserve Bank of Kansas City, Research Working Paper 17-11, September 2017

[20]Trieu H. Trinhl Andrew M. Dai Minh-Thang Luong Quoc V. Le. (2018). Learning Longer-term Dependencies in RNNs with Auxiliary Losses. ICLR 2018.

[21]V.V.Kondratenko, Yu.A Kuperin.(2003).Using Recurrent Neural Networks To Forecasting of Forex. arXiv:cond-mat/0304469v1.

[22]V´ıctor Campos, Brendan Jouz, Xavier Gir´o-i-Nietox, Jordi Torresy, Shih-Fu Chang. (2018) "SKIP RNN: LEARNING TO SKIP STATE UPDATES IN RECURRENT NEURAL NETWORKS. ICLR 2018

[23]Wei Bao, Jun Yue, Yulei Rao. (2017). A deep learning framework for financial time series using stacked autoencoders and long-short term memory. PLoS ONE 12(7): e0180944. https://doi.org/10.1371/journal.pone.0180944

[24]Xiao Dingy, Yue Zhangz, Ting Liuy, Junwen Duan. (2015). Deep Learning for Event-Driven Stock Prediction. IJCAI 2015.

[25]Xingyu Zhou, Zhisong Pan, Guyu Hu, Siqi Tang, Cheng Zhao. (2018). Stock Market Prediction on High-Frequency Data Using Generative Adversarial Nets. Mathematical Problems in Engineering, Volume

DOI:10.6814/NCCU201900589

2018, Article ID 4907423, 11 pages.

[26]Xin-Yao Qian. (2017). Financial Series Prediction: Comparison Between Precision of Time Series Models and Machine Learning Methods. arXiv:1706.00948.

國立政治大學
National Chengchi University

DOI:10.6814/NCCU201900589