

元智大學  
資訊管理學系碩士班  
碩士論文

基於多尺度卷積與雙向門控循環單元研究  
-以比特幣價格預測為例

Multi-scale Residual Convolutional Neural Network  
with Bi-directional Gate Recurrent Unit  
- A Case Study on Bitcoin Price Prediction

研 究 生：張朝瑞  
指導教授：袁鳳清 博士

中華民國一一三年六月

基於多尺度卷積與雙向門控循環單元研究  
-以比特幣價格預測為例

Multi-scale Residual Convolutional Neural Network with  
Bi-directional Gate Recurrent Unit  
-A Case Study on Bitcoin Price Prediction

研究 生：張朝瑞                  Student : Chao-Rui Chang  
指 導 教 授：袁鳳清 博 士      Advisor : Dr. Fong-Ching Yuan



Submitted to Department of Information Management  
College of Informatics  
Yuan Ze University  
in Partial Fulfillment of the Requirements  
for the Degree of  
Master of Science  
in  
Information Management

June 2024  
Chungli, Taiwan, Republic of China  
中華民國一一三年六月

# 論文口試委員審定書



# 基於多尺度卷積與雙向門控循環單元研究 -以比特幣價格預測為例

學生：張朝瑞

指導教授：袁鳳清 博士

元智大學資訊管理學系碩士班

## 摘要

在過去 20 年裡，金融商品不斷地在創新，其中的加密貨幣已逐漸成為投資者偏好的選擇之一。加密貨幣因具備「全球性」、「全年無休、24 小時交易」、「可作為美股、黃金的風險對沖工具」等特性，使得其價格波動範圍較為劇烈，這種高風險、高報酬的特點，使得投資者需要隨時關切價格波動資訊，以降低風險並增加獲利機會。因此，本研究擬建立一套完整的加密貨幣價格預測系統，期能深入分析加密貨幣價格，發展出更為精確的價格預測系統。比特幣（Bitcoin, BTC）為發展最早且最具知名度的加密貨幣，具有「大量交易額」與「加密貨幣代表性」等性質，又有數位黃金之稱，深具分析的價值。因此，本研究選擇以比特幣價格代表加密貨幣價格，為本研究被預測之因變數。

過去許多研究者利用“長短期記憶網路”（Long Short-Term Memory, LSTM）、“門控循環單元網路”（Gate Recurrent Unit, GRU）等模式處理預測問題，但由於“多尺度殘差卷積神經網路”（Multi-scale Residual Convolutional Neural Network, MRC）能夠融合不同時間尺度的特徵，而“雙向門控循環單元網路”（Bi-directional Gate Recurrent Unit, Bi-GRU）保有循環神經網路與高效率計算等優勢。因此，本研究擬採用混合 MRC 和 Bi-GRU 的方法，預期能有效地獲得高度的特徵，更有效率地學習多元時間序列趨勢和交互作用，應能協助投資人更準確地預測加密貨幣的價格。

除此之外，本研究同時設計了三種不同時間尺度大小的實驗資料，分別測試於 MRC-LSTM、MRC-GRU、MRC-BiLSTM、MRC-BiGRU 等模型訓練後的預測結果。實證研究顯示，在時間尺度較小的實驗中，GRU 系列獲取較 LSTM 系列為佳的預測結果，且以雙向模型 MRC\_BiGRU 呈現出最佳的預測效果。而在時間尺度較大的實驗中，LSTM 系列獲得較 GRU 系列為佳的預測結果，且以雙向模型 MRC\_BiLSTM 呈現出最佳的預測效果。最後，本研究依據研究結果提出結論與建議，以俾後續相關領域研究者之參考。



關鍵字：比特幣價格預測、多尺度殘差卷積神經網路、雙向門控循環單元網路模型

# **Multi-scale Residual Convolutional Neural Network with Bi-directional Gate Recurrent Unit - A Case Study on Bitcoin Price Prediction**

**Student : Chao-Rui Chang**

**Advisor : Dr. Fong-Ching Yuan**

**Department of Information Management  
College of Informatics  
Yuan Ze University**

## **ABSTRACT**

In the past 20 years, financial products have continuously innovated, among which virtual currencies have gradually become one of the preferred choices for investors. Cryptocurrency possess characteristics such as "globality," "24/7 trading," and "use as risk hedging tools for US stocks and gold," which lead to more intense price fluctuations. This high-risk, high-return nature necessitates investors to constantly monitor price volatility information to reduce risks and increase profit opportunities. Therefore, this study aims to establish a comprehensive cryptocurrency price prediction system to deeply analyze virtual currency prices and develop a more accurate price prediction system. Bitcoin (BTC), the earliest and most well-known virtual currency, with properties such as "high trading volume" and "representativeness of cryptocurrencies," also dubbed as digital gold, holds significant analytical value. Hence, this study chooses Bitcoin prices to represent virtual currency prices as the dependent variable for prediction.

Many researchers in the past have utilized models like Long Short-Term Memory (LSTM), Gate Recurrent Unit (GRU), etc., for prediction tasks. However, due to the ability of Multi-scale Residual Convolutional Neural Network (MRC) to integrate

features from different time scales and the advantages of Bi-directional Gate Recurrent Unit (Bi-GRU) in retaining the benefits of recurrent neural networks and efficient computation, this study proposes to adopt a hybrid approach combining MRC and Bi-GRU, expecting to effectively acquire rich features and learn multiple time series trends and interactions more efficiently, thus assisting investors in more accurately predicting virtual currency prices.

Furthermore, this study designs three sets of experimental data with different time scale sizes to test the predictive results after training on models such as MRC-LSTM, MRC-GRU, MRC-BiLSTM, MRC-BiGRU, etc. Empirical research demonstrates that in experiments with smaller time scales, the GRU series achieves better predictive results than the LSTM series, with the bidirectional model MRC-BiGRU presenting the best predictive performance. In experiments with larger time scales, the LSTM series outperforms the GRU series, with the bidirectional model MRC-BiLSTM demonstrating the best predictive performance. Finally, based on the research findings, this study concludes and offers suggestions for subsequent researchers in related fields.

Keyword : Bitcoin price Forecasting 、 Multi-scale Residual Convolutional Neural Network (MRC) 、 Bi-directional Gate Recurrent Unit (Bi-GRU)

## 誌謝

時光荏苒，轉瞬間已在元智待了六年。回想起當初剛進入元智時的青澀、懵懂，隨著時間與學習的歷練，即將邁入人生下一個階段，不經感慨萬千。在論文完成之際，誠摯地感謝所有支持、鼓勵和幫助過我的人。首先，我要感謝我的指導教授－袁鳳清老師，老師的專業知識、耐心指導和無私支持，是我完成這項研究最重要的動力。老師的悉心指導讓我受益良多，對我來說這將是一段寶貴的學習經歷。

其次，我要衷心感謝我的家人。感謝我的父母，他們的無私奉獻和無限鼓勵，讓我也能夠專注於我的學業。他們的支持是我前進的動力，我永遠感激他們在我生命中的存在。此外，我還要感謝我的同學和朋友。在我學習和研究的過程中，他們的陪伴和支持給予了我很多幫助，讓我感到溫暖和鼓舞。

最後，我要特別感謝元智大學 (Yuan Ze University)。在我就讀期間，學校不僅提供了優秀的教育環境，更為我提供了豐富的工讀機會和獎學金資助。這些機會不僅幫助我減輕了經濟壓力，還讓我有機會實踐所學，獲得寶貴的工作經驗。我要感謝學校管理部門、老師和工作人員，由於有他們的支持和關懷，讓我在學業和生活上都能得到充分的成長和發展。元智大學將永遠在我心中佔有一席之地，我會永遠珍惜這段美好的校園時光。

我深知這份成果不僅是我個人的努力，更是許多人共同努力的結果。在此，我要向所有支持過我的人表示最誠摯的謝意。感謝你們！

元智大學資訊管理學系碩士班  
張朝瑞 謹誌  
中華民國一一三年六月

# 目錄

書名頁 .....	i
論文口試委員審定書 .....	ii
摘要 .....	iii
ABSTRACT .....	v
誌謝 .....	vii
目錄 .....	viii
表目錄 .....	ix
圖目錄 .....	x
第一章 緒論 .....	1
第一節 研究背景與動機 .....	1
第二節 研究目的 .....	6
第三節 研究架構與流程 .....	8
第二章 文獻探討 .....	9
第一節 比特幣價格預測 .....	9
第二節 單一預測模型預測比特幣價格 .....	12
第三節 混合預測模型預測比特幣價格 .....	13
第四節 訓練模型之特徵值 .....	15
第三章 研究方法 .....	16
第一節 卷積神經網路 (CNN) .....	16
第二節 多尺度殘差卷積神經網路 (MRC) .....	19
第三節 長短期記憶網路 (LSTM) .....	23
第四節 雙向門控循環單元網路 (Bi-GRU) .....	26
第五節 多尺度卷積雙向門控循環單元網路 (MRC-BiGRU) .....	29
第四章 實驗結果與分析 .....	31
第一節 訓練資料收集 .....	31
第二節 訓練資料處理 .....	34
第三節 建構比特幣價格預測模型 .....	37
第四節 預測模型評估指標 .....	39
第五節 預測模型評估 .....	41
第五章 結論與建議 .....	48
第一節 研究結論 .....	48
第二節 研究建議 .....	50
參考文獻 .....	51

## 表目錄

表 2.1 比特幣價格預測文獻整理.....	9
表 4.1 本研究重要變數與資料取得來源.....	33
表 4.2 各時間尺度資料集.....	36
表 4.3 實驗環境.....	38
表 4.4 預測模型的誤差 .....	41
表 4.5 預測模型的誤差 .....	43
表 4.6 預測模型的誤差 .....	45



## 圖目錄

圖 1.1 加密貨幣幣種流通量 .....	2
圖 1.2 加密貨幣持有者數量 .....	2
圖 1.3 Web 3 錢包 SDK 安裝數量 .....	3
圖 1.4 Ethereum SDK 安裝數量 .....	4
圖 1.5 各大加密貨幣市值 .....	5
圖 1.6 研究流程圖 .....	8
圖 3.1 卷積層概念圖 .....	17
圖 3.2 池化層概念圖 .....	17
圖 3.3 ResNet 概念圖 .....	19
圖 3.4 DenseNet 概念圖 .....	20
圖 3.5 MRC 模型架構圖 .....	21
圖 3.6 常見的激活函數 .....	22
圖 3.7 LSTM 記憶元示意圖 .....	24
圖 3.8 GRU 記憶元示意圖 .....	26
圖 3.9 Bi-GRU 概念圖 .....	28
圖 3.10 MRC-BiGRU 模型 .....	29
圖 3.11 研究進行步驟流程圖 .....	30
圖 4.1 標準化前後示意圖 .....	34
圖 4.2 滑動窗口切割示意圖 .....	35
圖 4.3 MRC-BiGRU 模型視覺化 .....	37
圖 4.4 MRC_BiGRU 與 MRC_BiLSTM 預測與實際價格結果比較 .....	41
圖 4.5 MRC_BiGRU 與 MRC_GRU 預測與實際價格結果比較 .....	42
圖 4.6 MRC_BiGRU 與 MRC_LSTM 預測與實際價格結果比較 .....	42
圖 4.7 MRC_BiGRU 與 MRC_BiLSTM 預測與實際價格結果比較 .....	43
圖 4.8 MRC_BiGRU 與 MRC_GRU 預測與實際價格結果比較 .....	44
圖 4.9 MRC_BiGRU 與 MRC_LSTM 預測與實際價格結果比較 .....	44
圖 4.10 MRC_BiLSTM 與 MRC_BiGRU 預測與實際價格結果比較 .....	45
圖 4.11 MRC_BiLSTM 與 MRC_GRU 預測與實際價格結果比較 .....	46
圖 4.12 MRC_BiLSTM 與 MRC_LSTM 預測與實際價格結果比較 .....	46

# 第一章 緒論

## 第一節 研究背景與動機

加密貨幣（Cryptocurrency）是 Nakamoto 於 2008 年首次提出，標誌著數位金融領域的重大革新。這項概念融合了區塊鏈技術、密碼學，以及共識機制，成功打造出一套去中心化且無需中介的交易系統，徹底改變了傳統金融體系的運作模式，其優勢在於其提供的高度安全性和效能。透過區塊鏈的不可變性和去中心化的特性，加密貨幣確保了交易的安全性。同時，其高效的交易速度和無國界的特點，使得加密貨幣在全球範圍內廣受歡迎。

自 2009 年起，加密貨幣經歷了持續的發展和創新，擴展了其應用範疇。這種數位資產的崛起，象徵著金融科技的演進，並在全球引起了廣泛的興趣。加密貨幣的興起，不僅是金融領域的一場變革，更是對傳統金融體系的挑戰，引領著未來金融發展的方向。

與其他傳統金融商品相比，加密貨幣的價格相對劇烈變動，無地理、時間限制的自由轉移、幾乎甚少的交易成本，形成了加密貨幣市場高度敏感性的主要原因。由於加密貨幣市場相對較小，即使一點點的需求或供應變化都可能對價格產生巨大影響。其次，市場信心和情緒對加密貨幣價格的波動也有很大的影響。由於缺乏實體資產支撐，投資者情緒波動性較大，市場可能受到情報、宣傳或市場環境變化等因素的影響而產生極端的波動。然而，正是因為它的劇烈價格變動性與產品創新的特性，吸引了大量的風險投資家與金融機構的目光。

根據知名獨立加密貨幣資料整合商 CoinGecko 統計，截至 2023 年 11 月止，約有 10,748 種不同的加密貨幣於市面上流通，如圖 1.1 所示，儘管並非所有加密貨幣仍然保持其核心價值，但扣除已經失去價值與失去活躍的加密貨幣後，仍存

在 8,848 種加密貨幣於市面上流通。除此之外，共有 982 個加密貨幣交易所、634 個去中心化交易平台可提供加密貨幣的持有者進行交易；且根據 Crypto.com 交易所提出的研究報告，至 2023 年全球加密貨幣持有者數量成長 34%，從 1 月的 4.32 億增加到 12 月的 5.8 億，如圖 1.2 所示；值得一提的是，薩爾瓦多、中非共和國等國家甚至於 2021 年 9 月後陸續宣布加密貨幣為國家的法定貨幣，部分國家也排定議程商討是否跟上腳步。

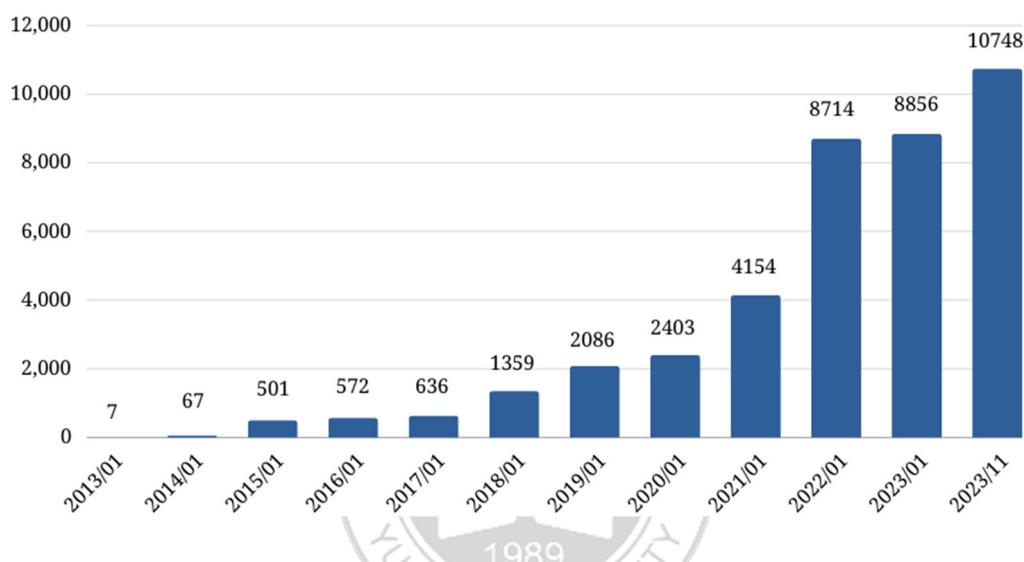


圖 1.1 加密貨幣幣種流通量

資料來源：CoinGecko

Total Number of Global Crypto Owners (In Millions)

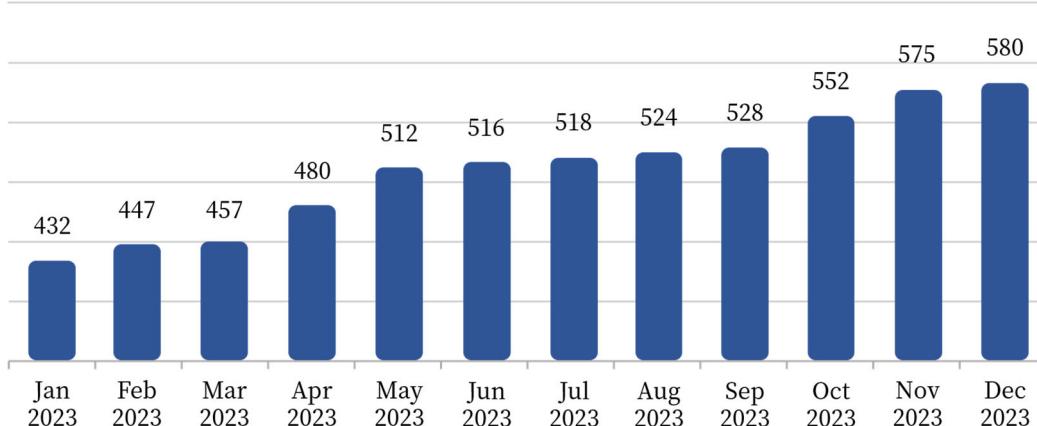


圖 1.2 加密貨幣持有者數量

資料來源：Crypto.com 交易所

近年來，因區塊鏈技術已逐漸成熟，而延伸出更多不同的 DeFi（去中心化金融）之應用、商品、組織或服務，如 GameFi（遊戲化金融）、NFT（非同質化代幣）、Ordinals（銘文）、DAO（去中心化自治組織）。與 CeFi（傳統中心化金融）仰賴如銀行、證卷行等中心化機構不同，DeFi 可以透過智能合約（Smart Contract）在開放、透明、任何人都能參與的去中心化環境下實現各式交易、轉帳、借貸等金融服務。與傳統上強調使用者生成內容、參與文化和互操作性的 Web 2.0 相比，上述應用正逐漸形成更強調去中心化、用戶共決的 Web 3.0 趨勢。

圖 1.3 與圖 1.4 為區塊鏈開發平台 Alchemy 的官方統計。自 2019 年以來，Ethereum SDK 與 Web 3 錢包 SDK 安裝量成長數據，從圖片中可觀測出，Web 3 錢包 SDK 安裝數量已從 2019 年第一季度的 4,800 個錢包，上升至 2023 年第二季度的 1,110 萬個錢包；Ethereum SDK 安裝數量已從 2019 年第一季度的 290 萬次下載，上升至 2023 年第二季度的 2,680 萬次下載，儘管加密貨幣市場在面臨熊市（Bear Market）的時期，也依舊能保持穩定成長的趨勢。



圖 1.3 Web 3 錢包 SDK 安裝數量

資料來源：Alchemy

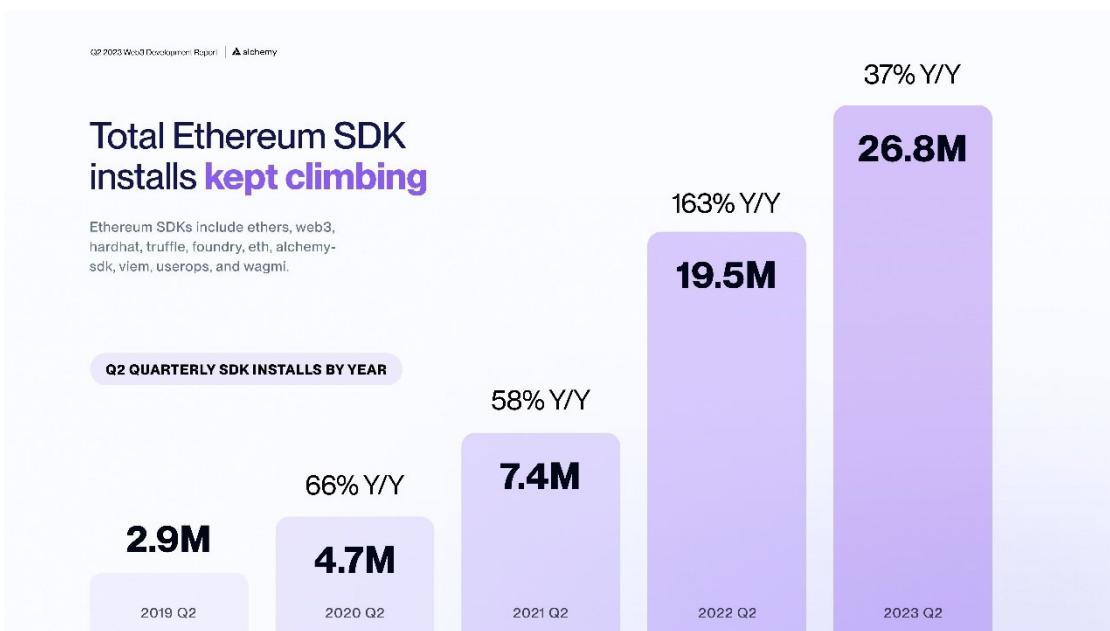


圖 1.4 Ethereum SDK 安裝數量

資料來源：Alchemy

在過去 20 年裡，金融商品一直在不斷創新，加密貨幣逐漸成為投資者偏好的選擇之一。相較於其他金融商品，加密貨幣的價格波動範圍顯著巨大，原因包括「加密貨幣的價值、實際應用仍有爭議性」，並同時具備「全球性」、「全年無休、24 小時交易」、「可作為美股、黃金的風險對沖工具」等特性，這樣劇烈的波動性質，使加密貨幣對投資者而言帶有高風險、高報酬的特點。投資者應密切關注加密貨幣價格波動資訊，以降低風險並增加獲利機會。因此，建立一套完整的加密貨幣價格預測系統對投資者至關重要，故本研究旨在深入分析加密貨幣價格，以期能發展更為精密的價格預測系統。

美國證管會 (SEC) 已於美國時間 2024 年 1 月 10 日核准首批比特幣 (Bitcoin, BTC) 的 ETF 上市，而其他的基金產業巨擘，如貝萊德、景順、富達、方舟投資 (Ark Investments)、21Shares、VanEck 和 Valkyrie 等公司申請的比特幣 ETF 也陸續獲准通過。這對於規模約 1.7 兆美元的數位資產產業而言，正式走向 ETF 化的新階段無疑具有里程碑意義。

加密貨幣的交易市場存在著許多不同種類的貨幣，其中以比特幣(Bitcoin, BTC)發展最早並最為知名，根據加密貨幣幣價追蹤網站 CoinMarketCap 統計，截至 2024 年 1 月止，比特幣佔有加密貨幣市場中最大市值，並且是擁有最大交易額的加密貨幣，因此有數位黃金之稱，如圖 1.5 所示。因比特幣具有「大量的交易額」，並有「加密貨幣的代表性」性質，深具分析的價值。因此，本研究選擇以比特幣價格代表加密貨幣價格，為本研究之被預測因變數。

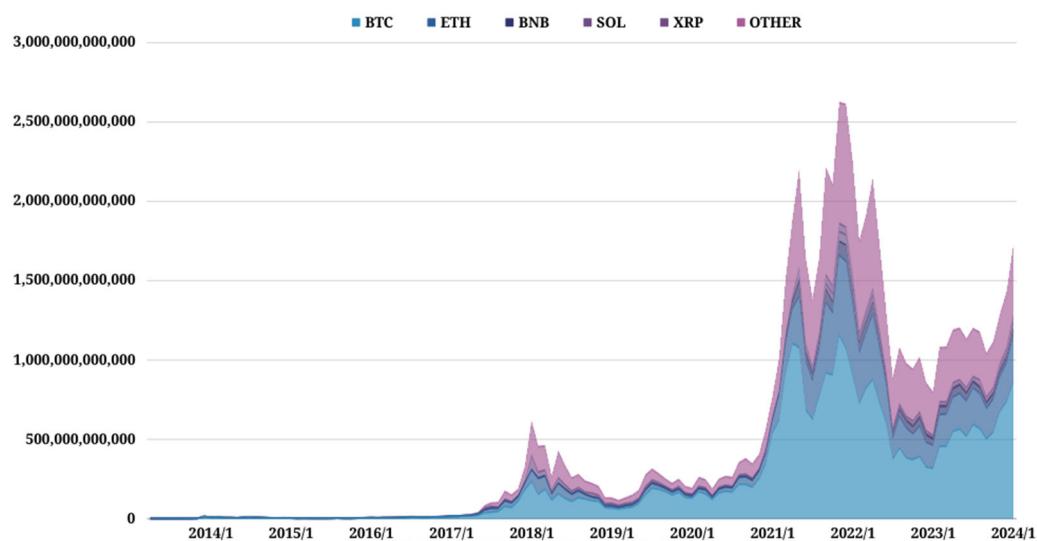


圖 1.5 各大加密貨幣市值

資料來源：CoinMarketCap

而加密貨幣的價格變動巨大，技術和法規的變動也是造成加密貨幣價格波動的原因之一。新的技術發展或法規的制定都可能引起市場的不確定性，從而影響加密貨幣的價格。此外，全球經濟和金融動盪也會對加密貨幣市場產生重大影響。全球經濟事件、政治動盪或金融危機都可能引發投資者對風險資產的恐慌性變動，進而影響加密貨幣價格。

總體而言，加密貨幣價格變動巨大的原因是一個複雜而多元的系統，由供需、市場情緒、技術變革和全球經濟環境等多方面因素共同塑造。投資者應謹慎應對這種高度不確定性的市場環境，並深入了解相關風險因素，以更有效地管理投資組合。

## 第二節 研究目的

在研究方法上，傳統股價和期貨市場之預測，多以統計時間序列方法為主，如多元迴歸分析（Multiple regression analysis）是一種用於探討多個自變數（解釋變數）與一個應變數（響應變數）之間影響關係的統計方法，它可以擴展到多變量時間序列模型，其中多個變數同時被考慮；卡爾曼濾波器（Kalman Filter）用於實時估計系統的狀態變數，廣泛應用於控制系統和時間序列分析；整合移動平均自我迴歸模型(Autoregressive Integrated Moving Average, ARIMA)由自迴歸(AR)、整合(I)和移動平均(MA)組成，整合(I)表示需要對數據進行差分以使其變得平穩；季節性自我迴歸整合移動平均模型（Seasonal Autoregressive Integrated Moving Average, SARIMA）、自我迴歸條件異方差模型（Autoregressive Moving Average - Generalized Autoregressive Conditional Heteroskedasticity model, ARMA-GARCH）皆是 ARIMA 模型的擴展，它們分別考慮時間序列中的季節性影響與時間序列中的自相關結構和變異數的波動；向量自我迴歸模型（Vector Autoregression, VAR）用於描述和分析多變量時間序列之間的相互關係；指數平滑法（Exponential Smoothing）包括單指數平滑、雙指數平滑和三指數平滑，用於平滑時間序列並進行預測。

近年來隨著人工神經網路（Artificial Neural Network, ANN）、機器學習（Machine Learning, ML）與深度學習（Deep Learning, DL）的發展，許多研究者也開始利用循環神經網路(Recurrent neural network, RNN)、長短期記憶網路(Long Short-Term Memory, LSTM)、門控循環單元網路(Gate Recurrent Unit, GRU)等模式處理股價預測問題，期能獲致更佳的預測結果 (Gao *et al.*, 2021)。Kuo and Xue (1998) 的研究就曾指出，人工神經網路模式較時間序列模式有更多的靈活性。Chang *et al.* (2021) 在股票市場之實證預測研究中指出，以 LSTM 的研究方法優於傳統的時間序列法，Xiao and Su (2022) 亦實證指出用 LSTM 法預測標普 500

指數，優於使用 ARIMA 法。最近兩年一些學者並提出結合 LSTM 與其他神經網路法的混合模型，以解決金融商品預測問題，並獲得較為優異的預測結果，如 CNN-LSTM (Li and Dai, 2020)、CNN-BILSTM-BIRNN (Das *et al.*, 2021)、MRC-LSTM (Guo *et al.*, 2021) 等。

由於多尺度殘差卷積神經網路 (Multi-scale Residual Convolutional Neural Network, MRC) 不僅能夠檢測多元時間序列中不同時間尺度的特徵，並且能融合這些特徵進行檢測。而 LSTM 雖具有分辨長序列資料之間的關聯之優勢，但因組成複雜，所以在處理龐大的資料時會較為耗時，因而有 GRU 的誕生改良，GRU 不但保有 LSTM 的優勢，內部結構更是相對於 LSTM 簡單，所以能夠節省計算時間而更有效率 (Chung *et al.*, 2014; Gruber and Jockisch, 2020)。若採用 Bi-GRU 的雙向結構，則能使模型能在理解過去資料影響輸出資料的同時，更是可以接受反向的輸入資料，以分析未來的資料如何對應到過去的資料。由於比特幣市場具有高度的波動性，同時市值容易受到世界各種重大之複雜因素影響變動，因此本計畫提出了一種結合多尺度殘差卷積神經網路 (Multi-scale Residual Convolutional Neural Network, MRC) 和雙向門控循環單元網路 (Bi-GRU) 的 MRC-BiGRU 新穎混合模型，來進行比特幣價格的預測，並將此混合模型與其他混合模型，如 MRC-BiLSTM、MRC-GRU 等之預測結果做比較，希冀能更精確的預測比特幣價格，能幫助投資人做出正確的投資決策以避免投資損失。

### 第三節 研究架構與流程

本研究內容共分成五章，研究流程圖如圖 1.6 所示。第一章為緒論，主要說明研究背景、動機、目的與整體架構等。第二章進行文獻探討，整理比特幣價格預測相關文獻，並根據模型種類、特徵值將其分析進行文獻歸類。第三章為研究方法，深度解析本研究使用的模型基礎原理、架構等。第四章為實驗結果與分析，說明訓練資料的收集與處理、模型建構參數、實驗規劃與環境，再使用本研究選定的預測模型評估指標，進行評估各項實驗模型。第五章為結論與建議，回顧本研究之實證結果與目的，提出未來研究的可行方向與建議。

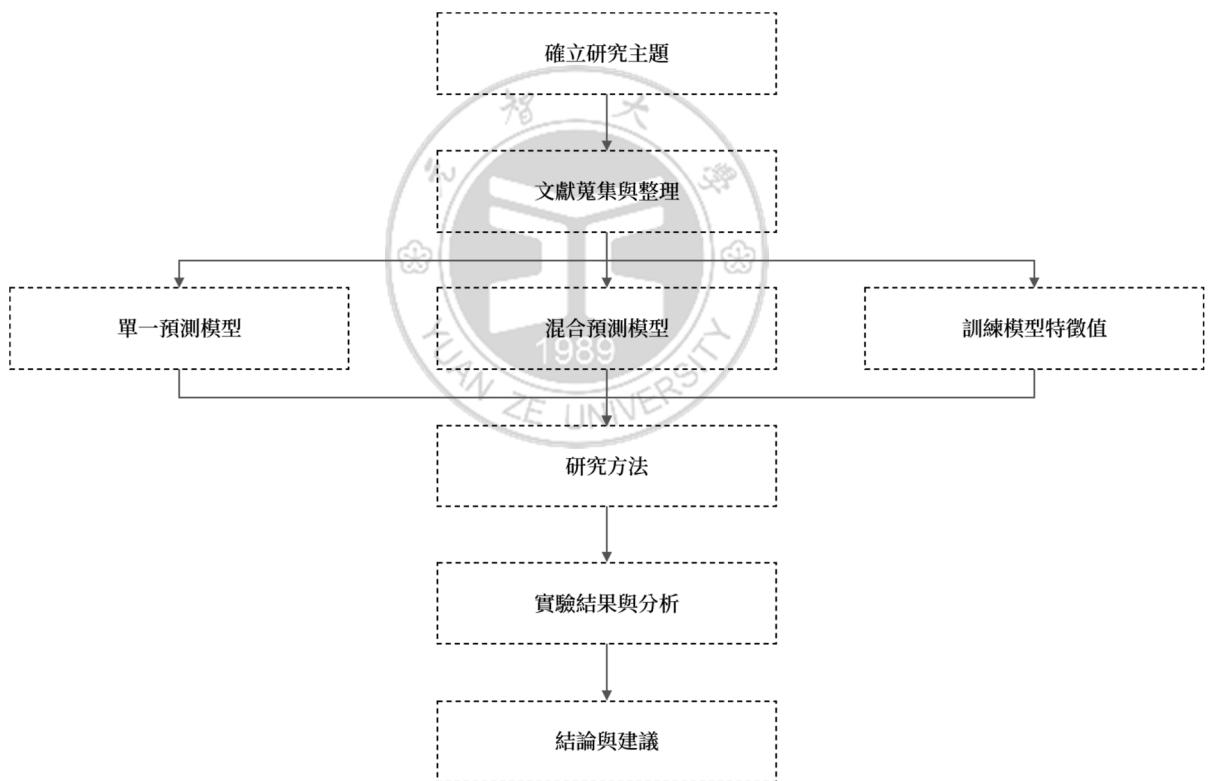


圖 1.6 研究流程圖

## 第二章 文獻探討

### 第一節 比特幣價格預測

由於人工神經網路（Artificial Neural Network, ANN）具有很強的非線性逼近能力（Approximation Ability）、自我學習能力、和適應能力，因此，近期多位學者紛紛利用人工神經網路模擬預測各種金融產品之非線性連續函數。而目前處理各種金融產品預測問題的人工神經網路方法，主要有卷積神經網路（Convolutional Neural Network, CNN）、循環神經網路（Recurrent neural network, RNN）、長短期記憶網路（Long Short-Term Memory, LSTM）、門控循環單元網路（Gate Recurrent Unit, GRU）與各種混合神經模型等。本研究整理近年來使用人工神經網路方法預測比特幣價格之相關文獻如表 2.1 所示。

表 2.1 比特幣價格預測文獻整理

學者	研究方法	資料描述
Dutta <i>et al.</i> (2020)	<b>Proposed Model :</b> GRU (Best) <b>Control Model :</b> LSTM	資料期間： 2010/1/1 ~ 2019/6/30 自變數 (X)： 比特幣價格、比特幣價格波動、比特幣礦工收入、交易量、交易金額、哈希率、資金供應、梅特卡夫-UTXO、區塊大小、Google Trend、VIX、黃金價格、美元指數、美國債券收益率、美國債券收益率、美國債券收益率、MACD、瑞波幣價格、標準普爾 500 指數滯後一天的市場回報、Interest Rates 因變數 (Y)： 比特幣價格
Alkhodhairi <i>et al.</i> (2021)	<b>Proposed Model :</b> LSTM (Best in 4 Hr data) 、 GRU (Best in 12 Hr and 24 Hr data)	資料期間： 2012/1/1 ~ 2020/9/14 自變數 (X)： 開盤價、最高價、最低價、收盤價 因變數 (Y)： 比特幣價格 (4 Hr、12 Hr、24 Hr)

學者	研究方法	資料描述
Liu <i>et al.</i> (2021)	<b>Proposed Model :</b> LSTM ( Best ) <b>Control Model :</b> SVM、MLP、TCN、Transformer	<b>資料期間 :</b> 2017/8/17 ~ 2021/2/27 <b>自變數 (X) :</b> 未特別提及 <b>因變數 (Y) :</b> 比特幣價格(1 Hr)
Nithyakani <i>et al.</i> (2021)	<b>Proposed Model :</b> Bi-LSTM ( Best ) <b>Control Model :</b> Linear Regression、XGBoost	<b>資料期間 :</b> 2012/1/1 ~ 2020/9/14 <b>自變數 (X) :</b> 開盤價、最高價、最低價、收盤價、交易量、每個時間點的加權價格信息 <b>因變數 (Y) :</b> 比特幣價格 (日)
Saadatmand and Chahoki (2023)	<b>Proposed Model :</b> Bi-GRU ( Best ) <b>Control Model :</b> Bi-LSTM(Best)、LSTM、GRU	<b>資料期間 :</b> 2012/1/1 ~ 2020/12/31 <b>自變數 (X) :</b> 價格變化方向、收盤價、Twitter 情緒分析 <b>因變數 (Y) :</b> 比特幣價格 (日)
Cavalli and Amoretti (2021)	<b>Proposed Model :</b> CNN ( Best ) <b>Control Model :</b> LSTM	<b>資料期間 :</b> 2013/4/28 ~ 2020/2/15 <b>自變數 (X) :</b> 比特幣收盤價、十個財務指標 (SMA、CCI、MFI、EMA、RSI、MACD、SRS、WMA、MVI、PVI)、Twitter 情緒分析、比特幣礦工分析 (Block Size、Fees、Mining Difficulty、Number of Transactions) <b>因變數 (Y) :</b> 比特幣趨勢 (日)
Li and Dai (2020)	<b>Proposed Model :</b> CNN-LSTM ( Best ) <b>Control Model :</b> BP、CNN、LSTM	<b>資料期間 :</b> 2016/12/30~2018/8/31 <b>自變數 (X) :</b> 開盤價、最高價、最低價、收盤價、交易量、交易金額、RSI、MFI、OBV 黃金價格、人民幣兌美元匯率、NYSE Index、NASDAQ Index、S&P 500 Index、聯邦基金匯率原油期貨價格 百度指數(百度瀏覽器搜尋比特幣次數) <b>因變數 (Y) :</b> 比特幣收盤價 (日)

學者	研究方法	資料描述
Das <i>et al.</i> (2021)	<b>Proposed Model :</b> CNN-BiLSTM-BiRNN ( Best ) <b>Control Model :</b> CNN、LSTM、 Bi-LSTM、LSTM、 Bi-RNN、RNN、2 Layer NN、GRU、 ARIMA	<b>資料期間 :</b> 2017/1/1~2019/12/31 <b>自變數 (X) :</b> Namely Date、開盤價、收盤價、最高價、最低價、調整後的收盤價、交易量 <b>因變數 (Y) :</b> 比特幣收盤價 (日)
Guo <i>et al.</i> (2021)	<b>Proposed Model :</b> MRC-LSTM ( Best ) <b>Control Model :</b> MLP、LSTM、CNN、 CNN-LSTM	<b>資料期間 :</b> 2015/10/25~2020/10/17 <b>自變數 (X) :</b> 開盤價、收盤價、最高價、最低價、加權價、比特幣交易量、市場幣種、標準普爾 500 指數、GVZ、VIX <b>因變數 (Y) :</b> 比特幣收盤價

資料來源：本研究整理

從表 2.1 可以得知，關於比特幣價格預測文獻各項研究中所使用的預測模型，可分為單一預測模型與混合預測模型。



## 第二節 單一預測模型預測比特幣價格

Dutta *et al.* (2020) 使用長短期記憶網路 (LSTM) 和門控循環單元 (GRU) 預測比特幣價格。實驗結果顯示，兩者皆優於傳統的機器學習模型，且門控循環單元 (GRU) 的訓練結果較為出色。Alkhodhairi *et al.* (2021) 亦使用長短期記憶網路 (LSTM) 和門控循環單元 (GRU) 作為深度學習框架，以預測比特幣不同級別時間單位 (4 小時、12 小時、24 小時) 的價格變動。實驗結果顯示，在 4 小時級別的長短期記憶網路 (LSTM) 模型優於門控循環單元 (GRU) 模型；而在 12、24 小時級別的門控循環單元 (GRU) 模型則優於長短期記憶網路 (LSTM) 模型。與 Alkhodhairi *et al.* (2021) 不同，Nithyakani *et al.* (2021) 使用雙向長短期記憶網路 (Bi-LSTM) 作為深度學習框架，並證實以 Bi-LSTM 進行的預測，優於以線性回歸 (Linear Regression) 與極限梯度提升 (XGBoost) 的方式。Liu *et al.* (2021) 亦使用長短期記憶網路 (LSTM) 作為比特幣交易策略的深度學習框架，並與 SVM、MLP、TCN、Transformer 所訓練的模型，實驗結果顯示，LSTM 擁有最優秀的結果。Saadatmand and Chahoki (2023) 使用雙向控循環單元 (Bi-GRU) 模型作為深度學習框架，與控循環單元 (GRU)、長短期記憶網路 (LSTM)、雙向長短期記憶網路 (Bi-LSTM) 所訓練出的模型比較，實驗結果顯示 Bi-GRU 與 Bi-LSTM 有不相上下的預測能力，但前者的訓練時間更迅速。Cavalli and Amoretti (2021) 提出了利用卷積神經網路 (CNN) 與長短期記憶網路 (LSTM) 分別作為深度學習框架。實驗結果顯示，卷積神經網路 (CNN) 的預測結果優於長短期記憶網路 (LSTM)。

上述各研究皆是訓練單一模型如長短期記憶網路 (LSTM)、門控循環單元 (GRU)、卷積神經網路 (CNN) 進行比特幣價格預測，學者嘗試使用各種不同的方法，以期找到最佳的預測模式，儘管單一模型之訓練結果優於傳統統計方法如線性回歸 (Linear Regression) 與極限梯度提升 (XGBoost) 等，然而單一模型有其侷限，以致無法取得優於混合模型預測的效果。

### 第三節 混合預測模型預測比特幣價格

Li and Dai (2020) 提出了一種基於卷積神經網路 (CNN) 和長短期記憶網路 (LSTM) 的混合神經網路模型 (CNN-LSTM)，此混合模型使用 CNN 進行特徵提取，再將特徵向量輸入到 LSTM 中，用於訓練和預測比特幣的短期價格。實驗結果顯示，卷積神經網路 (CNN) 在值的預測方面比長短期記憶網路 (LSTM) 更有優勢，但長短期記憶網路 (LSTM) 在漲跌方向預測方面表現較好，而混合模型 (CNN-LSTM) 無論是在數值的預測，或是方向預測方面的表現都比單一模型表現優秀。

然而在 LSTM 的使用上並未使用雙向長短期記憶網路 (即 Bi-LSTM)，忽略了反向資料分析的重要性。Das *et al.* (2021) 則嘗試將混合模型 (CNN-LSTM) 的長短期記憶網路 (LSTM) 改為雙向長短期記憶網路 (即 Bi-LSTM)，並加上一層雙向遞迴神經網路 (Bi-RNN)，最後將混合模型 (CNN-BiLSTM-BiRNN) 與其他單一方法如 CNN、LSTM、Bi-LSTM、LSTM、Bi-RNN、RNN、2 Layer NN、GRU、ARIMA 所訓練出的模型比較。實驗結果顯示，雙向混合模型 (CNN-BiLSTM-BiRNN) 擁有最優秀的結果。

除此之外，當類神經網路模型的深度過深時，容易出現梯度爆炸、消失等問題。為解決模型過深的梯度問題，Guo *et al.* (2021) 提出了一種稱為 MRC-LSTM 的新混合模型，基於卷積神經網路 (CNN) 的多尺度殘差卷積神經網路 (Multi-scale Residual Convolutional Neural Network，MRC) 能夠檢測並融合多元時間序列中不同時間尺度的特徵，將特徵向量輸入到 LSTM 中來實現比特幣收盤價預測。透過 MRC-LSTM 混合模型，能有效地學習多變量時間序列的趨勢和交互作用。該研究比較 MLP、LSTM、CNN、CNN-LSTM 所訓練的模型，結果顯示，混合模型 MRC-LSTM 擁有最優秀的結果。

然而 MRC-LSTM 並未使用效果良好的的雙向長短期記憶網路（Bi-LSTM），忽略了反向資料分析的重要性。因此，本研究擬使用多尺度殘差卷積塊（MRC）於不同的時間序列提取特徵，並與具有雙向分析資料能力的雙向門控循環單元網路（Bi-GRU）結合依時間序列學習關係模式，形成全新的混合網路架構（MRC-BiGRU）來執行比特幣時間序列預測。



## 第四節 訓練模型之特徵值

除此之外，本研究並整理前述文獻之訓練模型的特徵值，分為以下 4 種類別，做為本研究的參考：

### (一) 交易信息類：

開盤價、收盤價、最高價、最低價、交易量、交易金額、挖礦難度、區塊大小等。

### (二) 技術指標類：

RSI、KD、MACD、SMA 等。

### (三) 宏觀經濟變量類：

黃金價格、原油價格、美元匯率、標準普爾 500 指數、紐約證券交易所指數、納斯達克指數、聯邦基金利率。

### (四) 消息層面類：

Twitter 相關文章情緒分析、Google、百度趨勢。



## 第三章 研究方法

### 第一節 卷積神經網路 (CNN)

由於混合預測模型優於單一預測模型，而在眾多混合預測模型中，MRC-LSTM (Guo *et al.*, 2021) 擁有最優秀的預測結果。但根據 Dutta *et al.* (2021) 研究顯示，GRU 表現比 LSTM 好，而且雙向架構可使模型具備雙向分析資料能力，使模型具有更準確的預測效果 (Nithyakani *et al.*, 2021; Das *et al.*, 2021)。故本研究擬結合基於一維卷積 (Conv1D) 的多尺度殘差卷積神經網路 (MRC) 於不同的時間序列提取特徵，並與具有雙向分析資料能力的雙向門控循環單元網路 (Bi-GRU)，形成全新的混合模型 (MRC-BiGRU) 來執行比特幣時間序列預測，並一一分述如下。

MRC 結合 Bi-GRU 用於從數據集中提取在不同時間序列尺度上具有的潛在特徵。本研究使用了不同尺度的一維卷積 (Conv1D) 構建了一個多旁路分支的卷積層，並在此借鑒了 DenseNet 中結合不同卷積核的概念，透過 Channel-wise Concatenation 來組合多個時間尺度的局部特徵。

卷積神經網路 (Convolutional Neural Network, CNN) 是近年來最流行、應用最廣泛的人工神經網路 (Artificial Neural Network, ANN) 之一，其設計概念主要為模仿生物視覺細胞對局部信息的感知，基本構造包含卷積層 (Convolution Layer)、池化層 (Pooling Layer) 和全連接層 (Fully Connection Layer)，說明如下：

#### (一) 卷積層 (Convolution Layer)：

卷積層利用共享權值的卷積核 (Kernel) 進行資料特徵萃取，有效減少了參數數量並降低神經網路訓練的複雜度，之後得到的輸出結果稱為特徵圖 (Feature Map)，如圖 3.1 所示。

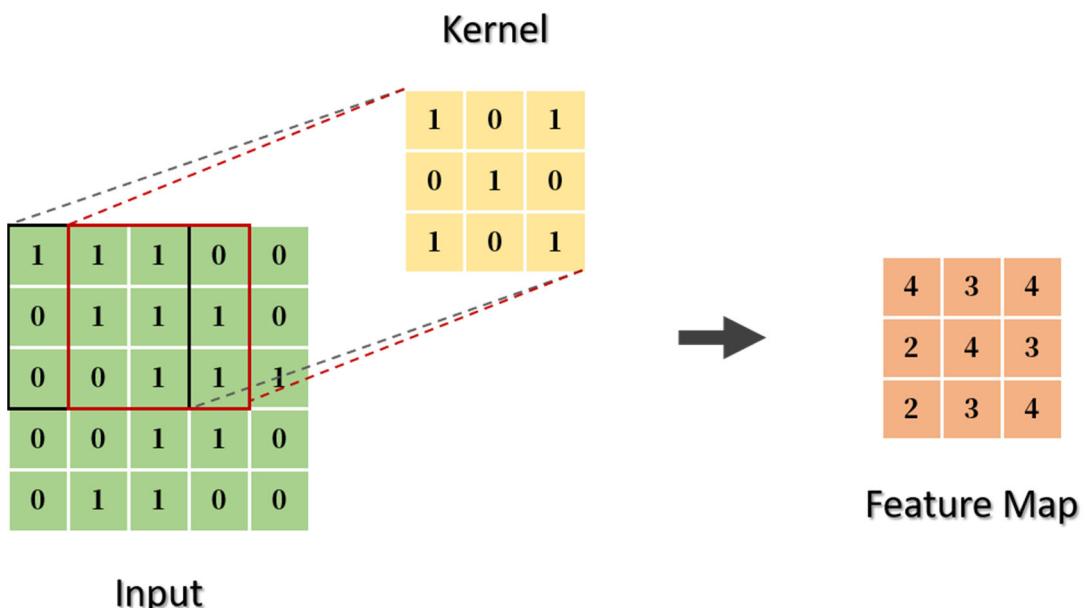


圖 3.1 卷積層概念圖

資料來源：本研究整理

## (二) 池化層 (Pooling Layer)：

池化層是一種非線性形式的降採樣層，是保留重要特徵資訊的方法，池化後的資訊保留了原本輸入資料各範圍的最大值 (Max pooling) 或平均值 (Average pooling)，在不大幅改變特徵的情況下，減少後續計算成本，其概念如圖 3.2 所示。

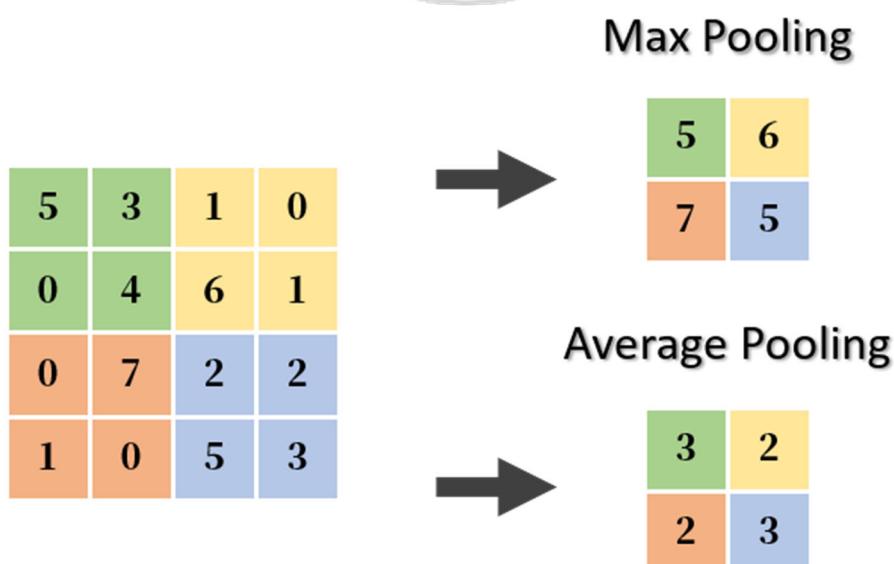


圖 3.2 池化層概念圖

資料來源：本研究整理

### (三) 全連接層 (Fully Connection Layer) :

經過上述特徵擷取與保留必要特徵等步驟後，全連接層對特徵進行加權運算，  
在經過權重的計算之後得以辨識出輸入資料屬於何種分類。



## 第二節 多尺度殘差卷積神經網路 (MRC)

一般來說，類神經網路模型的深度層數與訓練表現有極大的關係。然而當類神經網路模型的深度過深時，容易出現梯度爆炸、消失等問題，為了解決上述問題，He *et al.* (2016) 提出一種由多個 Residual Units 所組成的新型模型 Resnet (Deep Residual Learning for Image Recognition)，通過建立前面層與後面層之間的捷徑連接 (Shortcuts Connection)，連結有效聚合歷史信息，可在一定程度上減少模型加深時特徵和信息損失問題，有助於訓練過程中梯度的反向傳播，從而能訓練出更深的 CNN 網路，方法簡單卻使得深層網路訓練變得容易許多，在一定程度上減少模型加深時特徵和信息損失問題，其概念圖如圖 3.3 所示，網路各層之關係如式(1)所示。

$$x_l = H(x_{l-1}) + x_{l-1} \quad (1)$$

其中  $x_1$  代表網路在第 1 層的輸出， $H_l$  代表非線性轉化函數 (Non-linear Transformation)，它是一個組合運算過程，可以包括一系列的卷積層 (Convolution Layer)、激勵函數 (Activation Function)、池化層 (Pooling Layer) 等運算。

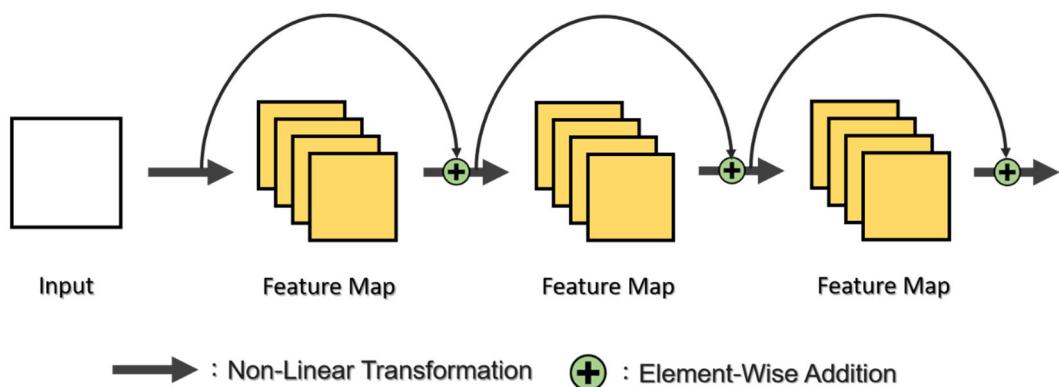


圖 3.3 ResNet 概念圖

資料來源：本研究整理

DenseNet 於 2017 年被 Huang, Maaten, and Weinberger 等學者提出，此種方法

受到 ResNet 的捷徑連接（Shortcuts Connection）思想啟發，其最大的特色在於特徵重用（Feature Reuse），將前面所有捲積層產生的 Feature Map 作為輸入，然後將其進行 Channel-Wise Concatenation 以組合不同的局部特徵聚合信息，如此一來便可以保留前面的特徵，讓模型整體結構更加緊密，提高了模型的信息、梯度流動，DenseNet 概念圖如圖 3.4 所示，網路各層之關係如式(2)所示。

$$x_l = H_l([x_0, x_1, \dots, x_{l-1}]) \quad (2)$$

其中  $x_1$  代表網路在第 l 層的輸出， $H_l$  代表非線性轉化函數（Non-linear Transformation），它是一個組合運算過程，可以包括一系列的卷積層（Convolution Layer）、激勵函數（Activation Function）、池化層（Pooling Layer）等運算。

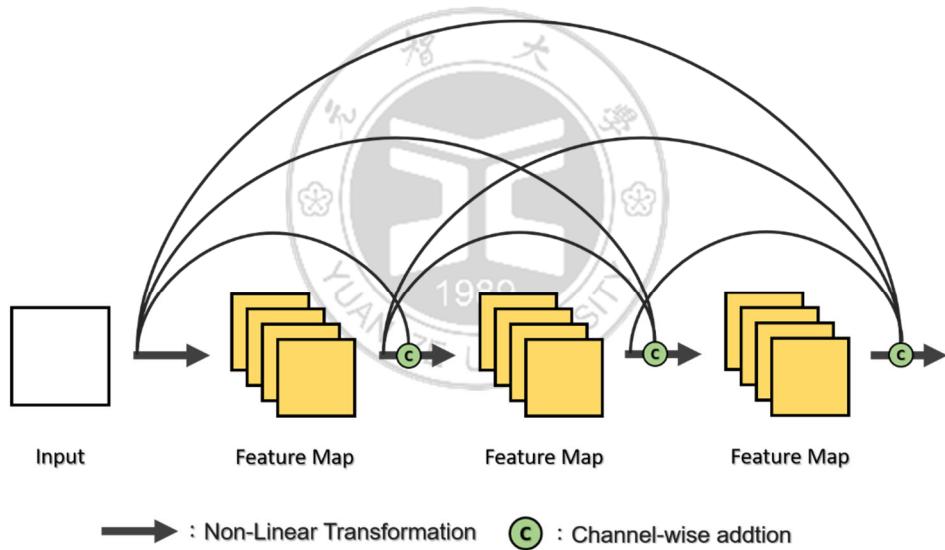
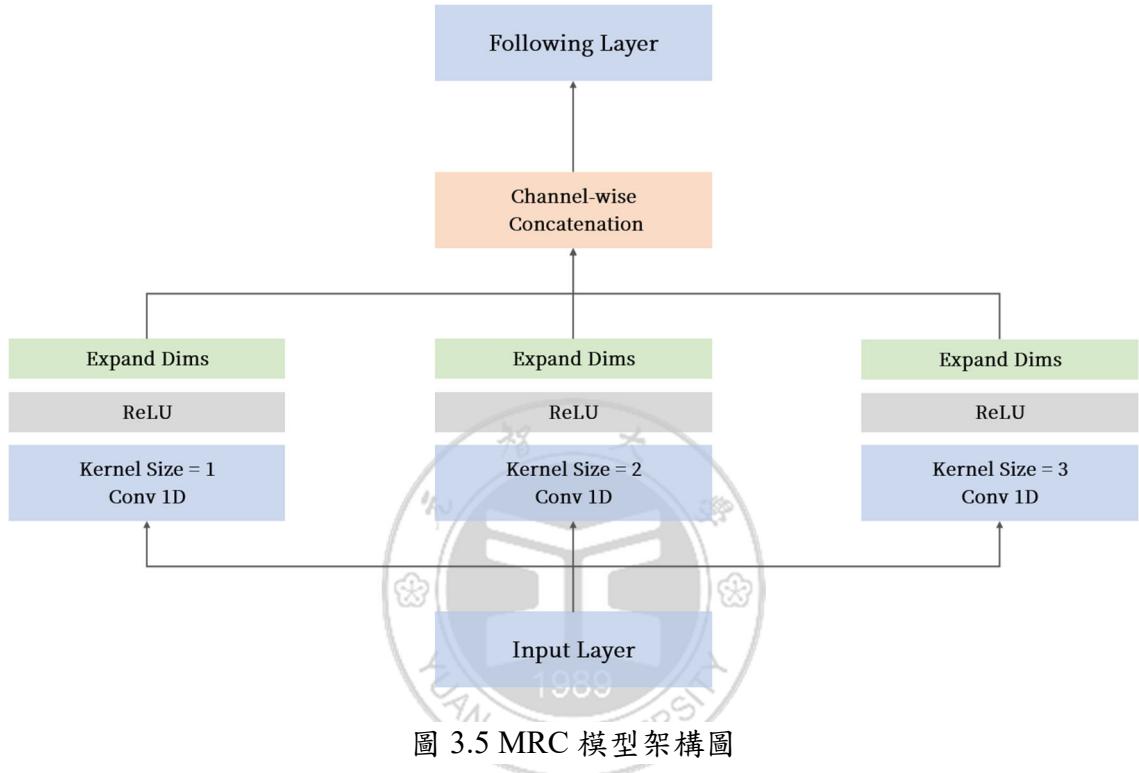


圖 3.4 DenseNet 概念圖

資料來源：本研究整理

本研究的多尺度殘差卷積神經網路（Multi-scale Residual Convolutional Neural Network，MRC）模型架構圖詳如圖 3.5 所示。首先，使用三個不同 Kernel Size 的一維卷積（Conv1D）構建了一個多旁路分支的卷積層提取 Feature map，並在此借鑒了 DenseNet 中結合不同卷積核的概念，透過 Channel-wise Concatenation 來組合多個時間尺度的局部特徵聚合信息，然而數據集經過不同尺度的 Kernel 時，會

產生不同大小的 Feature map，導致 Feature map 組合失敗。為了保持 Feature map 大小比例一致，需要在一維卷積時進行 Zero Padding 滿足 Feature map 大小一致，最後將不同大小的 Feature map 連接並傳入後續層，以進一步整合不同時間序列的歷史訊息。



同時，在保持 Feature map 比例不變的情況下引入激活函數( Activation functions )納入網路添加非線性，常見的激活函數包括 Sigmoid、Tanh、ReLU、SELU，如圖 3.6 所示。由於整流線性單位函數（Rectified Linear Unit, ReLU）是一種非常簡單且有效的激活函數，在 CNN 中使用 ReLU 作為隱藏層的激活函數是一種常見的選擇。因此，本研究選取整流線性單位函數（Rectified Linear Unit, ReLU）作為 MRC 網路的非線性激活函數，如式(3)所示。

$$f(x) = \begin{cases} x, & (x > 0) \\ 0, & (x \leq 0) \end{cases} \quad (3)$$

實現跨通道信息交互和擴展後，經過不同 Kernel 提取的 Feature map 將被融合，從這些分層特徵中提取有用的信息後將輸入至後續循環神經網路，如雙向門控循環單元網路（Bi-GRU）。

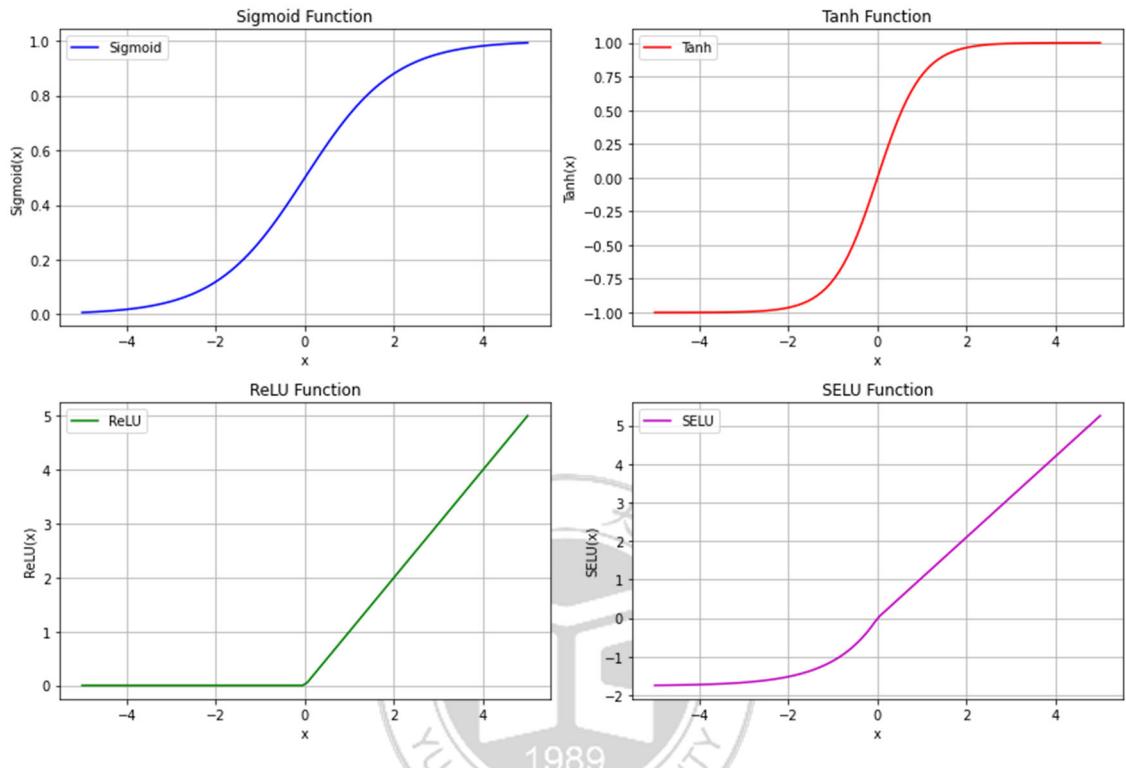


圖 3.6 常見的激活函數

資料來源：本研究整理

### 第三節 長短期記憶網路（LSTM）

循環神經網路 (Recurrent neural network, RNN) 的原型於 1986 年被 Rumelhart, Hinton, and Williams 等學者提出，該網路特點在於其獨特的循環結構，能將前一個時間步驟的隱藏層輸出作為下一個時間步驟的部分輸入，從而影響當前時間步驟的輸出。這種特性使得 RNN 在處理具有連續性和依賴性的輸入序列時表現出色，能夠捕捉數據中時間上的相關性，因此對於具有連續性的輸入序列有著良好的成果，而被廣泛應用時間序列領域的各項任務，如金融市場分析、意圖識別、文本生成、機器翻譯等。

對於短序列輸入之任務，RNN 普遍取得優秀的成果，然而在處理長序列之間的關聯時，通過實證結果表現相對不佳，原因是在進行反向傳播時，過長的序列導致梯度的計算異常，發生梯度消失、爆炸等問題。如果在訓練過程中發生了梯度消失問題，權重將無法被更新，最終導致訓練失敗；梯度爆炸將導致梯度過大，且大幅度更新網路參數，在極端情況下，結果將產生溢出 (NaN 值)。兩者皆是導致訓練結果不佳的主因。

為解決上述問題，Hochreiter and Schmidhuber (1997) 提出長短期記憶網路 (Long Short-Term Memory, LSTM)，此為 RNN 模型的變體，結構核心由遺忘門、輸入門、細胞狀態、輸出門組成，與 RNN 模型相比，可以有效緩解梯度消失或爆炸問題，如圖 3.7。

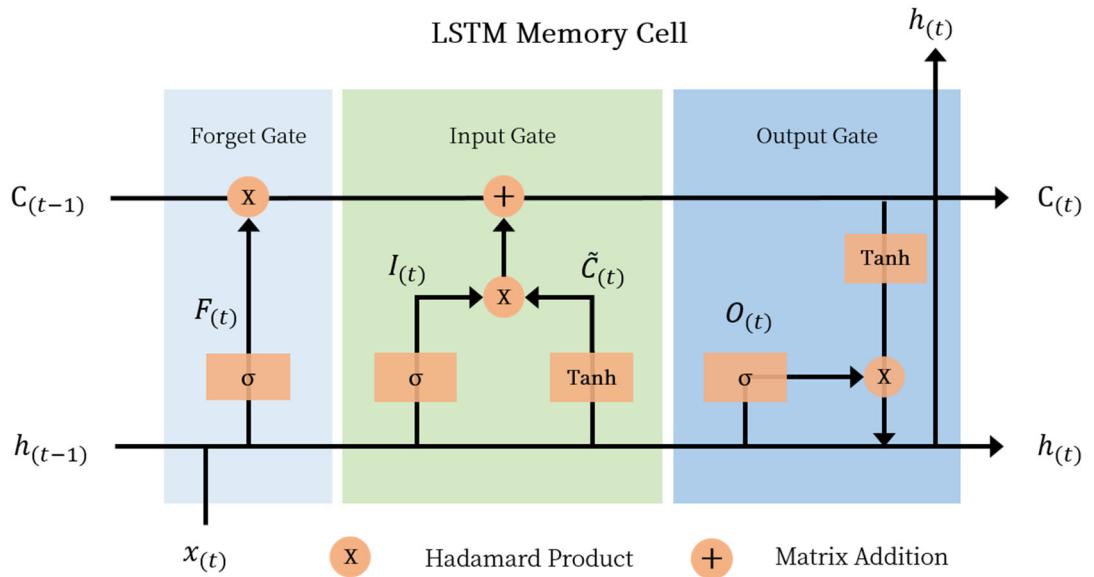


圖 3.7 LSTM 記憶元示意圖

資料來源：本研究整理

首先計算遺忘門 (Forget Gate)、輸入門 (Input Gate) 和輸出門 (Output Gate) 的門值，分別是  $F_t$ 、 $I_t$  和  $O_t$ ，如式(3)、(4)、(5)所示，計算方法使用上一個 LSTM 傳遞下來的資訊 ( $h_{t-1}$ ) 與當前節點的輸入 ( $x_t$ ) 拼接，進行 sigmoid 激活，其中  $W_f$ 、 $W_I$ 、 $W_O$ 、 $U_f$ 、 $U_I$ 、 $U_O$  屬於權重參數； $b_f$ 、 $b_I$ 、 $b_O$  屬於偏值參數。候選記憶元 (Candidate Memory Cell)  $\tilde{C}_t$  的計算方式與上述三個門值相似，但是使用 tanh 進行激活， $W_C$ 、 $U_C$  屬於權重參數； $b_C$  屬於偏值參數，如式(6)所示。

$$F_t = \sigma(W_f \cdot x_t + U_f \cdot h_{t-1} + b_f) \quad (4)$$

$$I_t = \sigma(W_I \cdot x_t + U_I \cdot h_{t-1} + b_I) \quad (5)$$

$$O_t = \sigma(W_O \cdot x_t + U_O \cdot h_{t-1} + b_O) \quad (6)$$

$$\tilde{C}_t = \tanh(W_C \cdot x_t + U_C \cdot h_{t-1} + b_C) \quad (7)$$

記憶元狀態 (Memory Cell State)  $C_t$  的計算方式，如式(7)所示，主要由遺忘門、輸入門控制，遺忘門決定保留多少的前一個記憶元內容  $C_{t-1}$ ；輸入門決定採用多少候選記憶元  $\tilde{C}_t$  的資訊，這種設計有效的緩解了梯度消失的問題，更加有效的捕捉長序列的依賴關係。

$$C_t = F_t \cdot C_{t-1} + I_t \cdot \tilde{C}_t \quad (8)$$

最後，隱藏狀態（Hidden state） $h_t$ 的計算方式，如式(8)所示，由輸出門 $O_t$ 與記憶元狀態 $C_t$ 組成，並使用  $\tanh$  進行激活，相對於記憶元狀態 $C_t$ 擅長保留遠期記憶， $h_t$ 更能保留近期記憶。

$$h_t = O_t \cdot \tanh(C_t) \quad (9)$$



## 第四節 雙向門控循環單元網路 (Bi-GRU)

LSTM 的組成相對傳統 RNN 複雜許多，因此訓練效率在同等運算能力之下較傳統 RNN 差，Chung *et al.* (2014) 為了解決 LSTM 訓練能力效率不佳的問題，提出了門控循環單元 (Gate Recurrent Unit, GRU)，結構核心由更新門、重置門組成。與 LSTM 相同，GRU 一樣可以有效緩解梯度消失或爆炸問題，且其組成結構更為單純，因此運算效率相對 LSTM 更高，如圖 3.8。

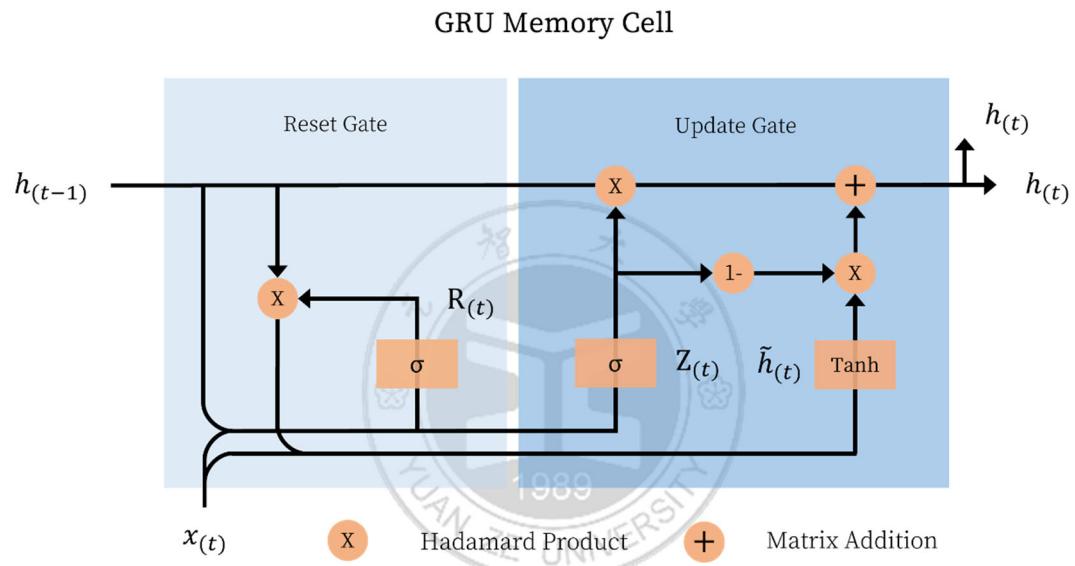


圖 3.8 GRU 記憶元示意圖

資料來源：本研究整理

首先計算重置門(Reset Gate)和更新門(Update Gate)的門值，分別是 $R_t$ 和 $Z_t$ ，如式(10)、(11)所示，計算方法使用上一個 GRU 傳遞下來的資訊 ( $h_{t-1}$ ) 與當前節點的輸入 ( $x_t$ ) 拼接進行線性變換，再經由 Sigmoid 激活，其中  $W_R$ 、 $W_Z$ 、 $U_R$ 、 $U_Z$  屬於權重參數； $b_R$ 、 $b_Z$  屬於偏值參數。

$$R_t = \sigma(W_R \cdot x_t + U_R \cdot h_{t-1} + b_R) \quad (10)$$

$$Z_t = \sigma(W_Z \cdot x_t + U_Z \cdot h_{t-1} + b_Z) \quad (11)$$

重置門的門值 $R_t$ 作用在 $h_{t-1}$ 上，代表控制上一時間步傳來的信息有多少可以被

利用，接著使用這個重置後的  $h_{t-1}$  進行基本的 RNN 計算，亦即與  $x_t$  合併進行線性變化，透過  $\tanh$  激活，得到候選隱藏狀態（Candidate Hidden State） $\tilde{h}_t$ ，如式(12)所示。

$$\tilde{h}_t = \tanh(W_h \cdot x_t + (R_t \cdot h_{t-1}) \cdot U_h + b_h) \quad (12)$$

最後更新門的門值會作用在新的  $h_{t-1}$  上，而 1-門值會作用在  $\tilde{h}_t$  上，隨後將兩者的結果相加，得到最終的隱藏狀態（Hidden State） $h_t$ ，如式(13)所示，這個過程意味著更新門有能力保留之前的結果。這種設計有效的緩解了梯度消失的問題，相對於重置門（Reset Gate）擅長捕捉近期依賴關係記憶，更新門（Update Gate）則是更擅長保留遠期依賴關係記憶。

$$h_t = (1 - Z_t) \cdot \tilde{h}_t + Z_t \cdot h_{t-1} \quad (13)$$

Schuster and Paliwal 於 1997 年提出了將常規的循環神經網路( Recurrent Neural Networks, RNN )擴展成雙向循環神經網路( Bidirectional Recurrent Neural Networks, Bi-RNN )，在原先只有正向傳遞訊息的隱藏層中，新增了反向傳遞訊息的設計，以便更靈活地處理資訊，奠定了雙向模型的基礎。

雙向門控循環單元網路 ( Bi-directional Gate Recurrent Unit, Bi-GRU ) 模型是門控循環單元網路 ( Gate Recurrent Unit, GRU ) 模型的一種延伸，概念圖如圖 3.9 所示。傳統單向的 GRU 只能正向捕捉輸入資料信息，然而 Bi-GRU，如式(16)所示，由兩個 GRU 組成：一個是正向 GRU 模型，如式(14)所示，接受正向的輸入資料，理解過去的資料如何影響現在的輸出；一個是反向 GRU 模型，如式(15)所示，接受反向的輸入資料，分析未來的資料如何對應到過去的資料，如以下所示。

$$h_t^{forward} = (1 - Z_t) \cdot \widetilde{h}_t^{forward} + Z_t \cdot h_{t-1}^{forward} \quad (14)$$

$$h_t^{backward} = (1 - Z_t) \cdot \widetilde{h}_t^{backward} + Z_t \cdot h_{t-1}^{backward} \quad (15)$$

$$h_t = [h_t^{forward}, h_t^{backward}] \quad (16)$$

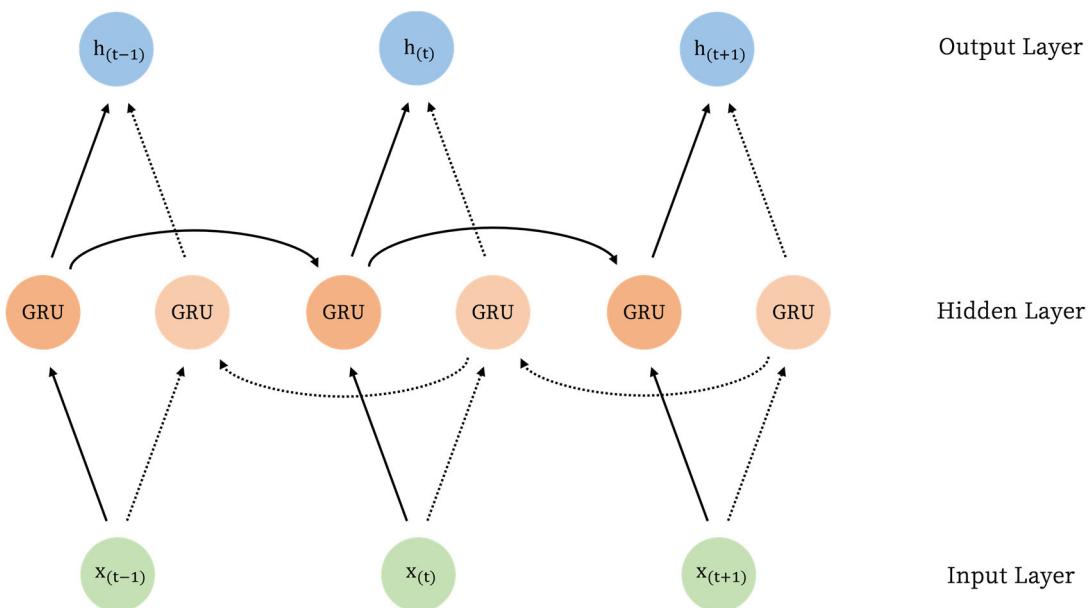


圖 3.9 Bi-GRU 概念圖

資料來源：本研究整理



## 第五節 多尺度卷積雙向門控循環單元網路 (MRC-BiGRU)

本研究以多尺度殘差卷積神經網路 (Multi-scale Residual Convolutional Neural Network, MRC) 結合雙向門控循環單元網路 (Bi-directional Gate Recurrent Unit, Bi-GRU) 建構比特幣價格預測模型，透過由一維卷積 (Conv1D) 組成的多尺度殘差卷積神經網路 (MRC) 於不同的時間序列提取特徵後，連接具有雙向分析資料能力的雙向門控循環單元網路 (Bi-GRU)，形成全新的混合模型 (MRC-BiGRU) 來執行比特幣時間序列預測，其流程圖如圖 3.10 所示。

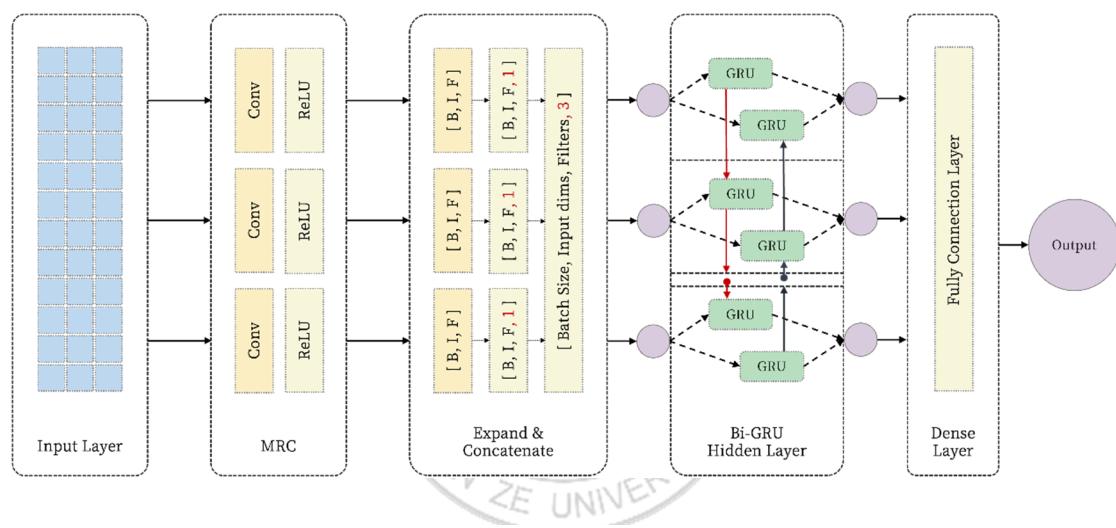


圖 3.10 MRC-BiGRU 模型

資料來源：本研究整理

本研究將採用三步驟程序建構比特幣價格的預測模型，其架構圖如圖 3.11 所示，包含：

- (一) 訓練資料收集 (Training data collection)、訓練資料處理 (Training data processing)。(二) 建構比特幣價格預測模型 (Constructing prediction model)。
- (三) 預測模型評估 (Model Evaluation)。

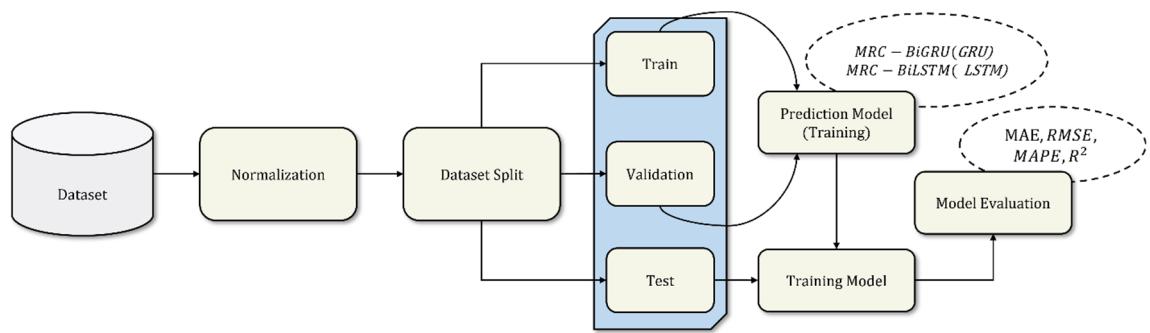


圖 3.11 研究進行步驟流程圖

資料來源：本研究整理



## 第四章 實驗結果與分析

### 第一節 訓練資料收集

經本研究整理比特幣價格預測文獻所使用模型自變數，發現自變數種類主要分成四種類別，包含交易信息類、技術指標類、宏觀經濟變量類與消息層面類，其中所有預測文獻所使用的自變數都包含交易信息類，因此本研究亦採用此類資訊當作預測模型的輸入特徵值。比特幣價格資訊包含開盤價（Open Price）、收盤價（Close Price）、最高價（High Price）、最低價（Low Price）、交易量（Volume）。

相對強弱指標（Relative Strength Index, RSI）是一種常見的技術分析指標，通過測量價格變化的速度和幅度提供超買和超賣的信號，被廣泛應用於衡量資產價格的方向和強度，有助於捕捉市場情緒的變化，引導交易者做出決策。RSI 作為一種經過歷史驗證的技術指標，在各種金融市場和資產類別中都顯示出一定的預測能力。儘管過去表現不能保證未來結果，但 RSI 的廣泛應用和可靠性為其在比特幣價格預測中的應用提供了一定的合理性，作為金融市場中成熟的技術指標，本研究使用 RSI 作為技術指標類模型自變數。

比特幣的價格波動被認為與其他宏觀經濟變量相關，Vassiliadis *et al.* (2017) 研究了「比特幣」與「股票、黃金、原油市場」的關聯，結果表明比特幣價格與股票市場指數、黃金價格、原油價格存在高度相關性。Maghyereh and Abdoh(2021) 研究了「比特幣」與「股票市場」的關聯，結果表明比特幣與美國股票市場之間的聯動是高度相關的。最終，本研究決定採用標準普爾 500 指數(Standard & Poor's 500, S&P500) 等宏觀經濟變量作為模型自變數。

投資者在交易市場的行為非常情緒化，且加密貨幣市場波動劇烈，因此更容易做出非理性的判斷，如 FOMO (Fear of missing out, 害怕錯過)、FUD ( Fear、Uncertainty、Doubt, 恐懼、疑惑、懷疑)，為了避免自變數中未考量投資者交易情

緒，本研究使用加密貨幣恐懼與貪婪指數（Crypto Fear & Greed Index, CFGI）作為消息層面類自變數，加密貨幣恐懼與貪婪指數會根據不同來源的統計資料進行情感分析（Sentiment Analysis）與情緒分析（Emotion Analysis）將市場情緒簡化至0~100 計量單位，0 表示極度恐懼；100 表示極度貪婪，以下為加密貨幣恐懼與貪婪指數組成來源：

(一) 波動性 (Volatility, 25 %)：根據比特幣的價格回撤與其過去 30 天、90 天的相應平均值比較，如波動性上升/下降通常屬於市場恐慌/貪婪的跡象。

(二) 市場動量/成交量 (Market Momentum / Volume, 25%): 根據比特幣的交易量與其過去 30 天、90 天的相應平均值比較，如積極的巨額購買/拋售量表示市場貪婪/恐懼的跡象。

(三) 社群媒體 (Social Media, 15%)：主要為 X (前身 Twitter)、Reddit 收集與統計相關貼文，並檢查它們在特定時間範圍內的互動速度和次數，如異常高的互動率通常導致大眾對比特幣的興趣增長，屬於貪婪的市場行為。

(四) 調查 (Surveys, 15%)：於公共民意調查平台 Strawpoll 進行每週的加密貨幣民意調查，並詢問人們如何看待市場。

(五) 統治力 (Dominance, 10%)：比特幣的統治力 (Dominance) 是指比特幣在整個加密貨幣市場中所佔的市場份額。比特幣的統治力直接反映了投資者對比特幣相對於其他加密貨幣的偏好程度，且會受到多種因素的影響，包括市場情緒和投資者信心、新型加密貨幣項目、宏觀經濟因素等。總的來說，比特幣的統治力是一個動態的指標，會隨著市場和環境的變化而變化。

(六) 趨勢 (Trends, 10%)：Alternative 提取了各種與比特幣相關的搜尋查詢的 Google 趨勢數據，特別是搜尋量的變化以及推薦的其他當前流行搜尋。

綜合上述，本研究採用相關自變數名稱及擬採用之資料來源整理如表 4.1。

表 4.1 本研究重要變數與資料取得來源

資料名稱	資料來源
比特幣價格資訊：開盤價（Open Price）、收盤價（Close Price）、最高價（High Price）、最低價（Low Price）、交易量（Volume）	Binance
相對強弱指標（Relative Strength Index, RSI）	Binance
標準普爾 500 指數（Standard & Poor's 500, S&P500）	Yahoo Finance
加密貨幣恐懼與貪婪指數（Crypto Fear & Greed Index, CFGI）	Alternative

資料來源：本研究整理



## 第二節 訓練資料處理

本研究將數據集切割成訓練集（Training Set）、驗證集（Validation Set）、測試集（Testing Set），訓練集用於初步模型訓練，驗證集用於確認神經網路結構或控制模型複雜度的參數，測試集用於檢測最終模型之性能。

在建立模型之前，先對資料集進行標準化（Normalization）處理，以消除各項數據間的尺度效應。常見的標準化處理有離差標準化（Min-Max Normalization）與標準分數標準化（Z-Score normalization）。本研究採用離差標準化（Min-Max Normalization），如式(17)所示，此方法可線性改變原數據，並映射數據值於 0 至 1 之間。 $x_t$  為標準化前數據； $x_{norm}$  為標準化後數據； $x_{max}$  與  $x_{min}$  分別為數據中的最大值與最小值。

$$x_{norm} = \frac{x_t - x_{min}}{x_{max} - x_{min}} \quad (17)$$

以圖 4.1 為例，比特幣收盤價經過離差標準化（Min-Max Normalization）處理後，將資料範圍固定於 0 至 1 之間，可以避免與其他資料之間的差異導致某些特徵在模型訓練中過於顯著或不顯著的問題，更能提高模型收斂速度。

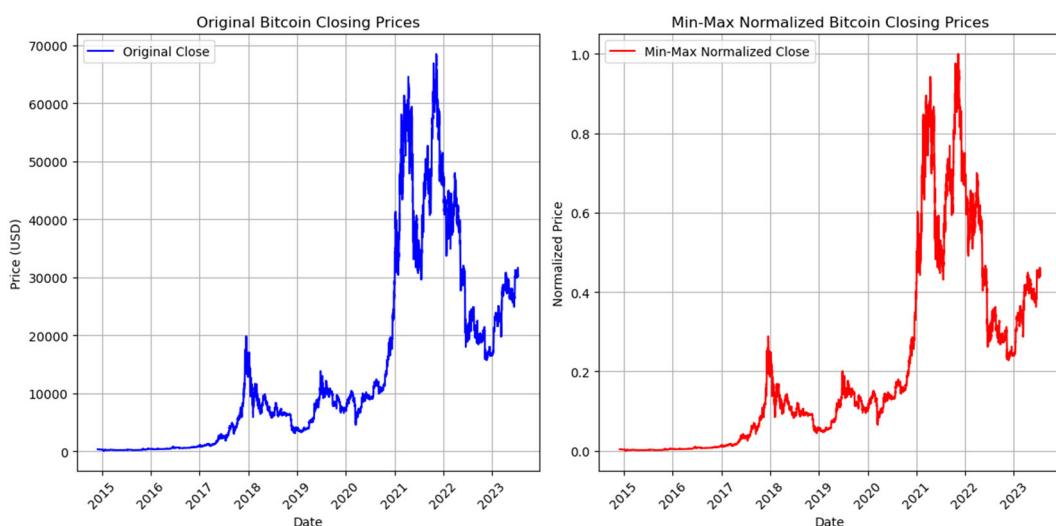


圖 4.1 標準化前後示意圖

資料來源：本研究整理

在建立模型之後，如式(18)所示，可將輸出結果還原標準化。 $y_t$ 為還原標準化後的模型輸出結果； $y_{norm}$ 為還原標準化前的模型輸出結果； $y_{max}$ 與 $y_{min}$ 分別為還原標準化前的模型輸出結果的最大值與最小值。

$$y_t = y_{norm}(y_{max} - y_{min}) + y_{min} \quad (18)$$

在模型訓練前，需要將標準化處理後的滑動窗口（Sliding Window）處理，意即使用前 N 個時間單位的資料預測下一個時間單位的資料。如圖 4.2 所示，以滑動窗口大小為 5 時所切割的訓練資料作為範例說明，由於使用前 5 個時間單位的資料預測下一個時間單位的收盤價，故第 N 組訓練資料的特徵值（X）為 1/1 00:00 至 1/1 16:00 的開盤價（Open Price）、收盤價（Close Price）、最高價（High Price）、最低價（Low Price）、交易量（Volume）、相對強弱指標（RSI）、標準普爾 500 指數（S&P500）、加密貨幣恐懼與貪婪指數（Crypto Fear & Greed Index, CFGI），目標值（Y）則是 1/1 20:00 的收盤價（Close Price）；第 N+1 組訓練資料的特徵值（X）與目標值則是相對於前一組向右滑動一個時間單位。

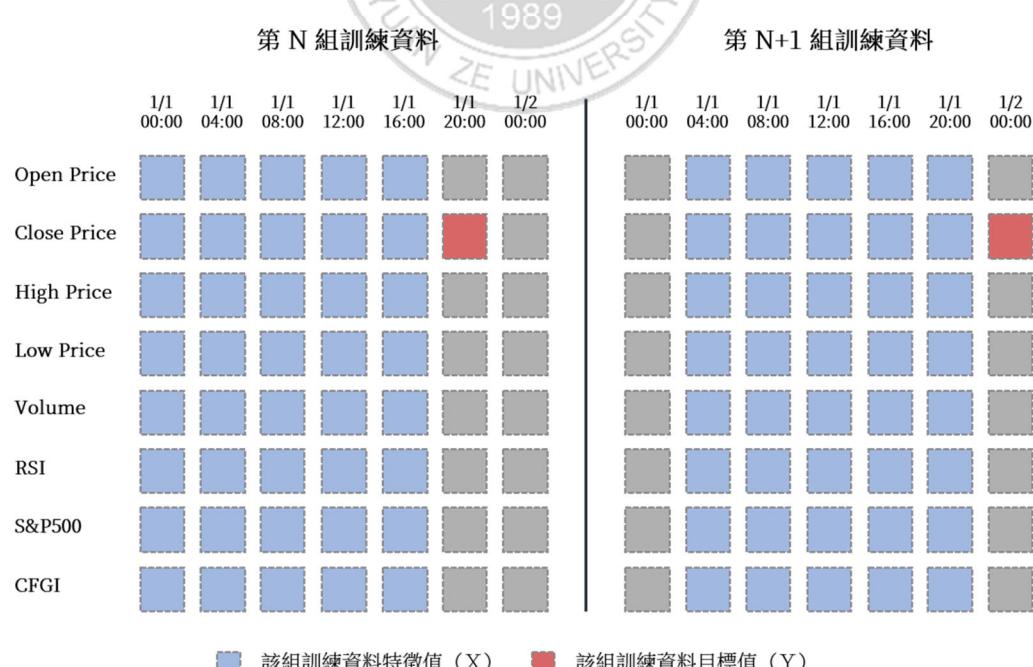


圖 4.2 滑動窗口切割示意圖

資料來源：本研究整理

考量到滑動窗口大小可能對於模型訓練結果的影響，本研究設計了三種大小的滑動窗口（5、10、15），將原始資料切割成3種不同時間尺度（Time Scale）的資料集，如表4.2所示，並以時間尺度（Time Scale）為5的訓練集（Training Set）為例說明，8352為資料的數量；5為各資料的時間尺度；8為各資料的特徵數。

表4.2 各時間尺度資料集

Time Scale	Training Set	Validation Set	Testing Set
5	(8352, 5, 8)	(1786, 5, 8)	(1787, 5, 8)
10	(8347, 10, 8)	(1781, 10, 8)	(1782, 10, 8)
15	(8342, 15, 8)	(1776, 15, 8)	(1777, 15, 8)

資料來源：本研究整理



### 第三節 建構比特幣價格預測模型

本研究建構 MRC-BiGRU 作為預測比特幣價格的模型，該混合模型於多尺度殘差卷積神經網路（Multi-scale Residual Convolutional Neural Network, MRC）將數據集中提取在不同時間序列尺度上具有的潛在特徵，並將所提取的特徵值送入後續連接的雙向門控循環單元網路（Bi-directional Gate Recurrent Unit, Bi-GRU）中，Bi-GRU 可對特徵值進行雙向分析，依時間序列學習關係模式來進行預測。並與其他混合模型 MRC-GRU、MRC-BiLSTM 及 MRC-LSTM，以研究各模型訓練於不同時間尺度資料的準確性及有效性。

如圖 4.3 所示，為本研究在時間尺度為 5 實驗中所建構 MRC-BiGRU 模型，並透過 Tensorflow 套件中的 plot\_model 函數結構視覺化圖片。

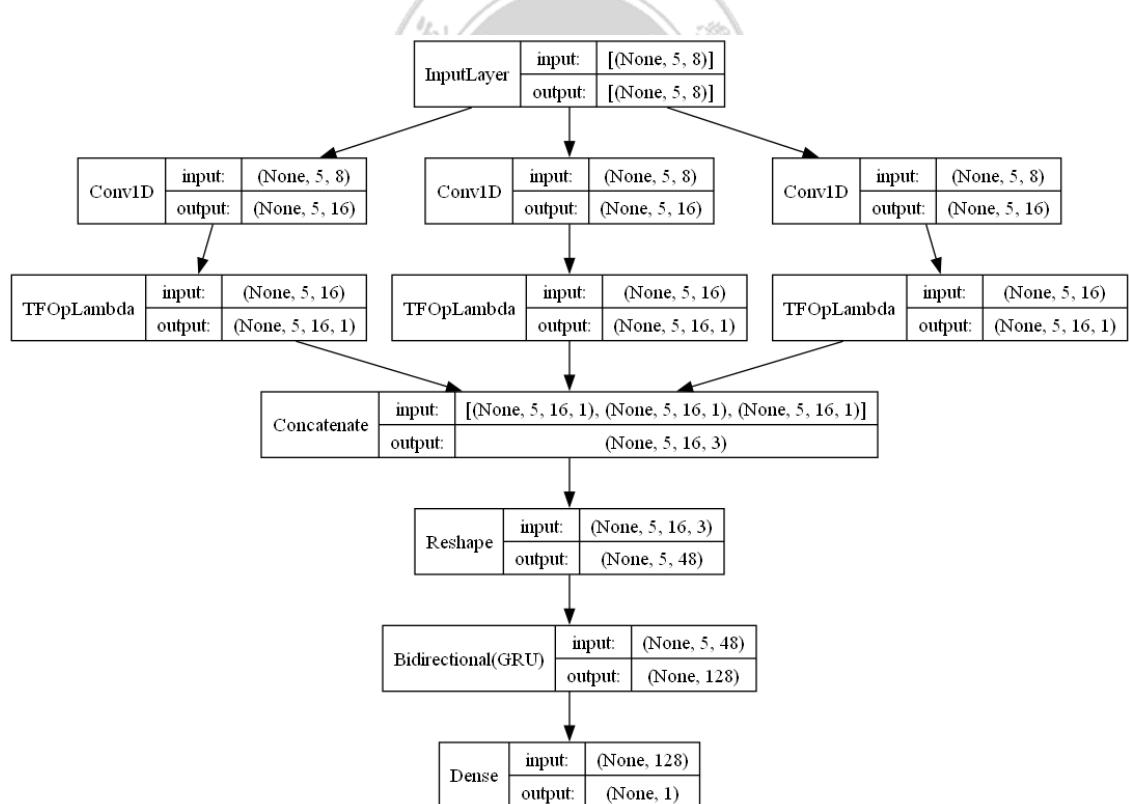


圖 4.3 MRC-BiGRU 模型視覺化

資料來源：本研究整理

本研究實驗環境可參考表 4.3，使用的程式語言為 Python 3.8.18、編程環境為 Jupyter、開源網路軟體庫為 Keras 2.8.0 與 Tensorflow 2.8.0。

表 4.3 實驗環境

<b>硬體設備</b>	CPU : Intel(R) Core(TM) i9-10980HK CPU @ 2.40GHz GPU : GeForce RTX 2070 Super with Max-Q Design
<b>程式語言</b>	Python 3.8.18
<b>編程環境</b>	Jupyter notebook 6.3.0
<b>開源網路軟體庫</b>	Keras 2.8.0、Tensorflow 2.8.0
資料來源：本研究整理	



## 第四節 預測模型評估指標

為了評估模型的性能，本研究選擇了四個評估指標，皆具備評估預測模型的解釋、預測能力，提供有價值的評估，並列式如下：

### (一) Mean Absolute Error (MAE)

MAE 是一種衡量模型預測值與實際值之間的平均絕對誤差的指標，如式(19)所示，以此來判斷預測模型的價格與實際價格之間的平均偏差，MAE 是一個直觀的指標，其取值範圍於  $[0, +\infty)$ ，越接近 0 代表模型的品質越良好。

$$MAE = \frac{1}{n} \sum_{t=1}^n |AP_t - PP_t| \quad (19)$$

### (二) Root-mean-square error (RMSE)

RMSE 是常用於衡量模型預測值與觀測值之間差異的一種指標，如式(20)所示，用於比較在特定數據集下所建立的不同模型的預測誤差，符合本研究的研究目的。RMSE 取值範圍於  $[0, +\infty)$ ，越接近 0 代表模型的品質越良好。

$$RMSE = \sqrt{\frac{\sum_{t=1}^n (AP_t - PP_t)^2}{n}} \quad (20)$$

### (三) Mean Absolute Percentage Error (MAPE)

MAPE 是絕對百分比誤差的平均值，如式(21)所示，能反映預測值誤差百分比的實際情況。MAPE 取值範圍於  $[0, +\infty)$ ，越接近 0% 代表模型的品質越良好。

$$MAPE = \frac{1}{n} \sum_{t=1}^n \left| \frac{AP_t - PP_t}{AP_t} \right| * 100\% \quad (21)$$

### (四) Coefficient of Determination ( $R^2$ )

$R^2$  在統計學中用於衡量「應變數的變異中可由自變量解釋部分」所占的百分比，如式(20)所示，以此來判斷預測模型的解釋能力，以殘差平方和 (Residual Sum

of Squares)、總平方和 (Total Sum of Squares) 組成。 $R^2$ 取值範圍於[0,1]，越接近1.0，代表此模式愈有解釋能力。

$$R^2 = 1 - \frac{\sum_{t=1}^n (AP_t - PP_t)^2}{\sum_{t=1}^n (AP_t - \bar{AP})^2} \quad (22)$$

$AP$ ( Actual Price )：比特幣實際價格； $AP_t$ ：在時間  $t$  的比特幣實際價格； $PP$ ( Predicted Price )：比特幣預測價格； $PP_t$ ：在時間  $t$  的比特幣預測價格。



## 第五節 預測模型評估

考量到滑動窗口大小對於模型訓練結果的影響，本研究設計了三種不同時間尺度大小的實驗資料，分別為 5、10、15，並下列將分別討論不同時間尺度資料於 MRC-LSTM、MRC-GRU、MRC-BiLSTM、MRC-BiGRU 模型訓練後的預測結果。

### (一) 時間尺度為 5：

表 4.4 預測模型的誤差

Architecture	MAE	RMSE	MAPE	R <sup>2</sup> (%)
MRC_LSTM	187.789	285.721	0.8352	99.640
MRC_GRU	179.09	261.589	0.8035	99.698
MRC_BiLSTM	166.085	258.213	0.7309	99.706
MRC_BiGRU	<b>159.864</b>	<b>252.725</b>	<b>0.7044</b>	<b>99.719</b>

資料來源：本研究整理

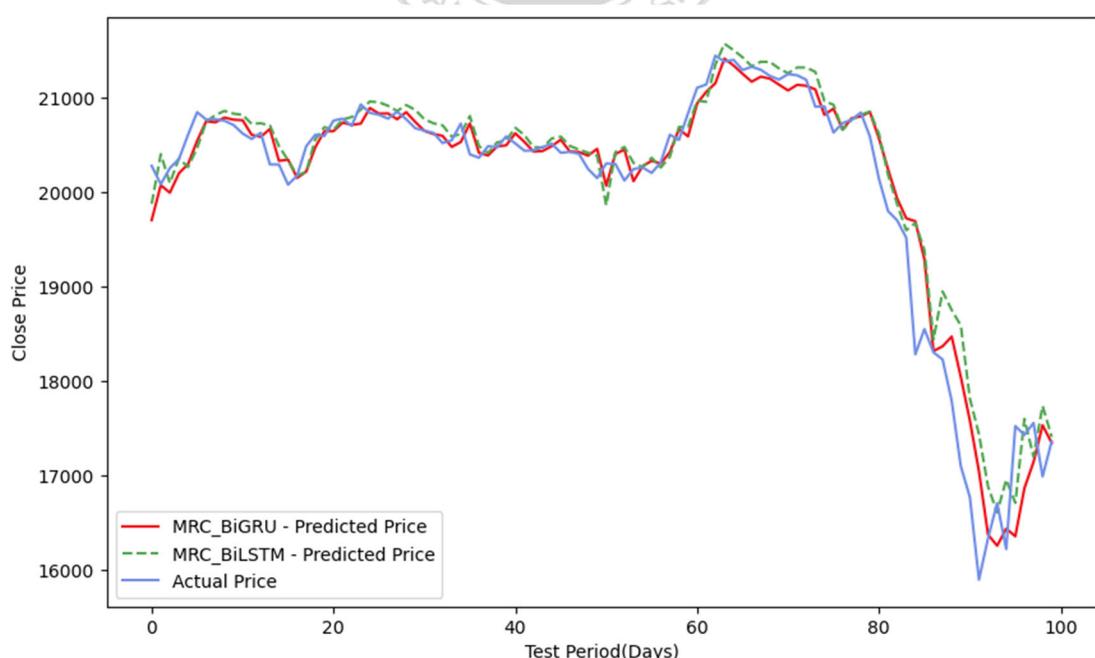


圖 4.4 MRC\_BiGRU 與 MRC\_BiLSTM 預測與實際價格結果比較

資料來源：本研究整理

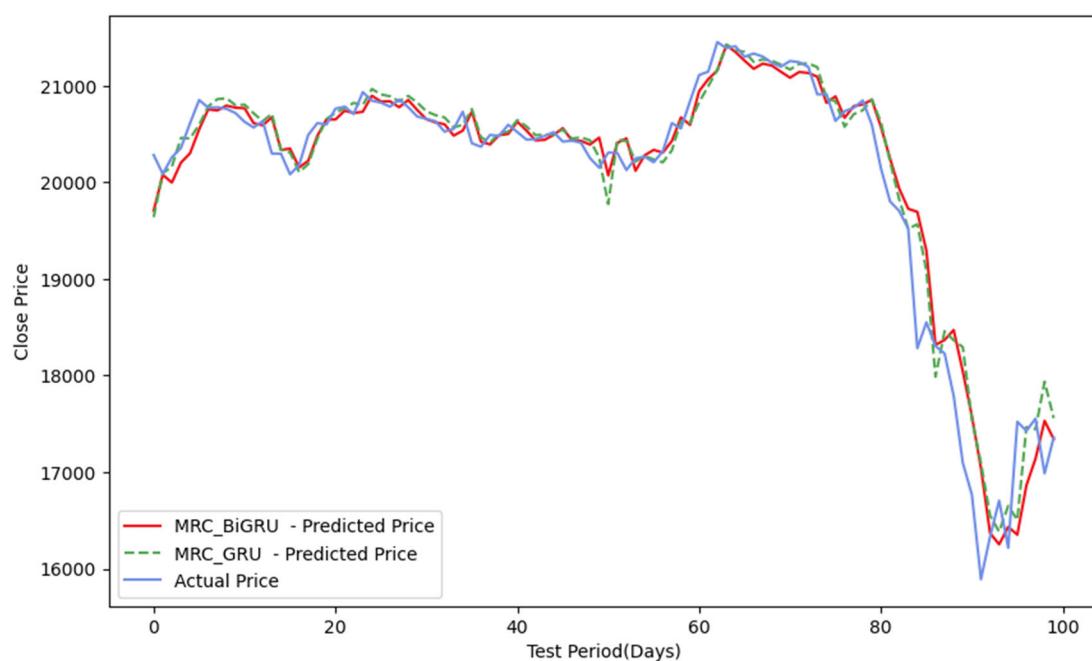


圖 4.5 MRC\_BiGRU 與 MRC\_GRU 預測與實際價格結果比較

資料來源：本研究整理

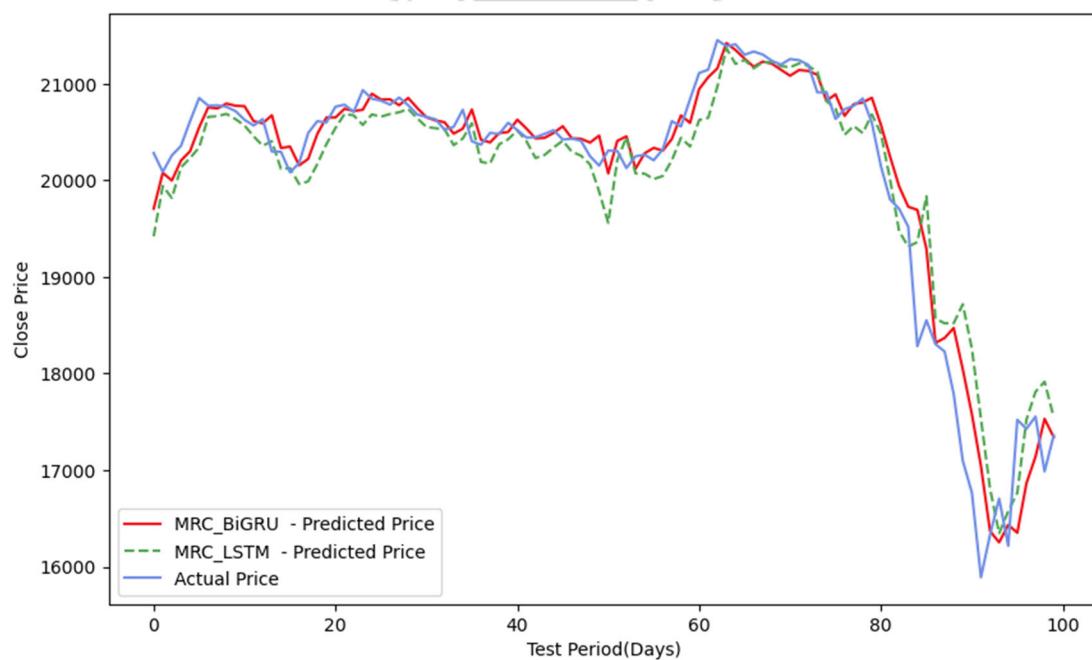


圖 4.6 MRC\_BiGRU 與 MRC\_LSTM 預測與實際價格結果比較

資料來源：本研究整理

對於時間尺度為 5 的實驗中，表 4.4 顯示了各模型訓練後，不同預測評估指標的準確性度量，從表 4.4 可以觀察出以雙向模型 MRC\_BiGRU 在各項評估指標中

皆呈現出最佳的預測效果，其次是雙向模型 MRC\_BiLSTM，最後是單向模型 MRC\_GRU 與單向模型 MRC\_LSTM。

## (二) 時間尺度為 10：

表 4.5 預測模型的誤差

Architecture	MAE	RMSE	MAPE	R <sup>2</sup> (%)
MRC_LSTM	185.821	276.694	0.8377	99.663
MRC_GRU	178.329	276.462	0.7863	99.663
MRC_BiLSTM	169.883	267.134	0.7360	99.686
<b>MRC_BiGRU</b>	<b>167.641</b>	<b>265.663</b>	<b>0.7310</b>	<b>99.689</b>

資料來源：本研究整理

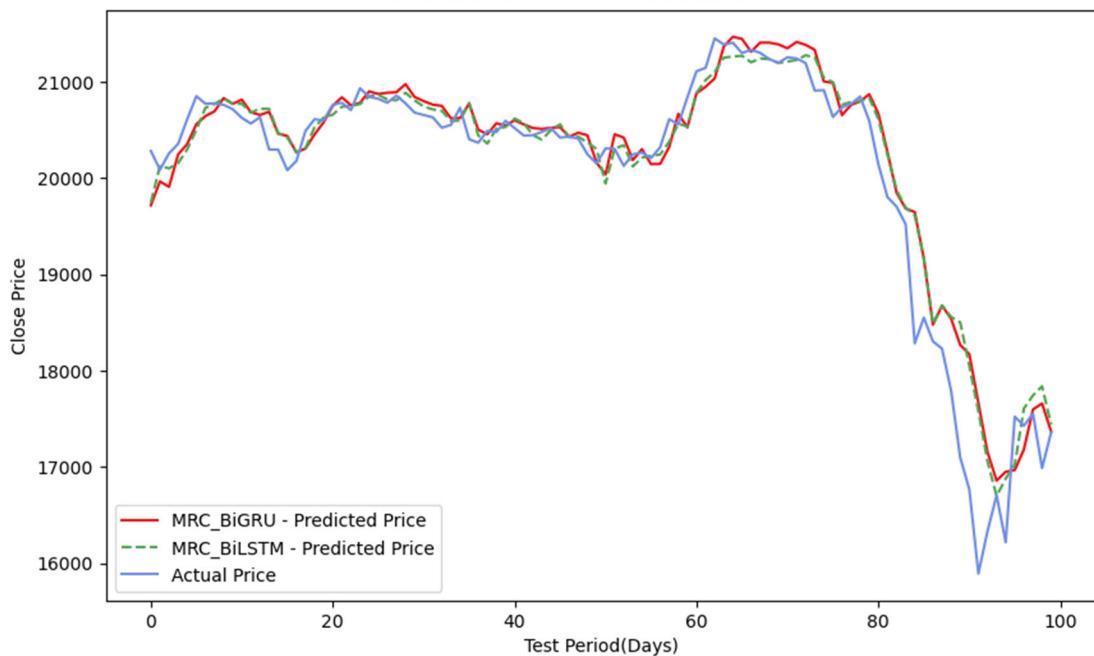


圖 4.7 MRC\_BiGRU 與 MRC\_BiLSTM 預測與實際價格結果比較

資料來源：本研究整理

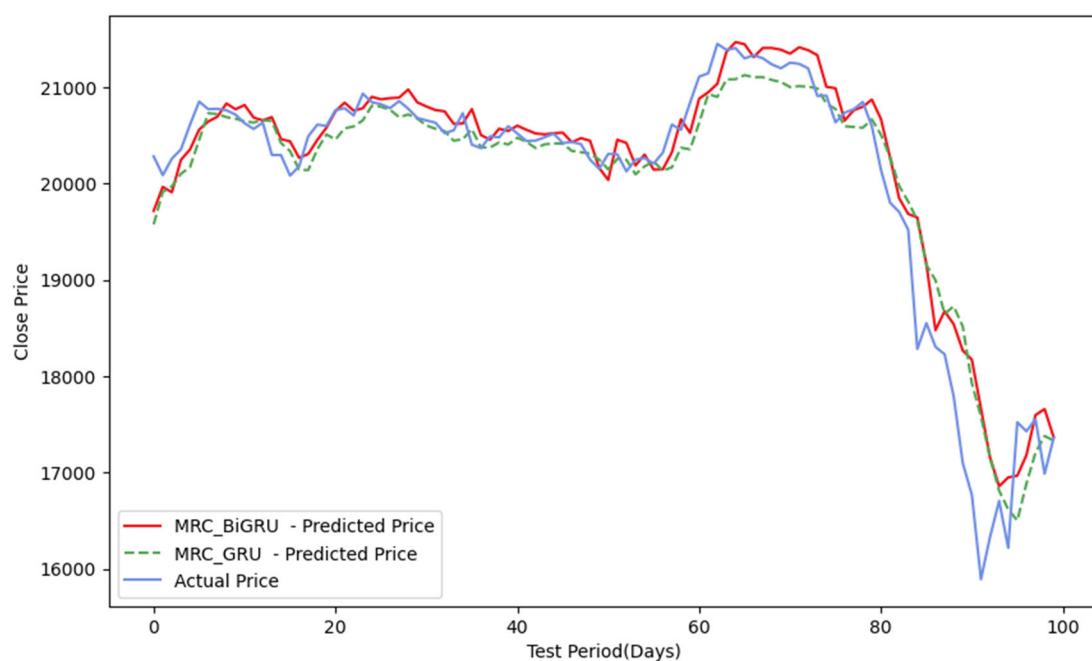


圖 4.8 MRC\_BiGRU 與 MRC\_GRU 預測與實際價格結果比較

資料來源：本研究整理

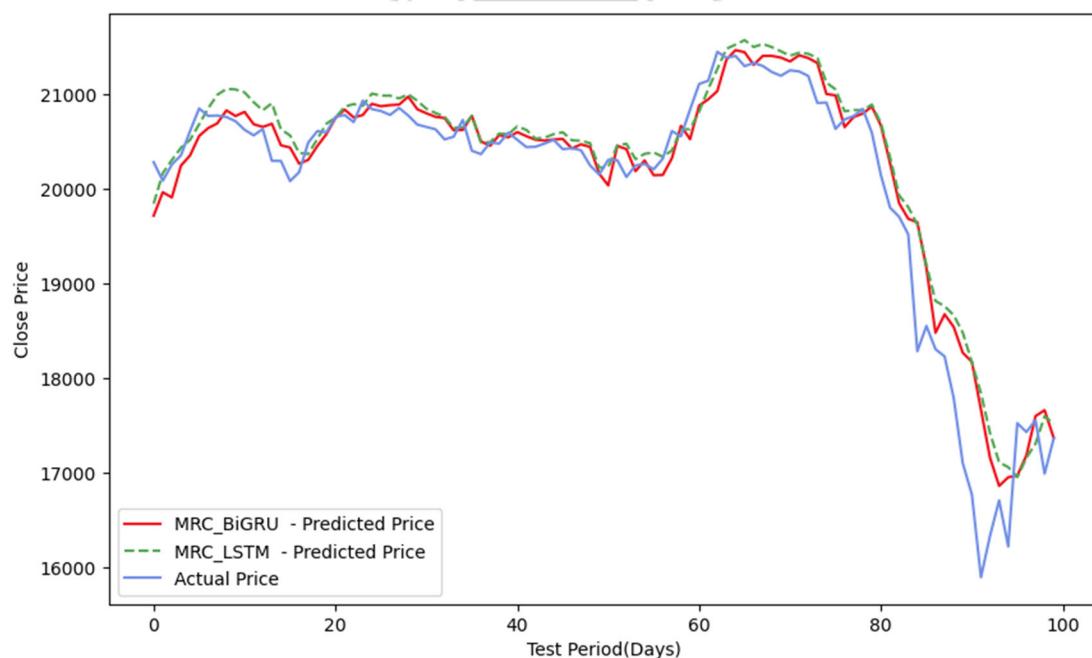


圖 4.9 MRC\_BiGRU 與 MRC\_LSTM 預測與實際價格結果比較

資料來源：本研究整理

對於時間尺度為 10 的實驗中，表 4.5 顯示了各模型訓練後，不同預測評估指標的準確性度量，從表 4.5 可以觀察出以雙向模型 MRC\_BiGRU 各評估指標中依

舊呈現最佳的預測表現，但相對時間尺度為 5 的實驗結果，與雙向模型 MRC\_BiLSTM 差距縮小；在單向模型的比較中，MRC\_GRU 依舊相對 MRC\_LSTM 取得較佳的預測成績，但相對時間尺度為 5 的實驗結果，兩者差距同樣縮小。

### (三) 時間尺度為 15：

表 4.6 預測模型的誤差

Architecture	MAE	RMSE	MAPE	R <sup>2</sup> (%)
MRC_LSTM	183.170	275.154	0.8120	99.667
MRC_GRU	191.795	285.654	0.8307	99.641
MRC_BiLSTM	<b>170.301</b>	<b>268.780</b>	<b>0.7432</b>	<b>99.682</b>
MRC_BiGRU	180.148	273.655	0.7999	99.670

資料來源：本研究整理

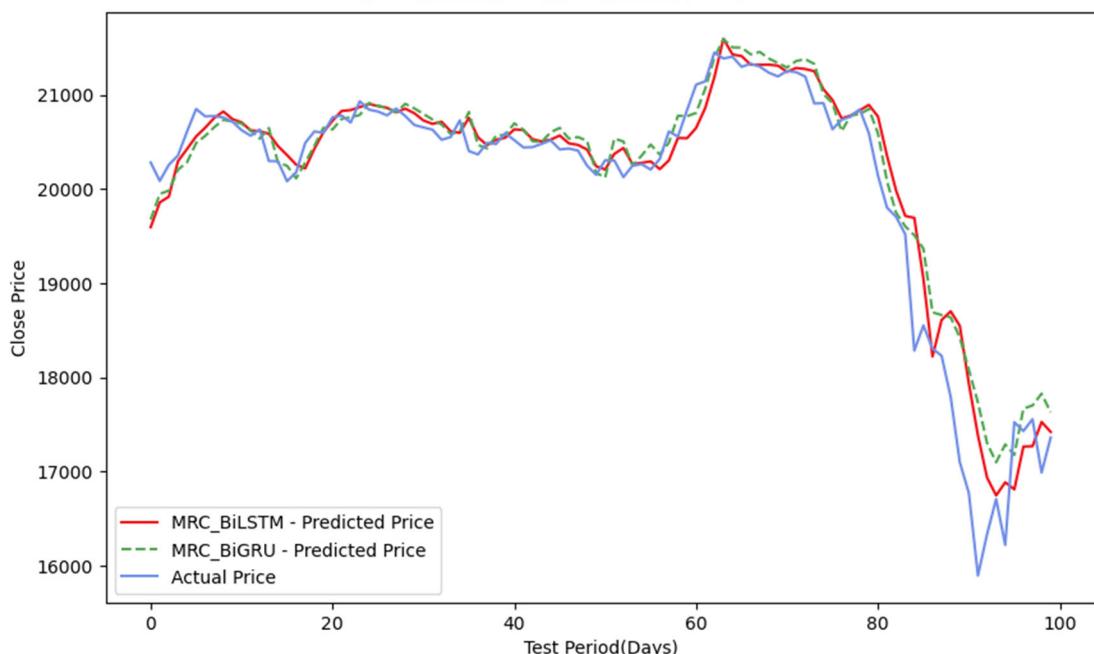


圖 4.10 MRC\_BiLSTM 與 MRC\_BiGRU 預測與實際價格結果比較

資料來源：本研究整理

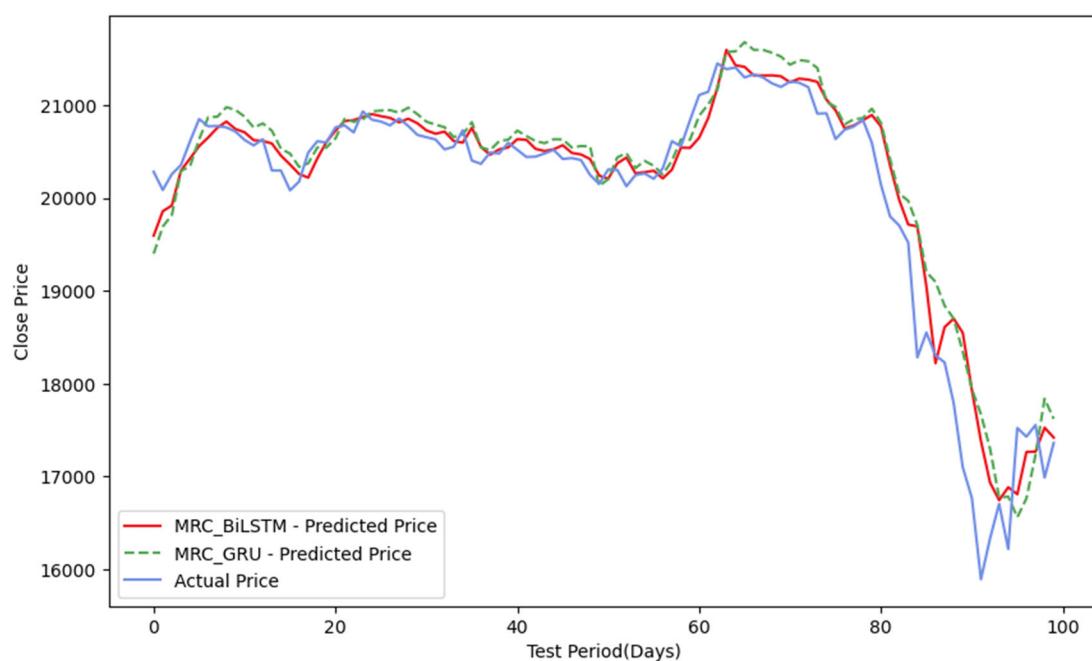


圖 4.11 MRC\_BiLSTM 與 MRC\_GRU 預測與實際價格結果比較

資料來源：本研究整理

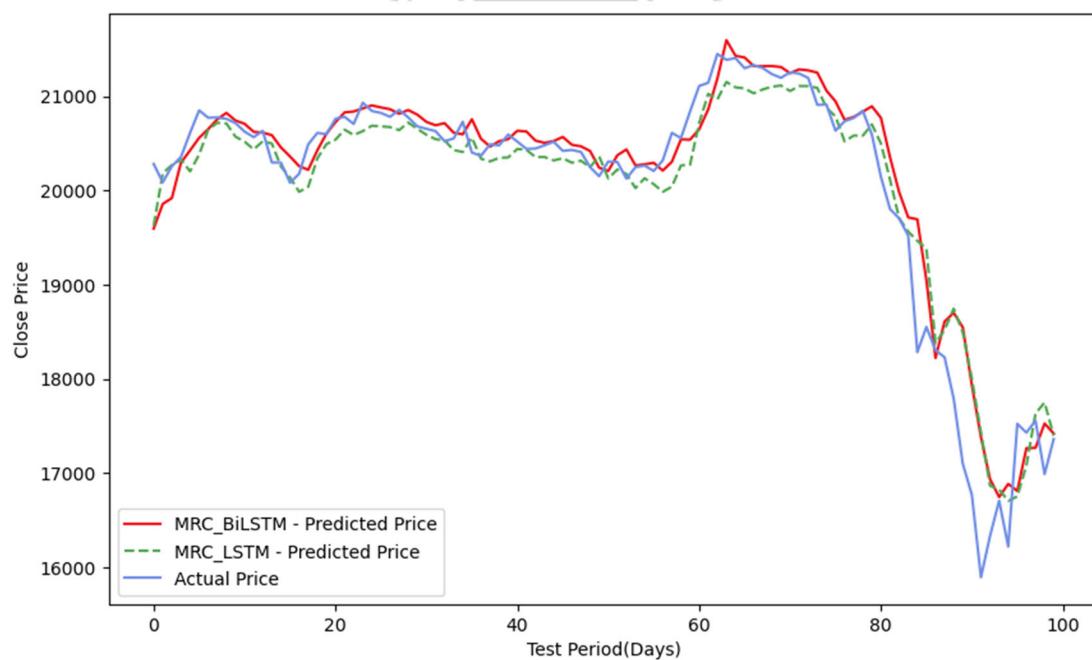


圖 4.12 MRC\_BiLSTM 與 MRC\_LSTM 預測與實際價格結果比較

資料來源：本研究整理

對於時間尺度為 15 的實驗中，表 4.6 顯示了各模型訓練後，不同預測評估指標的準確性度量，從表 4.6 可以觀察出以長短期記憶網路( Long Short-Term Memory,

LSTM) 系列的雙向模型 MRC\_BiLSTM 與單向模型 MRC\_LSTM 的預測結果在各評估指標中呈現較佳的預測表現，分別勝過雙向模型 MRC\_BiGRU 與單向模型 MRC\_GRU，這相對於前兩次實驗是不一樣的結果。



## 第五章 結論與建議

### 第一節 研究結論

對於比特幣價格的預測研究，學者起初嘗試使用各種不同單一模型如長短期記憶網路（LSTM）、門控循環單元網路（GRU）、卷積神經網路（CNN），以期找到最佳的預測模式，儘管單一模型之訓練結果優於傳統統計方法等，結果依舊有其侷限。後續學者致力於將模型以各種形式將混合來預測比特幣價格，其中以 Guo *et al.* (2021) 所提出的 MRC-LSTM 表現最佳，該方法特色在於其使用基於卷積神經網路（CNN）的多尺度殘差卷積神經網路（MRC）能夠檢測並融合多元時間序列中不同時間尺度的特徵，將特徵向量輸入到 LSTM 中來實現比特幣收盤價預測。

然而 MRC-LSTM 並未使用雙向模型分析反向資料。因此，本研究擬使用多尺度殘差卷積塊（MRC）於不同的時間序列提取特徵，並與具有雙向分析資料能力的雙向門控循環單元網路（Bi-directional Gate Recurrent Unit, Bi-GRU）結合依時間序列學習關係模式，形成全新的混合網路架構，即多尺度卷積雙向門控循環單元網路（MRC-BiGRU）來執行比特幣時間序列預測。考量到滑動窗口大小導致訓練資料時間尺度不同對於模型訓練結果的影響，本研究設計了三種不同時間尺度大小的實驗資料，分別測試不同時間尺度資料於 MRC-LSTM、MRC-GRU、MRC-BiLSTM、MRC-BiGRU 模型中訓練，並得其預測結果。

在訓練資料時間尺度為 5 的實驗中，以雙向模型 MRC\_BiGRU 在各項評估指標中皆呈現出最佳的預測效果，其次是雙向模型 MRC\_BiLSTM，最後是單向模型 MRC\_GRU 與單向模型 MRC\_LSTM。

在時間尺度為 10 的實驗中，雙向模型 MRC\_BiGRU 在各評估指標中依舊呈現最佳的預測表現，但相對時間尺度為 5 的實驗結果，與雙向模型 MRC\_BiLSTM 的

預測結果差距縮小；在單向模型的比較中，MRC\_GRU 依舊相對 MRC\_LSTM 取得較佳的預測成績，但相對時間尺度為 5 的實驗結果，兩者差距同樣縮小。

值得一提的是，在時間尺度為 15 的實驗中，長短期記憶網路（Long Short-Term Memory, LSTM）系列的雙向模型 MRC\_BiLSTM 與單向模型 MRC\_LSTM 的預測結果在各評估指標中呈現較佳的預測表現，分別勝過雙向模型 MRC\_BiGRU 與單向模型 MRC\_GRU，這相對於前兩次實驗是不一樣的結果。原因可能在於，LSTM 相對於 GRU 具有更多的門控結構，記憶單元複雜的結構能夠捕捉較長的時間依賴關係，因此在處理需要長期記憶的序列資料時可能會表現更好。



## 第二節 研究建議

在當今的學術研究中，對於金融商品的深度學習預測研究主要分為兩個方向，即價格預測研究和模擬交易預測研究。

價格預測研究專注於開發模型，以從歷史價格資料中學習模式，旨在預測價格的趨勢或未來走勢。準確預測金融商品未來的價格變化，這對投資者和交易者來說是至關重要的，因為它們能夠根據預測結果制定相應的交易策略。於本研究中，我們探討了多尺度殘差卷積神經網路(Multi-scale Residual Convolutional Neural Network, MRC) 在結合不同種類循環神經網路後，對比不同時間尺度輸入資料對比特幣收盤價格預測的影響。然而，本研究模型僅預測下一個時間單位的比特幣收盤價格，因此建議後續可以針對預測時間單位的擴充進行深入研究。

模擬交易預測研究則是更進一步，這類型的研究透過歷史數據回測、深度學習模型等方式制定交易策略，並提供了一個評估交易策略效果的實驗平台，讓交易者可以在不涉及真實資金的情況下測試和改進策略，進而提高交易的效率和收益。基於在本研究中觀察到的結果，MRC\_BiGRU 與 MRC\_BiLSTM 模型在不同時間尺度資料型態的比特幣價格預測結果各有千秋，因此建議後續可以針對不同需求，利用本研究模型建立模擬交易策略，並通過實驗和模型優化來確定哪種模型最適合特定的序列資料任務，幫助交易者測試和改進其策略，基於模型預測的結果生成交易信號，進而指導交易決策，並評估其在不同市場環境下的表現，提高交易的效率和收益。

## 參考文獻

- [1] Alkhodhairi, R.K., Aljalhami, S.R., Rusayni, N.K., Alshobaili, J.F., Al-Shargabi, A.A., and Alabdulatif, A. (2021). Bitcoin candlestick prediction with deep neural networks based on real time data. *Computers, Materials & Continua*, 68(3), 3215-3233.
- [2] Cavalli, S. and Amoretti, M. (2021). CNN-based multivariate data analysis for bitcoin trend prediction. *Applied Soft Computing*, 101(4), 107065.
- [3] Chang, G., Qiao, J., Liu, Y., Wang, Y., and Zhang, Z. (2021). A Perspective of LSTM Based Stock Prediction. *2021 International Wireless Communications and Mobile Computing (IWCMC)*, 576-580.
- [4] Chung, J., Gulcehre, C., Cho, K., and Bengio, Y. (2014), Empirical evaluation of gated recurrent neural networks on sequence modeling. *NIPS 2014 Deep Learning and Representation Learning Workshop*.
- [5] Das, S., Billah, M., and Mumu, S.A. (2021). A hybrid approach for predicting bitcoin price using Bi-LSTM and Bi-RNN based neural network. *Intelligent Data Engineering and Automated Learning -IDEAL 2021*, 223-233.
- [6] Dutta, A., Kumar, S., and Basu, M. (2020). A gated recurrent unit approach to bitcoin price prediction. *Journal of Risk Financial Management*, 13(2), 23.
- [7] Gao, Y., Wang, R., and Zhou, E. (2021). Stock prediction based on optimized LSTM and GRU models. *Scientific Programming*, 2021(4), 1-8.
- [8] Gruber, N. and Jockisch, A. (2020). Are GRU cells more specific and LSTM cells more sensitive in motive classification of text?. *Front Artif Intell*, 3, 40.

- [9] Guo, Q., Lei, S., Ye, Q., and Fang, Z. (2021). MRC-LSTM: A hybrid approach of multi-scale residual CNN and LSTM to predict bitcoin price. *2021 International Joint Conference on Neural Networks (IJCNN)*, 1-8.
- [10] He, K., Zhang, X., Ren, S., and Sun, J. (2016). Deep residual learning for image recognition. Computer Vision and Pattern Recognition, *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 770-778.
- [11] Hochreiter, S. and Schmidhuber, J. (1997). Long short-term memory. *Neural Computation*, 9(8), 1735-1780.
- [12] Huang, G., Liu, Z., Maaten, L.V.D., and Weinberger K.Q. (2017). Densely connected convolutional networks. *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2261-2269.
- [13] Kuo, R.J. and Xue, K.C. (1998). A decision support system for sales forecasting through fuzzy neural networks with asymmetric fuzzy weights. *Decision Support Systems*, 24(2), 105-126.
- [14] Li, Y. and Dai, W. (2020). Bitcoin price forecasting method based on CNN-LSTM hybrid neural network model. *The Journal of Engineering*, 2020(13), 344-347.
- [15] Liu, F., Li, Y., Li, B., Li, J., and Xie, H. (2021). Bitcoin Transaction Strategy Construction Based on Deep Reinforcement Learning. *2021 IEEE 2nd International Conference on Big Data, Artificial Intelligence and Internet of Things Engineering (ICBAIE)*, 180-183.
- [16] Maghyereh, A. and Abdoh, H. (2021). Time-frequency quantile dependence between Bitcoin and global equity markets. *The North American Journal of Economics and Finance*, 56(4), 101355.

- [17] Nithyakani, P., Tom, R.J., Gupta, P., Shanthini, A., John, V.M., and Sharma, V. (2021). Prediction of bitcoin price using Bi-LSTM network. *2021 International Conference on Computer Communication and Informatics (ICCCI)*, 1-5.
- [18] Rumelhart, D.E., Hinton, G.E., and Williams, R.J. (1986). Learning representations by back-propagating errors. *Nature*, 323, 533-536.
- [19] Saadatmand, F. and Chahoki, M.A.Z. (2023). Time Series Analysis by Bi-GRU for Forecasting Bitcoin Trends based on Sentiment Analysis. *2023 13th International Conference on Computer and Knowledge Engineering (ICCKE)*, 323-328.
- [20] Schuster, M. and Paliwal, K.K. (1997). Bidirectional recurrent neural networks. *IEEE Transactions on Signal Processing*, 45(11), 2673-2681.
- [21] Vassiliadis, S., Papadopoulos, P., Rangoussi, M., Konieczny, T., and Gralewski, J. (2017). Bitcoin value analysis based on cross-correlations. *Journal of Internet Banking and Commerce*, 22(S7), 1-12.
- [22] Xiao, D. and Su, J. (2022). Research on stock price time series prediction based on deep learning and autoregressive integrated moving average. *Scientific Programming*, 2022(2), 1-12.