



國立中央大學

企業管理學系

碩士論文

使用 MFCC 特徵提取技術結合 LSTM 模型之股價預測—以特斯拉股價為例

Tesla Stock Price Prediction through MFCC Feature Extraction
and LSTM Modeling

研究生：洪子鈞

指導教授：許秉瑜博士

共同指導：陳以錚博士

中華民國 113 年 6 月

中文摘要

過去十幾年間，全球經濟快速成長，促進了金融業的繁榮發展。隨著生活水平的提高，股票市場不僅是企業籌集資金的主要途徑之一，也成為大眾普遍採用的投資方式。然而，由於股票市場具有隨機和非線性的特性，使得股價預測變得非常具有挑戰性。

過往研究中，學者們已經探索了多種數值數據和文本數據的方法來預測股價。受此啟發，本研究提出了一種基於梅爾頻率倒譜係數（MFCC）的特徵提取方法，將股價歷史數據轉換為波形數據，並提取多維度的特徵向量，以提供一種新的股價預測方法，期望在準確性和可靠性方面取得顯著提升。

本研究透過將 MFCC 提取的特徵數據集切割為短期、中期和長期數據，以符合股價技術線圖中的周線、季線和月線趨勢，從而有助於模型在不同時間粒度下的訓練特徵，輸入 LSTM 模型與 SVM 模型進行訓練。最終，本研究的結果證明使用 MFCC 提取多維特徵數據結合 LSTM 模型，在股價預測領域具有一定的準確率，為未來的研究提供了一種新的特徵提取方法和方向。

關鍵詞：股價預測、梅爾頻率倒譜係數、長短記憶網路（LSTM）、支持向量機（SVM）

Abstract

In the past decade, the global economy has grown rapidly, promoting the prosperity of the financial industry. With the improvement of living standards, the stock market has not only become one of the main ways for companies to raise funds but also a common investment method for the public. However, due to the random and nonlinear nature of the stock market, stock price prediction has become very challenging.

In previous studies, scholars have explored various methods of numerical data and text data to predict stock prices. Inspired by this, this study proposes a feature extraction method based on Mel Frequency Cepstral Coefficients (MFCC), converting historical stock price data into waveform data and extracting multi-dimensional feature vectors to provide a new method of stock price prediction, aiming to achieve significant improvement in accuracy and reliability.

This study cuts the feature data set extracted by MFCC into short-term, medium-term, and long-term data to match the weekly, monthly, and quarterly trends in stock price technical charts, thereby assisting the model in training features at different time granularities, and inputs the LSTM model and SVM model for training. Ultimately, the results of this study demonstrate that using MFCC to extract multi-dimensional feature data combined with LSTM models has a certain accuracy in the field of stock price prediction, providing a new feature extraction method and direction for future research.

Keywords: Stock Price Prediction, Mel Frequency Cepstral Coefficients, Long Short-Term Memory (LSTM), Support Vector Machine (SVM).

誌謝

兩年的碩士生活迎來了尾聲，雖然忙碌但卻過得非常充實，我想感謝的人非常多，尤其是過程中曾給予我幫助的人，才使得我能擁有如此精彩的碩士生活。

首先，能夠順利完成碩士學位要特別感謝我的論文指導教授—許秉瑜老師，謝謝老師從碩一就開始督促我們讀論文，並耐心教導我們在 AI 領域相關的知識，在我研究論文卡關的時候，給我許多方向與建議，在 meeting 的時候也時常關心學生的課業及身心狀況。給予我們最大的支持與強大的後盾。

再來，謝謝在碩士班遇到的朋友們，在忙碌的課業之餘，能與朋友相處、聚餐，能讓我在壓力大的時候得以放鬆，特別感謝在寫論文期間陪伴我的朋友：室友兼鬧鐘：張嘉芸；AI 好夥伴：楊千郁；飯友：張嘉芸、楊凱崴、丁予思、李結衣；碰群：曾意軒、林靖璽，謝謝你們在碩二生活的陪伴。

最後，感謝我的家人，在我碩士生活中，給予我最大的支持與鼓勵，歲然能夠回家的日子不多，卻總能在每次回家時感受到溫暖，來到碩士班的最後一哩路，謝謝爸爸媽媽對我一路上的栽培。

洪子鈞 謹於

國立中央大學企業管理學系碩士班

中華民國 113 年 6 月

目 錄

摘要.....	2
Abstract.....	3
誌謝.....	4
第一章 緒論.....	1
1.1 研究背景.....	1
1.2 研究動機.....	1
1.3 研究目的.....	3
1.4 研究架構.....	4
第二章 文獻探討.....	5
2.1 梅爾倒頻譜係數 (MFCC).....	5
2.1.1 MFCC 技術應用於影像與聲音領域的識別任務.....	5
2.1.2 MFCC 特徵應用於時間序列的文獻.....	6
2.2 遞迴式神經網路 (Recurrent neural network, RNN).....	6
2.3 以深度學習預測股價之研究.....	8
第三章 研究方法.....	10
3.1 研究流程.....	10
3.2 資料前處理.....	10
3.2.1 數據轉換.....	10
3.2.2 特徵提取.....	11

3.3 研究模型：LSTM 模型架構.....	17
3.3.1 LSTM 層.....	18
3.3.2 時間分佈層（Time Distributed Layer）.....	20
3.4 損失函數（Loss Function）.....	21
第四章 實驗結果.....	22
4.1 資料蒐集.....	22
4.2 資料前處理.....	23
4.2.1 特徵生成.....	23
4.2.2 資料集切分.....	27
4.3 模型評估指標.....	29
4.4 實驗設定.....	30
4.4.1 LSTM 預測模型架構與超參數設置.....	30
4.4.2 支持向量機（Support Vector Machine, SVM）.....	34
4.5 實驗結果與分析.....	35
第五章 結論.....	38
5.1 研究結論.....	38
5.2 研究限制與未來研究建議.....	39
參考文獻.....	40



圖 次

圖 1	研究架構圖4
圖 2	RNN 示意圖7
圖 3	LSTM 示意圖8
圖 4	研究流程圖	11
圖 5	收盤價插值圖	11
圖 6	梅爾倒頻譜係數特徵萃取流程	12
圖 7	梅爾濾波器組	14
圖 8	LSTM 架構圖	17
圖 9	LSTM 單元結構圖	18
圖 10	收盤價差值圖	24
圖 11	音框步長 0.008 秒示意圖	24
圖 12	音框步長 0.0015 秒示意圖	25
圖 13	頻域圖	25
圖 14	每個音框的特徵維度	26
圖 15	LSTM 模型架構，實驗一	31
圖 16	LSTM 模型架構，實驗二	33

表 次

表 1	實驗參數表	26
表 2	資料集實驗參數表	28
表 3	實驗一的訓練集、驗證集與測試集	28
表 4	實驗二的訓練集、驗證集與測試集	29
表 5	混淆矩陣	29
表 6	SVM 超參數設定	34
表 7	實驗評估指標	36
表 8	混淆矩陣	37

第一章 緒論

1.1 研究背景

過去十幾年間，全球經濟迅速增長，帶動了金融業的蓬勃發展。同時，資本市場對經濟的影響力逐漸加強，股票市場的角色愈來愈重要。股票市場不僅是企業籌資的重要途徑之一，也隨著生活水準的提升，成為民眾普遍的投資方式。

投資股市需能預見股價走勢，股價預測這個跨領域研究因此變得非常有學術與實務價值。其核心技術在於通過分析歷史股價數據和相關財經資訊，預測未來股價的走勢。由於股票市場的隨機、非線性特性，使股價預測極具挑戰性。同時，股票市場的高風險和高回報特性，常使投資者希望能通過合理的投資策略來降低風險、增加收益。因此如何提高股價預測的準確性一直是學術界和業界關注的焦點。

1.2 研究動機

基於機器學習的研究

早期傳統的預測方法包括線性回歸（Sahoo and Charlapally, 2015）[1]、自回歸移動平均（ARMA）模型（Fu and Li, 2012）[2] 等，儘管這些技術可成功用於短期預測，但它們不適用於非線性問題，且長期預測性能較差（W. Lu, J. Li, J. Wang and L. Qin, 2021）[3]。為了解決這個問題，越來越多的研究者開始引入機器學習來分析時間序列，且成功應用於股價預測領域，如支持向量機（SVM）、隨機森林（RF）等，能夠捕捉數據中的複雜模式，顯著提高預測精度（H. Xu, B. Xu and K. Xu, 2020）[4]。

基於深度學習的研究

由於近年來深度學習技術快速發展，許多神經網路架構被用來建模於各種具有非線性時間序列的模式。因此，深度學習方法在股價預測領域的應用逐漸成為學者們的研究重點，如門控循環單元（GRU）、循環神經網絡（RNN）、卷積神經網絡（CNN）、長短期記憶（LSTM）和雙向長短期記憶（BiLSTM）。S. Mehtab, J. Sen 和 A. Dutta,(2020)[5]，使用 LSTM 回歸模型預測印度 NIFTY 50 指數，結果顯示，基於深度學習的 LSTM 模型表現優於傳統機器學習方法，對於處理時間序列數據方面表現出色，能夠更好地應對股價數據的動態特性和長期依賴關係。

股價預測的特徵數據

過往用於股票市場預測的數據類型可以分為兩種類型：基於數值數據和基於數值和文本數據，歷史市場數據如價格與交易量，通過技術指標來預測股價，例如，Nazario et al. (2017)[6] 使用技術指標如 EMA、RSI、隨機指標和變動率來預測各種股市的走勢；Picasso et al. (2019) [7] 使用 10 個常見技術指標生成買賣信號，進行 NASDAQ 100 指數中 20 家公司股票的交易策略。而文本數據則是來自社會政治事件和金融新聞的文本數據，通過對股票相關新聞、財經和其他文本信息的情緒進行分類，可以幫助用戶對股票的正面和負面影響做出投資判斷（Gao et al., 2010）[8]，例如，Yu et al. (2019) [9] 提出了一種使用 K-means 聚類的新文本融合模型，使用 K-means 聚類方法生成文本類別，並使用多層感知器算法結合 15 個原始價格指標（如開盤價和收盤價）來預測市場價格。W. Long, Z. Lu, 和 L. Cui 學者 (2019) [10] 提出了一種具有卷積和遞歸神經元層的深度學習模型，該模型使用價格和交易量的歷史數據作為輸入，將價格趨勢分為三類（如上漲、持平和下跌），準確率可達 55.5%。Cui et al. (2023)[11] 提出一種新的股價預測的特徵，使用小波變換（WT）對原始數據進行分解，求取特徵，將此特徵輸入 LSTM 模型進行預

測。小波轉換在高頻時段會取較少時間的資料而在低頻時段取較多的時間資料進行轉換，所獲取特徵會有不同時間尺度的問題。再將此變換時間尺度的資料輸入 LSTM, 使其預測結果較難解釋。

受此啟發，本研究提出了一種基於梅爾頻率倒譜係數（MFCC）的特徵提取方法來預測股價。與小波變換不同，MFCC 取固定時段資料萃取特徵值，使得股價的預測更符合直覺。

1.3 研究目的

本研究基於上述股價預測研究，旨在提出一種新的特徵數據應用於股價預測。將原始股價數據轉換為波形數據，並使用 MFCC 技術提取特徵向量，這些特徵向量能夠反映股價數據的頻譜特性，進而提高預測模型的準確性。具體來說，本研究提出以下創新點：

1. 數據預處理：將股價歷史數據轉換為波形數據，準備進行 MFCC 特徵提取。
2. 特徵提取：使用 MFCC 技術從波形數據中提取反映股價數據頻譜特徵的特徵向量。
3. 模型構建：構建一個基於 LSTM 的深度學習模型，使用提取的 MFCC 特徵進行股價預測分類任務。
4. 模型驗證：以 SVM（支持向量機）與 LSTM 的深度學習模型來衡量模型之表現，並使用準確率 (Accuracy)、精確度 (Precision)、召回率 (Recall)、F1-score 做為主要績效指標，得出的結果驗證本方法的有效性和準確性。

通過這些創新點，本研究將為股價預測提供一種新的特徵提取方法，並期望在準確性和可靠性方面取得顯著的提升。

1.4 研究架構

本研究分爲五個章節，第一章說明本研究的背景、研究動機、研究目的與研究架構；第二章爲文獻探討，描述應用 MFCC 技術於提取特徵數據與及介紹過往演算法應用於股價預測所用之模型；第三章爲研究方法，闡述本研究之研究數據轉換、MFCC 特徵提取和數據分割方法；第四部分描述了實驗結果和模型評估；第五部分總結了貢獻並提出了進一步的研究方向。

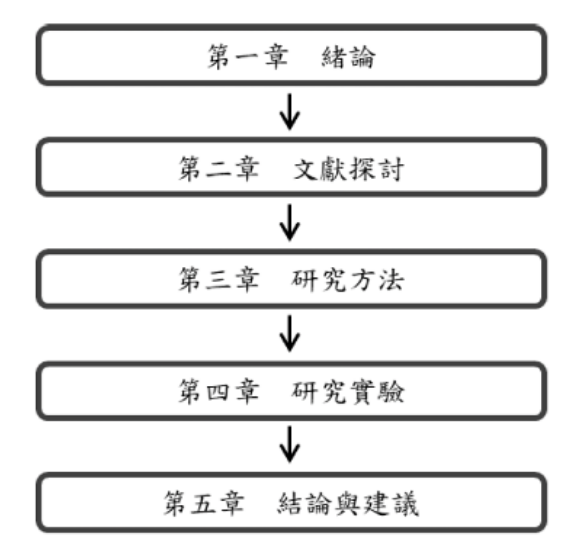


圖 1: 研究架構圖

第二章 文獻探討

此章節闡述本研究之文獻，第一部分將探討 MFCC 特徵提取的運作原理以及應用領域；第二部分將介紹本研究所使用之遞迴式神經網絡；第三部分將介紹 LSTM 預測股價之研究。

2.1 梅爾倒頻譜係數 (MFCC)

梅爾倒頻譜係數 (Mel Frequency Cepstral Coefficients) 是一種在語音訊號處理和語音識別中廣泛使用的特徵提取方法。它由 Davis 和 Mermelstein (1980) [12] 所提出，MFCC 的基於人耳對聲音頻率的感知特性，將語音信號轉為頻譜圖後，將頻譜圖畫分為對數間隔，模擬人類聽覺系統對頻率差異的感知。MFCC 提取過程包括將音頻信號分割成若干時間窗口，對每個窗口應用傅立葉變換以獲得頻譜，然後通過梅爾濾波器組進行過濾，將頻譜轉換為梅爾頻率域，接著進行對數變換，最後通過離散餘弦變換 (Discrete Cosine Transform, DCT) 得到梅爾頻率倒譜係數。這一過程有效地壓縮了數據，提取了反映聲音特徵的主要成分，並且大幅降低了計算複雜度，使其在各種聲音識別任務中表現出色。

2.1.1 MFCC 技術應用於影像與聲音領域的識別任務

MFCC 技術雖然最初設計用於語音識別，但其靈活性和優越的特徵提取能力使其在其他領域的識別任務中也獲得了廣泛應用。例如，V. Bansal 等人 (2020) [13] 將 MFCC 應用於咳嗽分類，基於 MFCC 特徵和頻譜圖影像作為 CNN 網路的輸入。Yusuf 和 Hidayat (2019) [14] 則利用 MFCC 從心電圖波形中提取特徵，並結合 KNN 分類器和 k 折交叉驗證，成功構建了心血管疾病診斷模型。J. Dutta 和 D. Chanda (2021)[15] 使用 MFCC 作為識別音樂情緒的特徵，可以達到 93.75% 的準確

率。T. Maeda 和 M. Yajima (2017) [16] 使用高速攝像機記錄排球比賽的運動影像數據，並通過快速傅里葉變換 (FFT) 和梅爾頻率倒譜系數 (MFCC) 進行分析。

2.1.2 MFCC 特徵應用於時間序列的文獻

在時間序列數據的應用中，MFCC 特徵同樣顯示出其強大的適應性和靈活性。Othman 等人 (2009)[17] 從 EEG 信號中提取了 MFCC 特徵，用於識別和分類患者的情感（如快樂、恐懼、悲傷和冷靜），並將這些特徵輸入多層感知器（MLP），準確率可達 90%。Yusuf 和 Hidayat (2019)[18] 評估了 13 個 MFCC 特徵和離散小波變換（DWT）的性能於心電圖（ECG）分析，結果顯示 13 個 MFCC 的性能優於 DWT。Mingsi 等人 (2021)[19] 提出了自適應頻率倒譜係數（AFCC）方法，並將其輸入 XGBoost 算法來分類軸承故障。Wang 等人（2022）通過時間序列數據提取 MFCC 特徵並預測銷售變化，將其輸入多層感知器（MLP），準確率可達 57.1%。MFCC 能夠根據需求提取多種頻譜特徵，從而在時間序列分析中提供更多的數據維度和信息，以上文獻顯示了 MFCC 特徵在處理時間序列數據中的廣泛應用和有效性，MFCC 均能提供高準確率和可靠性，並展示了其在實際應用中的成功案例。

基於這些前人的研究成果，本研究認為 MFCC 提取特徵技術在跨領域與時間序列數據的潛力和有效性，可以有效捕捉股價變動中的關鍵模式和趨勢，從而提高股價預測模型的性能和準確性，為股價預測提供一種新的數據集。

2.2 遞迴式神經網路 (Recurrent neural network, RNN)

遞迴式神經網路 (Recurrent neural network, RNN) 是時間序列中常見的模型之一，有別於傳統的神經網絡模型，RNN 接收輸入序列的每個元素，並將前一次輸出的結果，作為下一個隱藏層的輸入，使其能夠保留前面輸入的記憶，與當前的

輸入一起學習。

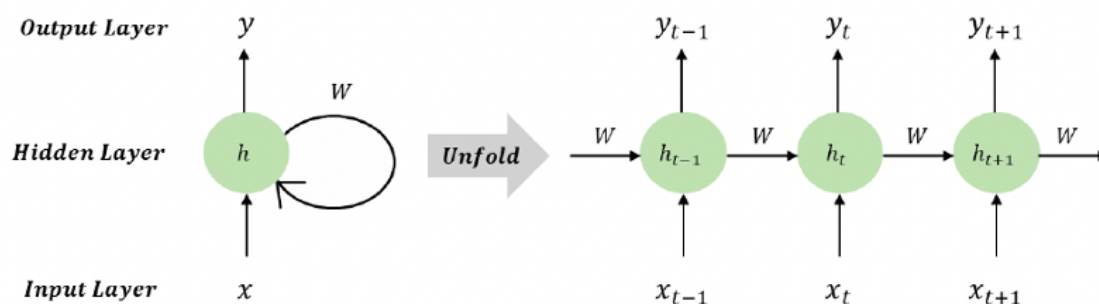


圖 2: RNN 示意圖

由於，在長序列數據的訓練中，RNN 會遇到梯度消失或梯度爆炸問題，導致訓練困難。梯度消失是指隨著神經網路層數的增加，前面層的資訊在向後傳遞時呈指數遞減，導致這些層的權重逐漸變得越來越小，使得網絡無法有效地記住和利用早期輸入的訊息。因此，Hochreiter 和 Schmidhuber (1997) [20] 於 1997 年提出長短期記憶 (Long Short-Term Memory, LSTM) 來改善 RNN 在長時間序列上的不足。

LSTM 的結構可以區分為長期記憶和短期記憶，如圖 3 所示。短期記憶主要為原 RNN 中的狀態向量，通過將當前輸入與前一時間點的狀態結合來獲取當前的狀態，這種短期記憶狀態稱為隱藏狀態 (hidden state)。長期記憶則是通過單元狀態 (cell state) 來儲存過去資訊，這個過程結合了當前輸入、前一時間點隱藏狀態以及前一時間點的單元狀態。

LSTM 引入了單元狀態，並且新增了三個控制閥門，分別是遺忘門 (forget gate)、輸入門 (input gate) 以及輸出門 (output gate)，用來決定記憶的儲存和使用。其中：

- 遺忘門：控制上一時間點的單元狀態中哪些資訊應該被保留到當前的單元狀態，哪些資訊應該被遺忘。

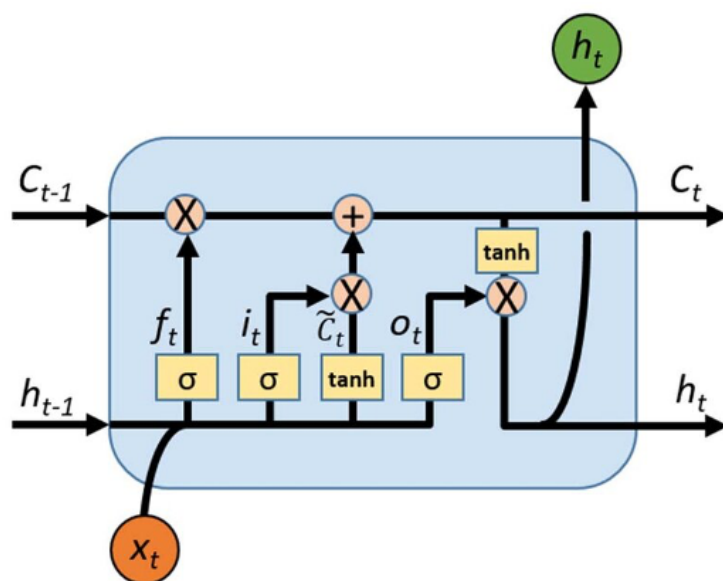


圖 3: LSTM 示意圖

- 輸入門：控制當前輸入中的哪些新資訊可以加入到當前的單元狀態中。
- 輸出門：決定當前的單元狀態中哪些資訊應該作為當前時間點的隱藏狀態並輸出。

這些門控機制使得 LSTM 能夠有效地管理和使用長期和短期記憶，從而在處理長序列數據時表現出色。

2.3 以深度學習預測股價之研究

因 LSTM 成功應用於處理時間序列的資料，許多學者開始研究 LSTM 在股價預測的應用。W. Long, Z. Lu, 和 L. Cui 學者（2019）[10] 提出了一種具有卷積和遞歸神經元層的深度學習模型，該模型使用價格和交易量的歷史數據作為輸入，將未來資產價格趨勢分為三類，準確率可達 55.5%。C. K. Vignesh 學者（2020）[21] 收集了 2011 年 1 月至 2015 年 12 月的 Yahoo 和 Microsoft 資產的開盤價、收盤價、最低價和最高價，並計算了五個技術指標作為特徵：動量、波動率、指數動量、指數波動率、股票動量、股票價格波動率。該論文進一步比較了 SVM 和

LSTM 模型在股票趨勢二元分類問題上的表現，結果顯示 LSTM 算法的準確性最高。M. Nabipour 等學者（2020）[22] 從 2009 年 11 月到 2019 年 11 月的資產開盤價、收盤價、最低價和最高價計算了十個技術指標，並應用決策樹、袋裝算法、隨機森林、Adaboost、梯度提升、XGBoost、人工神經網絡（ANN）、遞歸神經網絡（RNN）和長短期記憶（LSTM）算法來預測價格，結果顯示 LSTM 算法表現優於其他算法。

基於以上研究發現在股價預測任務中選擇使用 LSTM 模型可以有效的捕捉到股價波動中的時間序列特徵，並得到良好的預測結果，因此本研究採用基於 LSTM 的模型來分析股價預測的分類任務。

第三章 研究方法

此章節闡述本研究之方法，分為以下四節，第一節為研究流程；第二節為資料前處理與特徵提取之方法；第三節為時間序列模型建構之方法；最後第四節將說明模型使用之損失函數。

3.1 研究流程

本研究之流程分為四個部分，主要步驟如圖4所示。首先，進行資料蒐集 (Data Collection)，收集歷史股價中的每日收盤價和每日交易量。接著進行資料前處理 (Data Preprocessing)，包括數據轉換、特徵提取和切分資料集。在數據轉換步驟中，將原始數據轉換為波形數據的格式，以便在特徵提取步驟中，使用 MFCC 技術從數據中提取有用特徵；在切分資料集步驟中，將數據切分為三組數據集，分別為五日趨勢特徵、二十日趨勢特徵和六十日趨勢特徵。這些數據集再進一步分為訓練集、驗證集和測試集，以便於模型訓練和評估。第三部分是模型訓練 (Model Training)，建立 LSTM 模型，利用前處理後的數據集進行訓練，以學習股價變動的規律。最後是模型評估 (Model Evaluation)，使用多種評估指標，如準確率 (Accuracy)、精確度 (Precision)、召回率 (Recall) 和 F1-score，對訓練好的模型進行評估，以確定模型的性能和預測能力。

3.2 資料前處理

3.2.1 數據轉換

使用股票市場歷史股價中的每日收盤價和每日交易量作 原始數據，分別透過 MATLAB 進行線性插值法，如圖5所示：

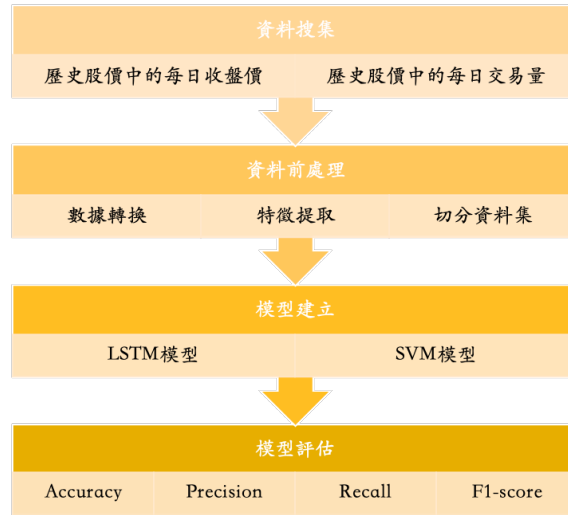


圖 4: 研究流程圖

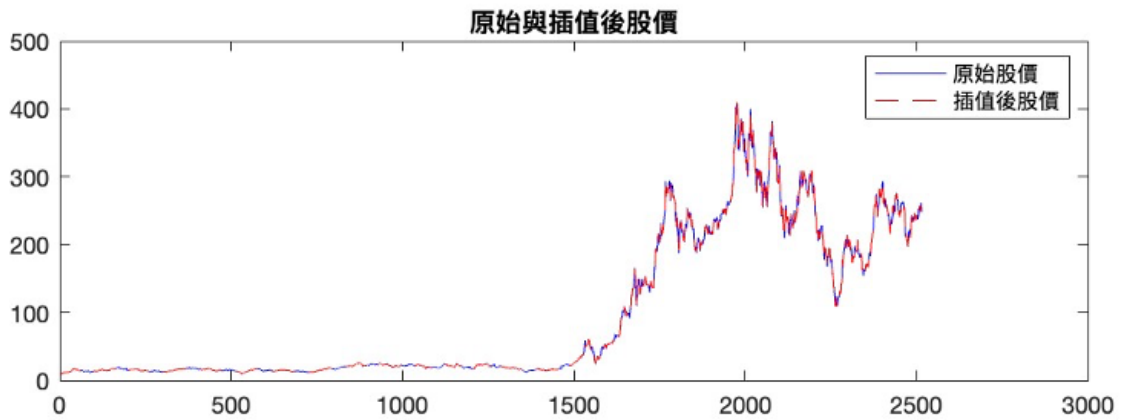


圖 5: 收盤價插值圖

而後將插值後的數據進行標準化，縮放於 $[-1,1]$ 之間，確保音頻信號的振幅在處理和分析時保持在一個一致的範圍內，以便轉換成波形數據。如公式（1）所示：

$$\text{normalizedAudioSignal} = 2 \left(\frac{Y - \min(Y)}{\max(Y) - \min(Y)} \right) - 1 \quad (1)$$

3.2.2 特徵提取

本研究使用梅爾倒頻譜係數（Mel-Frequency Cepstral Coefficients MFCC）作為特徵提取的技術，來提取波形數據集的特徵。一般提取語音訊號的 MFCC 特徵主

要有七個步驟，分別為的預強調、音框化、漢明窗、快速傅立葉轉換、三角帶濾波器、對數轉換以及離散餘弦轉換，即可將語音訊號轉換成 MFCC 特徵向量。如圖6所示：

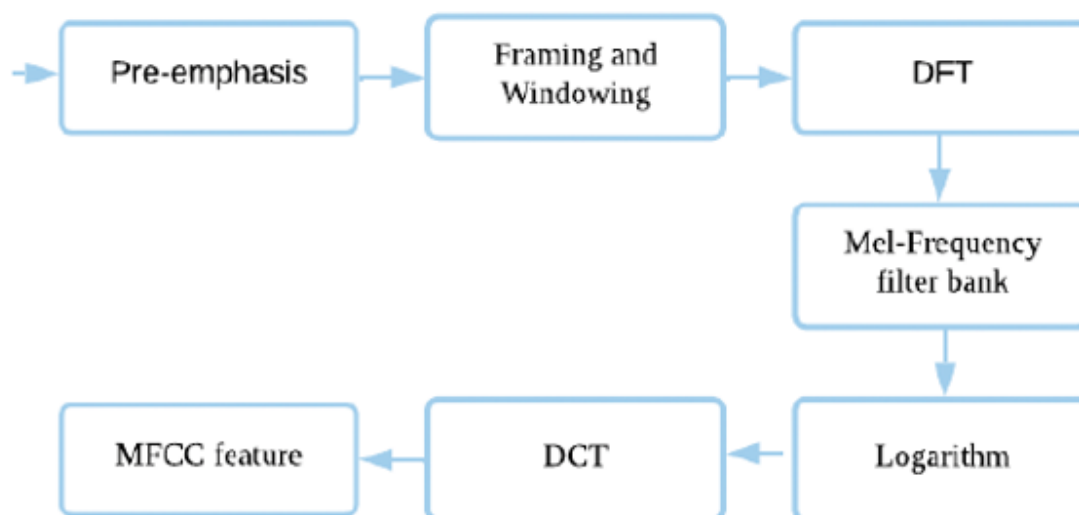


圖 6: 梅爾倒頻譜係數特徵萃取流程

然而，預加重主要應用於語音信號處理，用以補償高頻成分的能量衰減。但由於本研究股價數據不同於語音信號，數據的能量主要集中在低頻範圍，幾乎沒有顯著的高頻成分。因此，預加重對於股價數據的處理並沒有顯著的必要性，反而可能會導致不必要的失真。本研究選擇省略預加重步驟，保持數據的原始形態，以確保特徵提取的準確性和有效性。以下將詳細說明本研究特徵提取的步驟：

1. 音框化（Framing）：將連續的聲音信號分割成多個短時距音框。因聲音信號是非平穩信號，透過將其分成短時距音框，使信號在每個音框內會相對平穩，通過設定窗口長度和重疊時間，可以捕捉到聲音信號的時間和頻率變化特徵，從而達到時間和頻率分辨率的平衡。音框數量計算公式2如下：

$$\text{Number of frames} = \frac{\text{signal length (s)} - \text{frame length (s)}}{\text{frame step (s)}} + 1 \quad (2)$$

其中，signal length 為音檔總時長，本研究以 0.008 秒作為音框長度（frame

length)，並設置兩種重疊時間（frame step）來進行實驗。

2. 漢明窗（Hamming Window）：是對音框進行加窗處理的一種方法。為了確保相鄰音框的左右兩端信號連續，需要在音框化時使用加窗函數，即對每一個音框乘以一個漢明窗。其主要作用是消除相鄰音框之間的不連續性，在音框的兩端逐漸減少信號的幅度，從而達到平滑效果，避免在信號處理過程中出現頻譜泄露和諧波失真等問題。如公式3,4所示：

$$\tilde{s}[n] = s'[n]d[n] \quad (3)$$

$$d[n] = \begin{cases} 0.54 - 0.46 \cdot \cos\left(\frac{2\pi n}{N}\right), & 1 \leq n \leq N \\ 0, & \text{otherwise} \end{cases} \quad (4)$$

3. 快速傅立葉轉換（Fast Fourier Transform, FFT）是一種將聲音信號從時間域轉換到頻率域的技術。在時間域中，聲音信號的頻率和幅度會隨時間快速變化，難以有效地表現出聲音信號的特性。通過將聲音信號轉換到頻率域，快速傅立葉變換可以顯示聲音信號的頻率和能量分佈特性。如公式5 所示：

$$s_a(f) = \sum_{n=1}^N \tilde{s}(n) \cdot e^{-\frac{j2\pi f n}{F}}, \quad 0 \leq f \leq F \quad (5)$$

公式5中， $s_a(f)$ 為 \tilde{s} 在第 f 個頻率取樣本點的聲音訊號值， N 為音框內聲音訊號之樣本數， F 為頻率取樣個數。

4. 三角帶濾波器（Triangular filter bank）是一種將音頻信號轉換為梅爾尺度（Mel scale）下的能量分佈的濾波器，如圖7所示。

通過將梅爾尺度下的頻率劃分為固定寬度的三角帶濾波器，可以更好地模擬

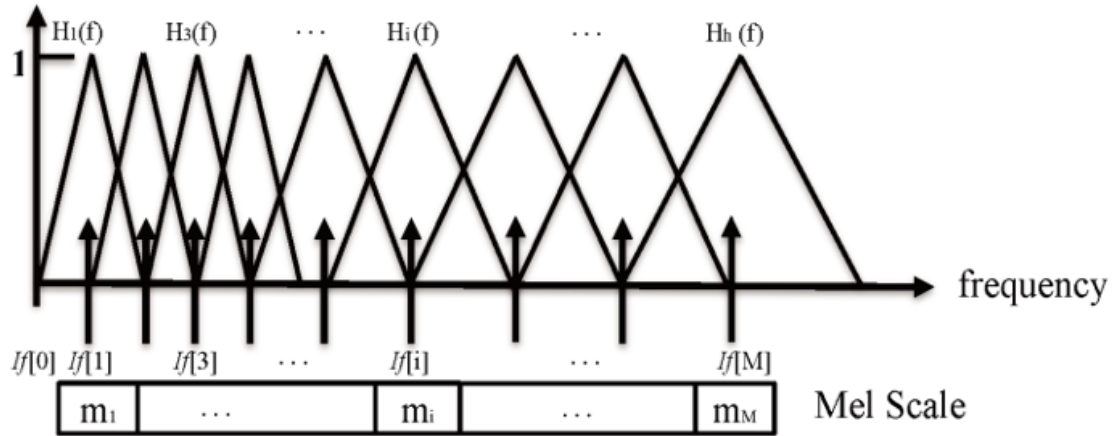


圖 7: 梅爾濾波器組

人耳對不同頻率的聲音感知，因為人耳的感知是對數刻度。每個三角帶濾波器的中心頻率位於梅爾尺度上的一個固定點，並通過一個線性上升和下降的斜坡來濾波中心頻率周圍的頻率帶。三角帶濾波器的數量和中心頻率的位置根據具體應用和需求來決定。通常三角帶濾波器的數量為 20 到 40 個，本研究設置了 22 個三角帶濾波器。梅爾尺度的轉換公式如公式6所示：

$$\text{Mel}(f) = 2595 \ln \left(1 + \frac{f}{700} \right) \quad (6)$$

公式6中， $\text{Mel}(f)$ 為梅爾頻率， f 為頻率。

再將各頻率的能量乘上三角帶濾波器並累加起來，取對數值求得對數能量，如公式8所示：

$$H_m(f) = \begin{cases} 0, & \text{if } f < f_{m-1} \\ \frac{f-f_{m-1}}{f_m-f_{m-1}}, & \text{if } f_{m-1} \leq f < f_m \\ \frac{f_{m+1}-f}{f_{m+1}-f_m}, & \text{if } f_m \leq f < f_{m+1} \\ 0, & \text{if } f \geq f_{m+1} \end{cases} \quad (7)$$

$$E_m = \ln \left[\sum_{f=0}^F |s_a[f]|^2 H_m(\text{Mel}(f)) \right], \quad m = 1, 2, \dots, M \quad (8)$$

公式7,8中， E_m 為第 m 個三角帶濾波器輸出的對數能量值， $|s_a[f]|^2$ 為第 f 個頻率的能量， $H_m(k)$ 為頻率 k 在第 m 個三角帶濾波器之權重， M 代表三角帶濾波器總個數。

5. 離散餘弦轉換 (Discrete Cosine Transformation, DCT) 的主要目的是將頻率軸上的能量譜值劃分為多個子帶，每個子帶對應一個梅爾濾波器的輸出能量。首先，對每個子帶的能量值取自然對數，然後應用 DCT 轉換，將這些對數能量值轉換為一組 MFCC 係數，最終獲得一個更加緊湊的低維度特徵向量。DCT 轉換可以有效地將高維度的數據向量壓縮成較低維度的特徵向量，就總共 M 個頻帶中，取 C 個組成數據的主要特徵。離散餘弦轉換的第 i 個特徵計算如公式9 所示：

$$\text{MFCC}_i = \sum_{l=1}^C \log(E_m) \cos \left[\frac{\pi}{M} (i - 0.5)(l - 1) \right], \quad l = 1, 2, \dots, C \quad (9)$$

公式9中， MFCC_i 稱為梅爾頻率倒譜係數， E_m 為三角帶濾波器得出的對數能量值， M 為三角帶濾波器的個數， C 為梅爾頻率倒譜係數的總維度個數。

離散餘弦轉換讓信號從頻率域轉回時間域，可求得梅爾倒頻譜係數。通常只取通常只取前 13 個或 20 個係數作為特徵向量，因前幾個係數包含了大部分的能量和主要信息，能夠很好地描述信號的主要特徵，同時會捨棄第一個倒譜係數（即零階係數），因我們對音頻信號進行了梅爾頻率濾波器組的處理，零階係數的數值會較大，包含了信號的整體能量訊息，對後續的分析來說，會干擾其他有用特徵。

因此本研究從離散餘弦轉換取得第 2 至 13 維的梅爾倒頻譜係數，作為特徵

向量。

6. 對數能量 (Logarithmic Energy, LE)：是用來表示每個音框中能量變化的一個指標，能夠反映信號的振幅變化和強度。對數能量的計算公式如10：

$$LE = \ln \left[1 + \sum_{n=1}^N (\tilde{s}[n])^2 \right] \quad (10)$$

公式10中， $\tilde{s}[n]$ 是經過加窗處理後的信號， N 為音框內聲音訊號之樣本數。

7. 幅度譜 (Magnitude Spectrum)：是用來表示信號在頻域中各個頻率成分上的強度，能夠代表信號中不同頻率成分的能量分布。其計算過程是從快速傅立葉變換的結果中計算幅度譜，計算公式如11,12:

$$s_a(f) = \sum_{n=1}^N \tilde{s}(n) \cdot e^{-\frac{j2\pi f n}{F}}, \quad 0 \leq f \leq F \quad (11)$$

$$MS(f) = \ln [|s_a(f)|] \quad (12)$$

其中， $s_a(f)$ 為 \tilde{s} 在第 f 個頻率取樣點的訊號值， $|s_a[f]|$ 為第 f 個頻率的能量。本研究提取幅度譜特徵值來代表股價數據在不同時間段的頻率能量，其值可以顯示股市的波動性。高頻成分的強度較大，說明股市在短時間內波動劇烈；而低頻成分的強度較大，則說明股市有較明顯的長期趨勢

本研究透過以上過程整合 12 個梅爾倒頻譜係數、1 個對數能量和 7 個幅度譜特徵，使每個幀有 20 維的信號特徵向量。

3.3 研究模型：LSTM 模型架構

本研究使用 LSTM（長短期記憶網絡）來建構股價預測模型。在股價預測的分類任務中，歷史股價資料具有時間序列的特性，因此使用 LSTM 可以有效捕捉資料中的時間依賴性和序列模式。同時，本研究將不同時間粒度的數據集作為 LSTM 模型的輸入，使模型能夠更好地學習股價的短期、中期和長期趨勢，從而做出更加準確的預測。LSTM 模型架構圖如圖8所示：

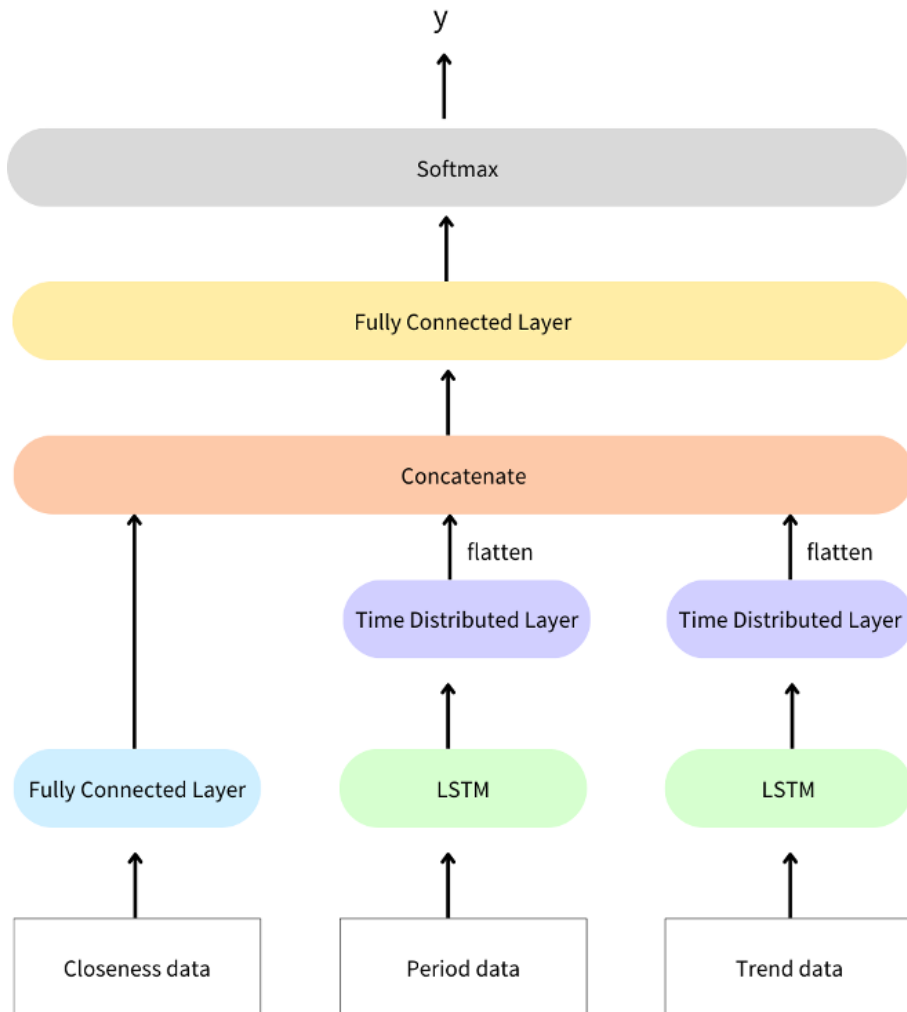


圖 8: LSTM 架構圖

3.3.1 LSTM 層

透過 LSTM 層來捕捉股價趨勢的時間依賴性和序列模式，將每一週期的特徵向量 $\{x_1, x_2, \dots, x_t, \dots, x_T\}$ 輸入到 LSTM 層，由 LSTM 層的三個控制閥門（如圖9所示）：遺忘門、輸入門、輸出門以及單元狀態所組成，分別用於重置過去不必要的資訊、寫入新資訊、輸出所需的資訊以及作為儲存記憶的空間。在 LSTM 中，每個控制閥會通過 Sigmoid 激活函數將值映射到 0 到 1 之間，從而控制信息的流動。而單元狀態則通過 tanh 激活函數將值映射到 -1 到 1 之間，進行信息的更新和使用，這樣可以避免梯度消失或爆炸問題。

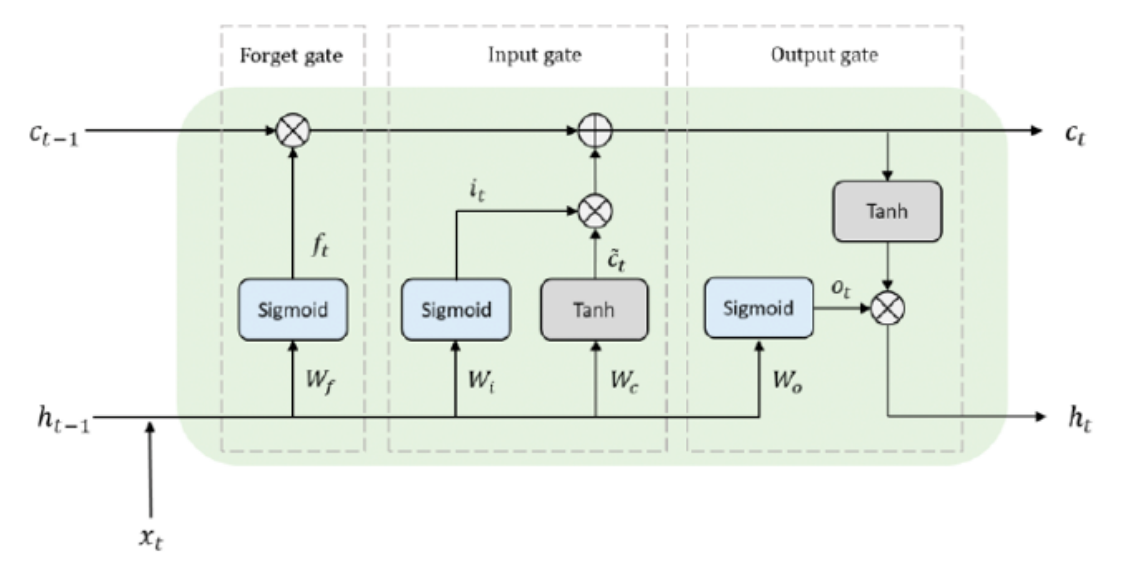


圖 9: LSTM 單元結構圖

遺忘門用於判斷前一個時間點的記憶單元 C_{t-1} 是否要遺忘或是保留，將當前輸入值 X_t 與前一期隱藏狀態 h_{t-1} 乘上一個權重矩陣，並經過 Sigmoid 激活函數來決定單元狀態中應該丟棄或保留多少資訊，其值越接近 0 意味著應該丟棄，越接近 1 意味著應該保留，如公式13所示：

$$f_t = \text{Sigmoid}(W_f[x_t; h_{t-1}] + b_f) \quad (13)$$

其中 x_t 為當期輸入， h_{t-1} 為前一期的隱藏狀態， f_t 為遺忘門的輸出， W_f 和 b_f 分別為遺忘門的權重和誤差。

輸入門有兩個步驟，第一步為決定哪些資訊要被寫入單元狀態，將當前輸入值 X_t 與前一期隱藏狀態 h_{t-1} 乘上一個權重矩陣，並經過 Sigmoid 激活函數來決定要更新哪些資訊，其值為 0 表示不重要，1 表示重要，如式10所示。第二步為生成候選狀態，這表示當前需要考慮的新資訊，通過學習當前輸入和前一時空步隱藏狀態的權重，並經過 tanh 激活函數來生成候選狀態，如公式11所示。接著，將這兩個步驟的輸出進行相乘，通過第一步 Sigmoid 的輸出來決定第二步候選狀態中的哪些訊息是重要且需要更新到單元狀態中的。最後，將遺忘門的輸出與輸入門的輸出合併，獲得當前新的單元狀態，如公式12所示。

$$i_t = \text{Sigmoid}(W_i[x_t; h_{t-1}] + b_i) \quad (10)$$

$$\tilde{c}_t = \tanh(W_c[x_t; h_{t-1}] + b_c) \quad (11)$$

$$c_t = f_t \cdot c_{t-1} + i_t \cdot \tilde{c}_t \quad (12)$$

其中， i_t 為輸入門的輸出， W_i 和 b_i 分別為輸入門的權重和誤差。 \tilde{c}_t 為候選狀態， W_c 和 b_c 為計算候選狀態的權重和誤差， c_t 和 c_{t-1} 分別為當期和前一期的單元狀態。

輸出門透過當前輸入值 X_t 與前一期隱藏狀態 h_{t-1} 乘上一個權重矩陣，並經過 Sigmoid 激活函數來決定單元狀態需要輸出哪些資訊，如式14所示。將更新後的新單元狀態透過 tanh 激活函數，輸出所需資訊以取得當期的隱藏狀態，如公式15所示。最後，將隱藏狀態作為當期 LSTM 單元的輸出，並將新的單元狀態和新的隱

藏狀態傳遞給下一期。

$$o_t = \text{Sigmoid}(W_o[x_t; h_{t-1}] + b_o) \quad (14)$$

$$h_t = o_t \cdot \tanh(c_t) \quad (15)$$

其中， o_t 為輸出門的輸出， W_o 和 b_o 分別為輸出門的權重和誤差。 h_t 為當期的隱藏狀態。取得 LSTM 層的隱藏狀態向量 $\{h_1, h_2, \dots, h_t, \dots, h_T\}$ 後，將其輸入至下一層的時間分佈層（Time Distributed Layer）。

3.3.2 時間分佈層（Time Distributed Layer）

在本研究中，我們使用時間分佈層（Time Distributed Layer）來處理不同時間粒度的數據集。將中期（period）和長期（trend）的數據集分別輸入到 LSTM 層後，使用時間分佈層（Time Distributed Layer）來應用的全連接層操作，以便在每個時間步上進行一致的處理，並同時保持輸入數據的時間維度，這樣的設計有助於捕捉每個時間步的特徵，並保留時間序列信息，從而提高模型對股價趨勢的預測能力。如公式15所示。

$$y_t = W^T \cdot h_t + b \quad (15)$$

其中， h_t 是第 t 個時間步 LSTM 的輸出， W^T 是全連階層的權重矩陣， b 是全連接層的誤差項， y_t 是第 t 個時間步的 time distributed layer 的值。

3.4 損失函數 (Loss Function)

損失函數 (Loss Function) 用於衡量模型的預測結果與實際值之間的差距，是神經網路中優化的目標函數。當損失函數的值越小，表示模型的預測結果越接近實際值。通過反向傳播算法，損失函數的梯度會被用來更新模型的權重，從而逐步優化模型。

在本研究中，我們將股價預測任務分為三種類別：上漲、持平、下跌，因此使用了 Categorical Cross entropy 作為損失函數，這是一種常用於多分類問題的損失函數，可以有效處理多類別分類問題中機率分佈的差異，確保模型輸出的概率分佈與真實標籤分佈盡可能接近，使模型具有分類能力。它的計算公式如16下：

$$\text{loss} = - \sum_{i=1}^N p_i \cdot \log(\hat{p}_i) \quad (16)$$

其中， p_i 是真實標籤， \hat{p}_i 是模型的預測機率。

第四章 實驗結果

該章節闡述本研究之實驗，分別為資料蒐集、將資料輸入模型前的預處理過程、衡量模型之評估指標、實驗設定以及最終的實驗結果與分析。

4.1 資料蒐集

本研究透過 Python 的 `yfinance` 庫取得特斯拉 (Tesla, Inc.) 在美國股票市場的歷史股價資料，作為實驗的原始數據。蒐集了 2014 年 1 月 1 日至 2023 年 12 月 31 日為期 9 年的每日收盤價和每日交易量。並對數據進行了標籤化處理，將每日收盤價的變動分為持平、上升和下降三種類別，我們設定價格變動的閾值為 1% (0.01)，用以判斷股價變動的顯著性。具體步驟如下：首先，將第一天的股價標記為持平 ('0')。接下來，從第二個交易日開始，計算每一天的收盤價相對於前一天的變化百分比。根據變化百分比的大小，將股價的變動分為三類：

- 若當天股價相較於前一天的變化百分比大於 1%，則標記為上升 ('1')。
- 若當天股價相較於前一天的變化百分比小於 -1%，則標記為下降 ('-1')。
- 若當天股價的變化百分比在 -1% 到 1% 之間，則標記為持平 ('0')。

標籤生成後，我們將標籤轉換為獨熱編碼 (One-Hot Encoding)，以適應分類模型的輸入需求。獨熱編碼是一種將分類標籤轉換為二進制向量的方法，每個標籤都被轉換為一個向量，該向量的維度等於分類的總數，每個維度對應一個特定的分類。如：持平 ('0') 標籤被編碼為 [1, 0, 0]，上升 ('1') 標籤被編碼為 [0, 1, 0]，下降 ('-1') 標籤被編碼為 [0, 0, 1]。

4.2 資料前處理

資料前處理將會分成兩部分介紹，第一部分將詳細說明特徵生成的前處理以及 MFCC 的參數設定；第二部分將說明訓練資料集與測試集的切分。

4.2.1 特徵生成

特斯拉歷史股價中的每日收盤價和每日交易量，將股票數據轉換為音頻信號的過程中，需要增加數據點使得每秒中有足夠多的樣本來表示數據的變化，因此，分別透過 MATLAB 進行線性內插值法，讓數據生成多一倍的數據。原始數據點共有 2516 個點，線性插值後的數據點共有 5032 個點，並將插值後的數據進行標準化，縮放於 $[-1,1]$ 之間，使音頻信號的振幅在處理和分析時保持在一個一致的範圍內，以利模型的權重更新會更加穩定，從而加快訓練速度。後轉換成波形數據，以便輸入至 MFCC 提取特徵。

其中，本研究起初嘗試了生成 8 倍、5 倍、3 倍等不同數量的數據，但最終決定只生成一倍數據。主要是基於特斯拉股價的特殊性，從圖10可以看出，大約有一半的原始數據點都處於一個相對水平的數值範圍內，如果生成過多的數據，這些相對穩定區域的數據會被過度擴展，從而在模型訓練過程中導致失真，產生模型過度擬合於那些並不代表真實市場變動的平穩數據段。

本研究使用 MFCC 作為股價數據的特徵，為了搭配 5 天為單位預測第六天股價變化，本研究使用採樣率 1258 作為目標採樣值，並設置每個長度為 0.008 秒的窗口，使得每個窗口內包含 10 個數據點，等同於 5 天的股價特徵值。並在窗口之間的步長本研究實驗了兩種參數，第一，是設置與窗口長度一樣長度為 0.008 秒的步長，如圖11，第二，是設置 0.0015 秒的步長，如圖12。對每個窗口的訊號應

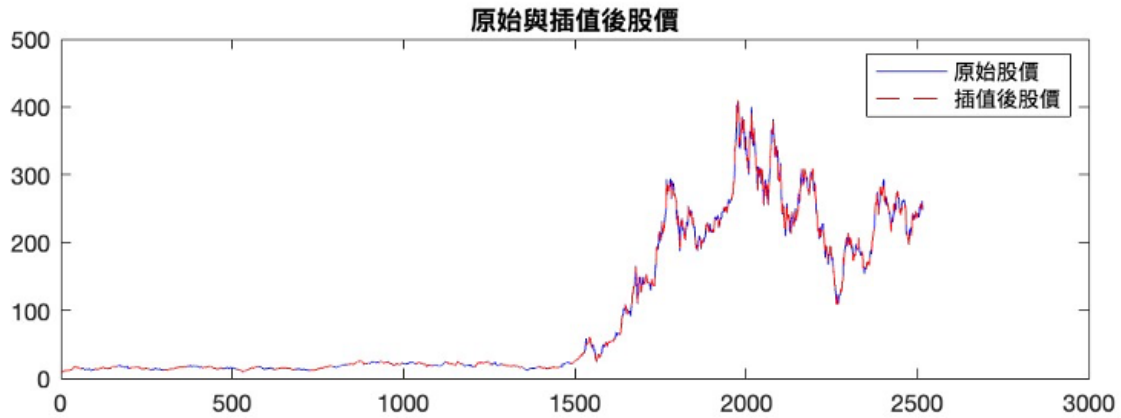


圖 10: 收盤價差值圖

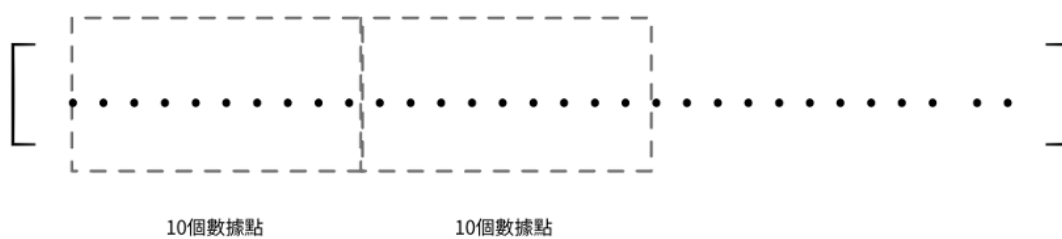


圖 11: 音框步長 0.008 秒示意圖

用 512 點的傅立葉轉換，取得頻域圖，並使用 22 個梅爾濾波器進行濾波後，取得頻域能量特徵，透過離散餘弦轉換讓信號從頻率域轉回時間域，取得梅爾倒頻譜係數。最終，本研究從 22 個梅爾倒頻譜係數中，提取前 2 到 13 個 MFCC 特徵。

其中，本研究爲了確定最適合的梅爾濾波器數量，嘗試了不同的梅爾濾波器數量，包括 40、30、22 等，發現使用 40 與 30 個梅爾濾波器時，模型容易出現過擬合現象。過多的濾波器數量雖然可以捕捉到更多的細節特徵，但也更容易將噪聲納入模型，導致模型在訓練數據上表現良好，但在驗證和測試數據上表現不佳。而 22 個梅爾濾波器則能有效避免這一問題，提供了更具泛化能力的特徵。

同時，本研究透過傅立葉轉換後取得頻域圖，可以從頻域圖中取得能量特徵，以增加數據特徵的多樣性，如圖13所示。從圖型上可以看出，較亮的部分表示在該時間段和頻率範圍內，股價數據有較高的能量，即在這些頻率上變動較大，較

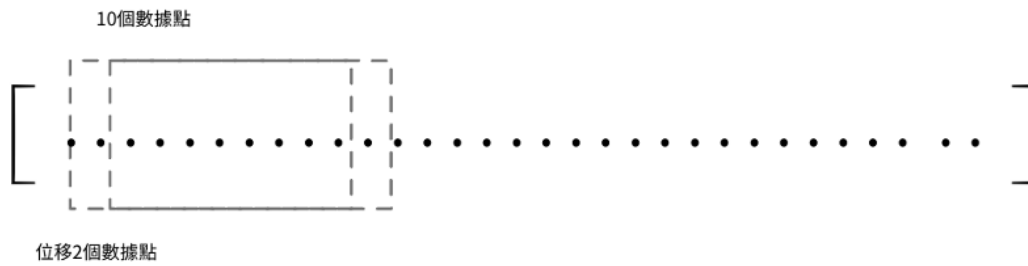


圖 12: 音框步長 0.0015 秒示意圖

暗的部分表示能量較低，這些頻率上的變動較小。本研究對訊號應用 14 點的傅立葉轉換後，可以觀察到前 7 個頻率分段的能量較高，而超過 7 個頻率分段的能量顯著降低。這代表大部分有用的訊息集中在前 7 個頻率分段中，而後續頻率分段的能量則較低，包含的訊息較少。

最終，我們取前 7 個特徵值作為分析的頻率特徵，用來代表信號在不同頻率上的能量。

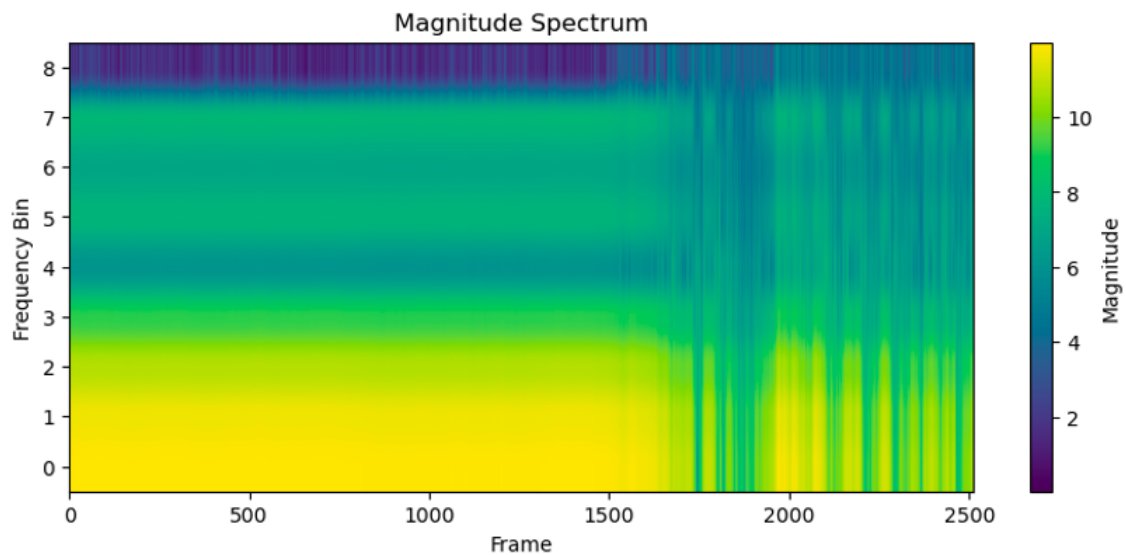


圖 13: 頻域圖

最後，我們對每個音框取對數能量特徵，此特徵能夠反映信號的振幅變化和強度，結合以上特徵數據，我們從 MFCC 中對每一音框提取共 20 維的特徵值，如圖14。請參照表格1中 MFCC 特徵萃取所設置的參數。

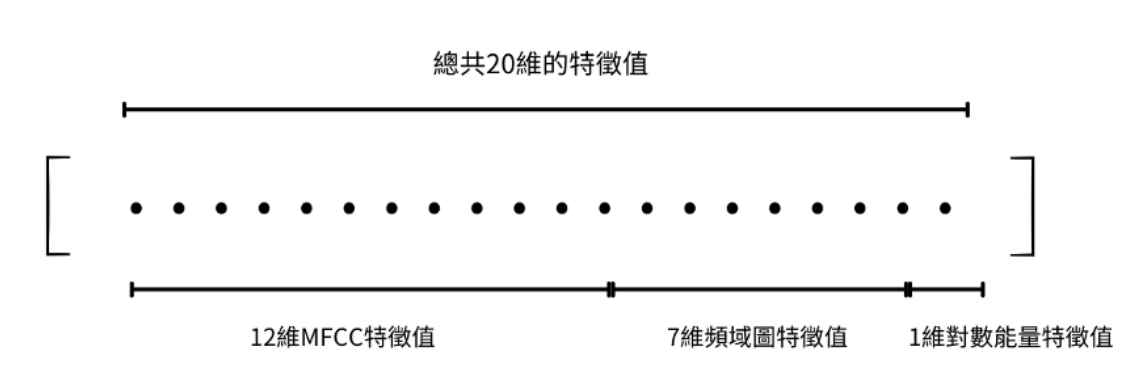


圖 14: 每個音框的特徵維度

表 1: 實驗參數表

特徵萃取參數	實驗一	實驗二
取樣率	1258	1258
音框長度	0.008s	0.008s
音框步長	0.008s	0.0015s
音框數量	504	2512
三角帶濾波器	22	22
傅立葉轉換的頻率切格點數	512	512
頻域圖能量特徵	7	7
對數能量特徵	1	1
MFCC 特徵數	12	12

其中，使用了兩種不同的音框步長參數設置，以研究不同步長對特徵提取效果的影響。第一種設置方法是將音框步長與窗口大小設置為相同的長度，即 0.008 秒，代表音框之間沒有重疊，每一個音框都包含獨立的 5 天股價數據。第二種設置方法是將音框步長設置為 0.0015 秒，基於本研究 MFCC 的參數設定相當於位移 1 天的股價數據。換句話說，每個音框之間會有四天的數據是重疊的。

本研究設計兩種參數的原因在於探索不同的數據重疊方式對模型性能的影響。當音框步長與窗口大小相同時，數據之間沒有重疊，適合用來捕捉每一段獨立的數據特徵；而當音框步長設置為 0.0015 秒時，每個音框之間的重疊部分提供了連續的數據訊息，能夠捕捉更細緻的變化趨勢和關聯性。

4.2.2 資料集切分

在經過 MFCC 提取特徵值後，我們獲得了一組形狀為 (num frames, features) 的數據集，每個音框包含 5 天的信息，來代表周線股價的趨勢特徵。為了進一步捕捉股價的中期和長期趨勢，我們對數據集進行了以下切分：

- 月線數據集：基於第一組數據集，我們將四組音框合併為一組，每組音框包含 20 天的信息，這樣就得到了形狀為 (num frames, 4, features) 的數據集，代表月線股價的趨勢特徵。
- 季線數據集：在此基礎上，我們將十二組音框合併為一組，每組音框包含 60 天的信息，從而得到形狀為 (num frames, 12, features) 的數據集，代表季線股價的趨勢特徵。

這樣的數據切分可以幫助模型在不同時間粒度下學習股價的趨勢變化，從而提高預測的準確性。本研究分別將特斯拉歷史股價中的每日收盤價和每日交易量，分別透過 MFCC 提取特徵值後，將產生 (num frames, features) 的數據集，代表周線股價的趨勢特徵，並且透過資料集的重組、切分，將生成另外兩組數據集，分別代表月線數據集以及季線數據集。最後，我們將每日收盤價與每日交易量的數據集合併，得到適合輸入到模型訓練的特徵數據集。資料集參數表如表格2。

在輸入模型訓練前，我們將數據集切割為為訓練集 (Training set)、驗證集 (Validation set)、測試集 (Testing set)，分別為 0.7、0.2、0.1，如表格3與表格4所示。

表 2: 資料集實驗參數表

資料集		實驗一	實驗二
每日收盤價	周線特徵資料集	(493, 20)	(2501, 20)
	月線特徵資料集	(493, 4, 20)	(2501, 4, 20)
	季線特徵資料集	(493, 12, 20)	(2501, 12, 20)
每日交易量	周線特徵資料集	(493, 20)	(2501, 20)
	月線特徵資料集	(493, 4, 20)	(2501, 4, 20)
	季線特徵資料集	(493, 12, 20)	(2501, 12, 20)
特徵數據集	周線特徵資料集	(493, 40)	(2501, 40)
	月線特徵資料集	(493, 4, 40)	(2501, 4, 40)
	季線特徵資料集	(493, 12, 40)	(2501, 12, 40)

表 3: 實驗一的訓練集、驗證集與測試集

訓練集 (70%)，共 345 筆			
	下跌 (-1)	持平 (0)	上漲 (1)
資料標籤	114 (33.04%)	128 (37.10%)	103 (29.86%)
驗證集 (20%)，共 98 筆			
	下跌 (-1)	持平 (0)	上漲 (1)
資料標籤	37 (37.76%)	30 (30.61%)	31 (31.63%)
測試集 (10%)，共 50 筆			
	下跌 (-1)	持平 (0)	上漲 (1)
資料標籤	15 (30.00%)	23 (42.00%)	12 (28.00%)

表 4: 實驗二的訓練集、驗證集與測試集

訓練集 (70%)，共 1750 筆			
	下跌 (-1)	持平 (0)	上漲 (1)
資料標籤	555 (31.71%)	580 (33.14%)	615 (35.14%)
驗證集 (20%)，共 500 筆			
	下跌 (-1)	持平 (0)	上漲 (1)
資料標籤	182 (36.40%)	151 (30.20%)	167 (33.40%)
測試集 (10%)，共 251 筆			
	下跌 (-1)	持平 (0)	上漲 (1)
資料標籤	73 (29.08%)	103 (41.04%)	75 (29.88%)

4.3 模型評估指標

根據模型對資料的預測結果與實際類別，可以將結果分成 9 種類別，而這些類別所組成的矩陣稱為混淆矩陣（Confusion Matrix），在三類分類問題中，混淆矩陣是一個 3x3 的矩陣，如表5。

表 5: 混淆矩陣

	預測類別		
	類別 1	類別 2	類別 3
實際類別 1	TP	FN	FN
實際類別 2	FP	TP	FN
實際類別 3	FP	FN	TP

其中，TP 代表實際類別為 Positive，且預測結果亦為 Positive；FN 代表實際類別為 Positive，但預測結果為 Negative；FP 代表實際類別為 Negative，但預測結

果為 Positive；TN 代表實際類別為 Negative，且預測結果亦為 Negative。基於混淆矩陣，可以計算出用來評估模型性能的指標，包含準確度 (Accuracy)、精確度 (Precision)、召回率 (Recall)、F1-score，本研究將以上四個指標做為模型表現之評估表現，其公式分別為公式17,18,19,20。

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (17)$$

$$Precision = \frac{TP}{TP + FP} \quad (18)$$

$$Recall = \frac{TP}{TP + FN} \quad (19)$$

$$F1\ Score = 2 * \frac{Precision * Recall}{Precision + Recall} \quad (20)$$

4.4 實驗設定

本研究實驗是採用 Python 撰寫，利用 Keras 深度學習框架來進行模型訓練與測試，此節將說明 LSTM 模型之架構、超參數設置以及損失函數，同時本研究將周線特徵資料集輸入至支持向量機 (Support Vector Machine, SVM)，作為與 LSTM 模型之比較。

4.4.1 LSTM 預測模型架構與超參數設置

實驗一：

- Number of LSTM layer : 1
- Epoch : 500
- Batch Size : 128

- Learning rate : 0.0001
- Optimizer : Adam
- Loss Function : categorical cross entropy

5 days data				20 days data				60 days data			
Layer	Input shape	Output shape	Activation Function	Layer	Input shape	Output shape	Activation Function	Layer	Input shape	Output shape	Activation. Function
FC	40	20	LeakyReLU	LSTM	4 x 40	4 x 16	tanh	LSTM	12 x 40	12 x 16	tanh
				Time Distributed	4 x 16	4 x 16	tanh	Time Distributed	12 x 16	12 x 16	tanh
				Flatten	4 x 16	64	None	Flatten	12 x 16	192	None
concatenate (20 + 64 + 192)											
Layer		Input shape		Output shape		Activation Function					
Dropout(0.1)		276		276		None					
FC		276		128		LeakyReLU					
Dropout(0.1)		128		128		None					
FC		128		24		LeakyReLU					
FC		24		3		Softmax					

圖 15: LSTM 模型架構，實驗一

實驗一，我們研究了音框步長與音框大小設置為相同長度（0.008 秒）的效果。這種設置代表音框之間沒有重疊，每一個音框都包含獨立的 5 天股價數據，這樣的設置可以幫助我們捕捉每段獨立數據的特徵，適合用來分析獨立時間段的股價變化。並透過建構 LSTM 模型將三種不同時間粒度的數據集，輸入至模型訓練，以利學習不同時間趨勢的特徵。

首先，對於短期數據集，模型接收形狀為 40 的輸入數據，這些數據進入全連

接層進行處理，輸出 20 維的特徵向量，並應用 LeakyReLU 激活函數。

接著，中期的數據集以形狀為 4×40 的矩陣形式輸入模型，這些數據首先通過 LSTM 層提取時間序列中的特徵，輸出形狀為 4×16 的矩陣，並使用 tanh 激活函數。然後，特徵向量進入 Time Distributed 層，保持輸入形狀為 4×16 的矩陣並應用 tanh 激活函數。最後，數據經過一個 Flatten 層，將形狀轉換為 64 維的向量。

同樣地，長期的數據集以 12×40 的矩陣形式輸入模型，這些數據也首先通過 LSTM 層，該層的輸出形狀為 12×16 ，並使用 tanh 激活函數。隨後，數據進入一個 Time Distributed 層，這層應用 tanh 激活函數並保持輸入形狀為 12×16 的矩陣。最後，數據經過一個 Flatten 層，將形狀轉換為 192 維的向量。

各部分數據經過特徵提取及降維處理後，在 Concatenate 層進行合併，合併後的特徵向量形狀為 276 維。合併後的數據首先通過一個 Dropout 層，保持輸入形狀為 276 並應用 0.1 的 Dropout 比例。接著，數據進入全連接層，該層輸出形狀為 128 並應用 LeakyReLU 激活函數。之後，數據再次通過一個 Dropout 層，保持輸入形狀為 128 並應用 0.1 的 Dropout 比例。然後，數據進入另一個全連接層，該層輸出形狀為 24 並應用 LeakyReLU 激活函數。最終，數據進入 Softmax 層，該層將數據轉換為 3 維的輸出，用於分類任務。

實驗二：

- Number of LSTM layer : 1
- Epoch : 500
- Batch Size : 256
- Learning rate : 0.0001

- Optimizer : Adam
- Loss Function : categorical cross entropy

5 days data				20 days data				60 days data			
Layer	Input shape	Output shape	Activation Function	Layer	Input shape	Output shape	Activation Function	Layer	Input shape	Output shape	Activation Function
FC	40	20	LeakyReLU	LSTM	4 x 40	4 x 20	tanh	LSTM	12 x 40	12 x 20	tanh
				Time Distributed	4 x 20	4 x 20	tanh	Time Distributed	12 x 20	12 x 20	tanh
				Flatten	4 x 20	80	None	Flatten	12 x 20	240	None
concatenate (20 + 80 + 240)											
Layer		Input shape		Output shape		Activation Function					
Dropout(0.05)		340		340		None					
FC		340		128		LeakyReLU					
Dropout(0.05)		128		128		None					
FC		128		24		LeakyReLU					
FC		24		3		Softmax					

圖 16: LSTM 模型架構，實驗二

實驗二，我們研究了音框步長設置為 0.0015 秒，而音框大小設置為 0.008 秒的效果。這種設置代表音框之間會有四天的數據重疊。這樣的設置有助於捕捉連續的數據訊息，能夠更細緻地提取變化趨勢和關聯性。

首先，對於短期數據集，模型接收形狀為 40 的輸入數據，這些數據進入全連接層進行處理，輸出 20 維的特徵向量，並應用 LeakyReLU 激活函數。

接著，中期的數據集以形狀為 4x40 的矩陣形式輸入模型，這些數據首先通過 LSTM 層提取時間序列中的特徵，輸出形狀為 4x20 的矩陣，並使用 tanh 激活函

數。然後，特徵向量進入 Time Distributed 層，保持輸入形狀為 4x20 的矩陣並應用 tanh 激活函數。最後，數據經過一個 Flatten 層，將形狀轉換為 80 維的向量。

同樣地，長期的數據集以 12x40 的矩陣形式輸入模型，這些數據也首先通過 LSTM 層，該層的輸出形狀為 12x20，並使用 tanh 激活函數。隨後，數據進入一個 Time Distributed 層，這層應用 tanh 激活函數並保持輸入形狀為 12x20 的矩陣。最後，數據經過一個 Flatten 層，將形狀轉換為 240 維的向量。

各部分數據經過特徵提取及降維處理後，在 Concatenate 層進行合併，合併後的特徵向量形狀為 340 維。合併後的數據首先通過一個 Dropout 層，保持輸入形狀為 340 並應用 0.05 的 Dropout 比例。接著，數據進入全連接層，該層輸出形狀為 128 並應用 LeakyReLU 激活函數。之後，數據再次通過一個 Dropout 層，保持輸入形狀為 128 並應用 0.05 的 Dropout 比例。然後，數據進入另一個全連接層，該層輸出形狀為 24 並應用 LeakyReLU 激活函數。最終，數據進入 Softmax 層，該層將數據轉換為 3 維的輸出，用於分類任務。

4.4.2 支持向量機 (Support Vector Machine, SVM)

本研究使用周線特徵資料集來作為 SVM 的輸入資料集，與 LSTM 模型相同，我們將實驗一與實驗二的資料集分別輸入 SVM，並利用 Optuna 於自訂區間內找到最適解，如表6。

表 6: SVM 超參數設定

Hyperparameters	Range
PCA	13, 10
Kernel	Poly, RBF
Cost	0.01, 0.1, 1, 10, 100
Gamma	0.01, 0.1, 1, 10, 100

4.5 實驗結果與分析

該小節說明 LSTM 模型配置的選擇與 SVM 模型進行比較的實驗結果。透過探索不同的數據重疊方式對模型性能的影響，找出最適的 MFCC 特徵數據集，以應用於股價預測領域。

實驗結果如表 9 所示，最左側的資料集分別為實驗一：音框為獨立的 5 天的股價數據以及實驗二：音框之間會有四天的數據是重疊的股價數據，評估指標中的 A, P, R, F 分別代表 Accuracy, Precision, Recall, F1-score，一般來說，最常用的評估指標是 Accuracy，但對於多分類任務來說 Accuracy 不一定能全面評估各類別的預測能力，則需要使用其他評估指標如 Precision、Recall 或 F1-score，這些指標能夠提供更全面的評估，專注於少數類別的預測表現。其中，F1-score 相較於 Precision 和 Recall，可以更有效地區分模型的好壞。F1-score 是 Precision 和 Recall 的調和平均數，一個好的分類器應該能夠在保持高 Precision 的同時，達到較高的 Recall，從而獲得較高的 F1-score。

在實驗一中顯示，LSTM 模型對於「下跌」與「上漲」比「持平」類別有較好的表現，表示 LSTM 模型在判別「持平」這個類別有困難，而 SVM 模型在所有三個類別上的整體性能均優於 LSTM 模型，特別是在「下跌」和「上漲」類別上的 F1-score 均高於 0.68，顯示了其在這些類別上好的預測能力。

在實驗二中顯示，LSTM 整體性能與實驗一相比在所有類別上的判別能力都有所提升，特別是在「持平」類別上的 F1-score 從 0.23 提高到 0.47，顯示在數據重疊下，LSTM 模型的判別能力有所增加。而 SVM 模型在實驗二與實驗一相比整體判別性能都有所下降。

通過兩組實驗結果的比較，可以發現當音框之間存在重疊時，LSTM 模型的

性能明顯提升，本研究認為導致其最大差異的原因是來自兩組實驗的樣本數差異，實驗一的樣本數為 493，實驗二的樣本數為 2501，相差有五倍之多，而對於深度學習模型，特別是像 LSTM 這樣的複雜神經網絡，通常需要大量樣本來充分訓練和學習數據中的模式和特徵，使其能進行有效訓練，以避免過擬合（overfitting）和提高泛化能力。所以，當樣本數增加到 2501 時，有助於 LSTM 模型更好地學習和識別股價變動模式，並且三種類別的判別性能都在一個相對好的能力。

表 7: 實驗評估指標

Dataset	Model	units	Evaluation Metric				Class
			A	P	R	F	
實驗一	LSTM	16	0.42	0.40	0.62	0.48	Down
				0.60	0.14	0.23	Same
				0.40	0.62	0.49	Up
	SVM	13	0.6634	0.65	0.71	0.68	Down
				0.59	0.52	0.55	Same
				0.73	0.75	0.74	Up
實驗二	LSTM	20	0.4628	0.43	0.48	0.46	Down
				0.48	0.46	0.47	Same
				0.45	0.44	0.45	Up
	SVM	13	0.4533	0.44	0.5	0.47	Down
				0.43	0.42	0.43	Same
				0.5	0.44	0.47	Up

表 8: 混淆矩陣

Dataset	Model	Confusion Metric	
		Training	Testing
實驗一	LSTM		
	SVM		
實驗二	LSTM		
	SVM		

第五章 結論

5.1 研究結論

本研究旨在提出了一種新的特徵數據應用於股價預測，基於梅爾頻率倒譜係數（MFCC）的特徵提取方法，將股價歷史數據轉換為波形數據，並提取多維度的特徵向量。將資料集切割為短期、中期和長期數據，以符合股價技術線圖常見的周線、季線、月線趨勢，從而有助於模型在不同時間粒度下的訓練特徵。在多次實驗測試下，本研究找出 MFCC 的最適參數，並探討在音框之間沒有重疊與音框之間有重疊之下，輸入支持向量機（SVM）和本研究構建的長短期記憶（LSTM）模型，以評估數據集對股價預測的有效性。

在實驗一中，LSTM 模型對於「下跌」與「上漲」類別的預測比「持平」類別有較好的表現，這表明 LSTM 模型在判別「持平」這個類別時存在困難。然而，SVM 模型在所有三個類別上的整體性能均優於 LSTM 模型，特別是在「下跌」和「上漲」類別上的 F1-score 均高於 0.68，顯示了其在這些類別上的優異預測能力。

在實驗二中，LSTM 整體性能相比實驗一在所有類別上的判別能力都有所提升，特別是在「持平」類別上的 F1-score 從 0.23 提高到 0.47，這顯示在數據重疊的情況下，LSTM 模型的判別能力有所增加。然而，SVM 模型在實驗二中的整體判別性能相比實驗一有所下降。

本研究的結果證明使用 MFCC 提取多維特徵數據結合 LSTM 模型，在股價預測領域可達 46.28% 的準確率，為未來的研究提供一種新的特徵提取方法和方向。並且，SVM 模型在所有類別上的預測性能均優於 LSTM 模型，特別是在數據未重疊的情況下，其 F1-score 均顯著高於 LSTM 模型，可達 66.34% 的準確率，這顯示

了 SVM 模型在股價預測上的穩定性和準確性。我們可以推測當樣本數據較少時，SVM 模型的性能會優於 LSTM 模型，原因在於 SVM 模型在小樣本數據情況下能夠更有效地學習和識別數據中的模式和特徵，而 LSTM 模型則需要更多的數據來充分訓練和避免過擬合（overfitting）。因此，在樣本數據較少的情況下，SVM 模型能夠提供更好的預測效果。

5.2 研究限制與未來研究建議

本研究也存在一些限制。首先，本研究僅使用了特斯拉公司的歷史股價數據，未能涵蓋其他行業和公司的股票數據，未來可以擴展數據集以驗證模型的泛化性。同時，未來的研究可以結合更多不同類型的數據源，如社會經濟指標、公司財務報表等，進一步提升股價預測的準確性。

本研究提出基於 LSTM 模型建構預測模型，未來可以嘗試建構更複雜的深度學習模型，如 Transformer 模型等，進一步提升時間序列數據的預測能力；且可以從研究中發現，隨著樣本數據的增加，LSTM 模型的性能將會逐漸提升，但由於資料量的不足，LSTM 模型效果沒有較理想，因此，建議未來的研究可以增加實驗資料量，以獲得更充足的訓練資料，進而提高預測準確度。

參考文獻

- [1] P. Sahoo and K. Charlapally, “Stock price prediction using regression analysis.,” *International Journal of Scientific Engineering Research*, vol. 6, no. 3, p. 1655—1659, 2015.
- [2] Y. Fu and F. Li, “Application of arma model in sports stock price prediction in china,” *Statistics and Decision*, vol. 21, no. 3, pp. 101–103, 2012.
- [3] W. Lu, J. Li, J. Wang, and L. Qin, “A cnn-bilstm-am method for stock price prediction,” *Neural Comput. Appl.*, vol. 33, pp. 4741–4753, May 2021.
- [4] H. Xu, B. Xu, and K. Xu, “Overview of the application of machine learning in stock forecasting,” *Comput. Eng. Appl.*, vol. 56, no. 12, pp. 19–24, 2020.
- [5] S. Wang, “A stock price prediction method based on bilstm and improved transformer,” *IEEE Access*, vol. 11, pp. 104211–104223, 2020.
- [6] R. T. Nazário, E. S. e Silva, V. A. Sobreiro, and H. Kimura, “A literature review of technical analysis on stock markets,” *The Quarterly Review of Economics and Finance*, vol. 66, pp. 115–126, 2017.
- [7] A. Picasso, S. Merello, Y. Ma, L. Oneto, and E. Cambria, “Technical analysis and sentiment embeddings for market trend prediction,” *Expert Systems with Applications*, vol. 135, pp. 60–70, 2019.
- [8] Y. Gao, L. Zhou, Y. Zhang, C. Xing, Y. Sun, and X. Zhu, “Sentiment classification for stock news,” in *5th international conference on pervasive computing and applications*, pp. 99–104, IEEE, 2010.

- [9] C. Yu, Y. Gong, F. Wang, and L. An, “Predicting stock prices with text and price combined model,” *Data Analysis and Knowledge Discovery*, vol. 2, no. 12, pp. 33–42, 2019.
- [10] W. Long, Z. Lu, and L. Cui, “Deep learning-based feature engineering for stock price movement prediction,” *Knowledge-Based Systems*, vol. 164, pp. 163–173, 2019.
- [11] H. Cui, H. Yi, F. Tang, and K. Zhang, “Stock price prediction based on wavelet transform and group of long-short term memory network,” in *2023 IEEE 6th International Conference on Electronic Information and Communication Technology (ICEICT)*, pp. 1008–1011, 2023.
- [12] S. Davis and P. Mermelstein, “Comparison of parametric representations for monosyllabic word recognition in continuously spoken sentences,” *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 28, no. 4, pp. 357–366, 1980.
- [13] V. Bansal, G. Pahwa, and N. Kannan, “Cough classification for covid-19 based on audio mfcc features using convolutional neural networks,” in *2020 IEEE International Conference on Computing, Power and Communication Technologies (GU-CON)*, pp. 604–608, 2020.
- [14] S. A. A. Yusuf and R. Hidayat, “Mfcc feature extraction and knn classification in ecg signals,” in *2019 6th International Conference on Information Technology, Computer and Electrical Engineering (ICITACEE)*, pp. 1–5, IEEE, 2019.
- [15] J. Dutta and D. Chanda, “Music emotion recognition in assamese songs using mfcc features and mlp classifier,” in *2021 International Conference on Intelligent Technologies (CONIT)*, pp. 1–5, 2021.

- [16] T. Maeda and M. Yajima, "Motion frequency data analysis for sports skill," in *2017 Computing Conference*, pp. 113–116, 2017.
- [17] M. Othman, A. Wahab, and R. Khosrowabadi, "Mfcc for robust emotion detection using eeg," in *Proc. IEEE 9th Malaysia Int. Conf. Commun.*, p. 5532, 2009.
- [18] S. A. A. A. Yusuf and R. Hidayat, "Mfcc feature extraction and knn classification in ecg signals," in *Proc. 6th Int. Conf. Inf. Technol., Comput. Electr. Eng. (ICITACEE)*, pp. 1–5, 2019.
- [19] M. Qi, R. Zhou, Q. Zhang, and Y. Yang, "Feature classification method of frequency cepstrum coefficient based on weighted extreme gradient boosting," *IEEE Access*, vol. 9, pp. 72691–72701, 2021.
- [20] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural computation*, vol. 9, no. 8, pp. 1735–1780, 1997.
- [21] C. K. Vignesh, "Applying machine learning models in stock market prediction," *EPRA Int. J. Res. Develop.*, vol. 5, pp. 395–398, Apr 2020.
- [22] M. Nabipour, P. Nayyeri, H. Jabani, A. Mosavi, and E. Salwana, "Deep learning for stock market prediction," *Entropy*, vol. 22, p. 840, Jul 2020.