# 基於社群媒體情緒與圖神經網路進行股票趨勢預測
# Stock Trend Prediction with Social Media Sentiment and Graph Neural Network

**Yen-Tsang Wu** **Jenq-Haur Wang**
Department of Computer Science and Information Engineering
National Taipei University of Technology
Web Information Retrieval Lab
buddyswu@gmail.com
jhwang@ntut.edu.tw

**Po Chuan Huang**
Taiwan Semiconductor Manufacturing Company Limited
peter02589@gmail.com

## 摘要

本文提出一種結合歷史股價與社群情緒的股票漲跌預測多模態整合架構。首先，我們將過去幾天的移動平均使用 GRU 取得序列式特徵，再將過去幾天股價之相對價差透過圖注意力機制得出價差特徵，並將社群言論之情緒分析也透過圖注意力機制得出情緒特徵；我們將三種不同性質之特徵互相結合，透過圖注意力機制得出股票特徵。最後將不同股票透過超圖神經網路預測出股票漲跌之結果。實驗結果顯示，本論文提出的模型在結合了多種不同性質特徵與考量不同股票之間關係後，相較於先前的方法，能更有效的偵測出股票漲跌。

## Abstract

This paper presents a novel multimodal fusion approach for stock trend prediction, integrating historical stock prices and social sentiment data. The proposed method first extracts sequential features from stock price moving averages over recent days using a Gated Recurrent Unit (GRU). Simultaneously, it captures relative spread features by applying Graph Attention Networks (GAT) to the relative price spreads, and sentiment features are derived from social sentiment analysis, also using GAT. These three feature sets are then fused using a graph attention mechanism to obtain comprehensive stock representations. Subsequently, a stock correlation graph is constructed, where nodes represent individual stocks, and edges reflect their correlations. Hypergraph neural networks are employed to predict stock trends based on this graph structure. Experimental results demonstrate that the proposed method achieves an accuracy of 0.604. By incorporating diverse feature types and accounting for inter-stock relationships, this approach significantly outperforms previous models in predicting stock price movements.

關鍵字：圖神經網路、社群媒體情緒、股票趨勢預測

Keywords: Graph neural network, Social media sentiment, Stock trend prediction

## 1 介紹

近年來，隨著經濟發展和全球化，股票市場交易活躍，股票投資成了重要的經濟行為。由於股票市場是一個複雜的系統，受到多種因素的影響，傳統的股票預測方法主要依賴於歷史股價數據的分析，如價格走勢圖、交易量等些指標能夠反映出市場的整體走向和投資者的交易行為 (Duan et al., 2022; Zhang et al., 2017)。

隨著社群媒體影響力日益增強，投資者情緒和觀點越來越多地在網路上表達。眾多研究表明，社交媒體上的情感信息可以反映市場情緒的變化，並影響股票價格的走向 (Behrendt et al., 2018; Ben Cheikh et al., 2024)。學者們指出，僅依靠分析社交媒體上與股票相關的討論和情感提供的額外信息，就能有助於預測未來的價格走勢。Tabari 等人 (Tabari et al., 2018) 研究發現,平均而言,社交媒體上發文的情緒會導致當天股票收益上升 0.26%。除此之外，由於美國股市在全球交易中占據重要地位， Mora 等人 (Nuñez-Mora et al., 2023) 分析了約 5000 萬條推文，計算了 2557 家美國上市公司的正面和負面情緒因子，結果顯示，503 家公司的負面情緒因子影響大於正面情緒，證明社交媒體情緒已成為市場的一個影響源。多項研究都支持社交媒體情緒能夠反映市場

情緒變化，並影響股票價格走向的觀點。投資者可利用社交媒體情緒信息來預測股價。Wang 等人 (Wang et al., 2023) 提出一個名為 ECON 的框架，利用推文、宏觀經濟指標和歷史價格來預測股票運動和波動性。該研究強調了推文數據的質量和行業相關性對預測的影響。Asgarov 等人 (Asgarov, 2023) 評估了社交媒體情感在預測主要公司股價中的有效性，他們使用 LSTM 模型分析推文情感和歷史價格數據，結果顯示情感表達與股價波動之間存在強相關性。Amin 等人 (Amin et al., 2024) 探討了社交媒體上關於人工智慧進展的情感是否能預測相關公司的日常股價波動。研究利用自然語言處理技術分析推文情感，並使用多種機器學習模型進行預測，發現推文情感與市場價值之間存在潛在的相關性。

基於上述的研究，我們提出一個結合技術指標、相對價差與社群言論情緒，並使用圖神經網路融合多種特徵，最後使用 Hypergraph Neural Networks (Feng et al., 2019)結合股票之間的關聯性，結合一個多層次圖神經網路模型，以提升股票趨勢預測的準確性。我們的貢獻如下：1.結合歷史的股價、股價之間的變化、社群平台情緒三種資訊來進行股票趨勢預測 2.提出一個多項特徵結合圖神經網路的融合技術。

## 2 相關研究

本章節介紹股票趨勢預測的相關研究，主要分為基於股價、社群平台、以及結合兩者資訊的研究，並探討不同特徵下使用的遞歸神經網路和圖神經網路等模型架構技術。

### 2.1 基於序列式神經網路模型股票趨勢預測

2017 年由 Selvin 等人 (Selvin et al., 2017) 使用每日收盤價預測單間公司的股票走勢，他們使用三種不同的深度學習模型 RNN、CNN、LSTM 來進行訓練與預測。2020 年 Lu 等人 (Lu et al., 2021) 提出 CNN-BiLSTM-AM 方法，以歷史價格、成交量、漲跌變化作為輸入，由卷積神經網路(CNN)、雙向長短期記憶神經網路(BiLSTM)和注意力機制(AM)組合，CNN 用於提取輸入數據的特徵，BiLSTM 用於學習和預測提取的特徵數據，而 AM 用於捕捉過去不同時間的特徵對第二天收盤價的影響。2024 年 LI 等人 (Li et al., 2024) 提出 MASTER 模型利用市場訊息進行自動特徵選擇，以適應市場的動態變化。MASTER 能夠有效處理複雜的時間序列資料，提高股價預測的準確性。

### 2.2 基於社群平台言論股票趨勢預測

Nguyen 等人 (Nguyen & Shirai, 2015)於 2015 年提出 TSLDA 模型，透過社群平台上的言論進行情緒分析，並將其輸入至主題模型以捕捉相關主題，用於預測股價走勢。2018 年由 Hu 等人 (Hu et al., 2018)提出一個多層注意力網路模型，將單日的發文輸入至注意力網路以獲得其表示形式，然後將多日的發文表示形式輸入至序列式神經網路模型與注意力網路，用於預測股票趨勢。同年 Xu 等人 (Xu & Cohen, 2018) 提出 StockNet 模型，同時結合了股價特徵與 tweets 文本特徵，其中股價特徵使用序列式神經網路模型，而文本特徵亦是使用多層注意力網路進行編碼，並且比較了五種模型組合的表現。2024 年 Fan 等人 (Fan & Shen, 2024) 提出一種基於多層感知機（MLP）的簡單而強大的股價預測架構。該模型通過三種混合機制有效地捕捉股票數據中的複雜相關性。同年，Ranjith 等人 (Ranjith, 2024) 提出了一種新穎的可解釋人工智慧（XAI）模型，該模型結合了多種數據源，包括社交媒體情緒和技術指標，以預測股票市場趨勢並提供解釋性結果。

### 2.3 基於圖神經網路進行股票趨勢預測

近年許多研究提出使用圖神經網路來學習股票間複雜的關聯性，Chen 等人 (Chen et al., 2018) 在 2018 年提出了一種基於 LSTM+GCN 的模型，該模型使用歷史股價作為節點表示形式，並通過圖卷積神經網路（GCN）來學習股票之間的關係，進行訓練以預測股票趨勢。2019 年 Kim 等人 (Kim et al., 2019b)提出 HATS 模型，該模型將每檔股票的歷史股價作為輸入，並通過圖注意力機制聚合不同關係類型的股票表示形式。HATS 不僅用應用於預測個股價格，還可用於預測市場指數的走勢。2020 年由 Wei Li 等人 (Li et al., 2021)提出一種基於 LSTM-RGCN 模型用於股票趨勢預測。由於股票在休市期間沒有交易操作，休市後的新聞往往會對下個開盤日的股市產生影響。因此，該研究根據股市收盤期間的新聞來預測下個開盤日股票走勢。該模型結合了財經

新聞文本與歷史股價當作為特徵，使用 LSTM 學習，並通過 GCN 捕捉學習股票之間的依賴性，最終預測下個開盤日股票走勢。2020 年由 Sawhney 等人 (Sawhney et al., 2020)提出一種名為 MAN-SF 的模型架構，從歷史股價、社群平台、股票間關係中進行聯合學習，具體而言，模型通過 GRU 捕捉歷史股價的數值型時間序列關係與社群平台的文本上下文關係，並進一步通過圖神經網路學習股票之間的相互關係。該模型在真實世界的股市數據上進行了實驗，驗證了其有效性。

由於股票預測具有高度複雜性，部分研究選擇使用超圖神經網路模型來學習，2021 年 Sawhney (Sawhney et al., 2021)等人提出 STHAN-SR，一種用於選擇股票的模型。該模型通過 LSTM 與注意力機制的組合來提取股價的時間序列特徵，並使用圖模型學習股票之間的複雜依賴關係。最終，根據每檔股票的預期利潤對其進行排名。

## 3 方法

本研究提出的架構分為三個部分，包括特徵提取（Feature Extraction）、模型訓練（Model Training）和分類（Classification）。特徵提取負責擷取數據特徵與情緒特徵；模型訓練部分結合序列式神經網路模型與圖神經網路模型，針對不同性質的特徵進行相應的模型訓練；最後，分類部分將模型輸出的向量進行分類處理。架構圖如圖 1 所示。

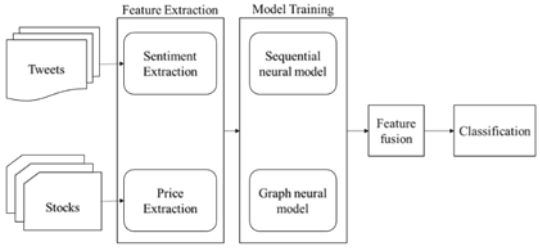



圖 1.系統架構圖

### 3.1 特徵擷取

特徵擷取分成股價特徵跟情緒特徵兩個部分。

#### 3.1.1 股價特徵擷取

股價特徵由每檔股票每日的日期、開盤價、收盤價、最高價、最低價及成交量構成。我們使用還原 K 線的收盤價來計算股價的漲跌情況。定義如下：若當天 $t$ 的收盤價大於或等於前一天 $t$-1 的收盤價，則當天 $t$ 被定義為「漲」；若當天 $t$ 的收盤價小於前一天 $t$-1 的收盤價，則當天 $t$ 被定義為「跌」，如公式 (1) 所示。公式中的 $p_t^{ac}$ 代表某檔股票在第 $t$ 天的股價，其中 ac 表示還原 K 線的收盤價。此時，若 $y_t = 0$，則表示第 $t$ 天股票下跌；若 $y_t = 1$，則表示第 $t$ 天股票上漲。

$$y_t = \begin{cases} 0, p_t^{ac} < p_{t-1}^{ac} \\ 1, p_t^{ac} \geq p_{t-1}^{ac} \end{cases} \quad (1)$$

根據 Sawhney 等人的研究，決定股票趨勢的關鍵在於價格變化，而非價格本身。雖然每日的股價確實具有重要意義，但真正引發股勢變化的是歷史股價的價差。因此，我們採用由開盤價、收盤價、最高價和最低價計算出的四種價差值作為特徵，具體如公式 (2) 所示。

$$c_t = p_t - p_{t-1} \quad (2)$$

公式中的 $p_t$ 代表某檔股票在第 $t$ 天的股價，將其與前一天 $t$-1 的股價相減，便可得出價差 $c_t$，相同的作法，使用開盤價得到開盤價價差 $c_t^o$、使用最高價得到最高價價差 $c_t^h$、使用收盤價得到收盤價價差 $c_t^c$、使用最低價得到最低價價差 $c_t^l$。此外，觀察股勢變化的另一重要指標是移動平均線，其目的是通過計算過去價格的平均值來平滑化價格波動，使價格走勢更為明確。因此，我們使用收盤價來計算五天、十天、二十天及六十天的移動平均，如公式 (3) 所示。

$$m_t^T = \frac{\sum_{i=t-T+1}^{t} p_i^c}{T} \quad (3)$$

公式中的 $p_i^c$ 代表某檔股票在第 $i$ 天的收盤價，$T$ 為觀察的過去天數。通過將過去 $T$ 天的收盤價總和除以 $T$，即可得到第 $t$ 天的 $T$ 天移動平均值 $m_t^T$。而五天的移動平均值表示為 $m_t^5$，十天的移動平均值為 $m_t^{10}$，二十天的移動平均值為 $m_t^{20}$，而六十天的移動平均值則為 $m_t^{60}$。

#### 3.1.2 情緒特徵擷取

首先，我們對原始推文移除 URL 及停用字，然後在文本中出現的股票代碼與股票名稱處加上[ASP]目標標籤，接著，我們使用 Feng 等人 (Feng et al., 2019) 提出的 ABSC 模型進行情緒辨識。該模型並不是將整篇推文進行整體情緒評估，而是針對推文中每一個目標標籤進行情緒分類，情緒類別包括負面（Negative）、中性（Neutral）和正面（Positive）。我們利用 Word2vec (Mikolov,

2013) 將這三種情緒類別轉換為嵌入向量（embedding），作為情緒特徵向量。資料集中所使用的股票均屬於高討論度金融商品，因此，我們將單一股票在一天內的推文情緒特徵向量相加並取平均，作為該天的情緒特徵向量。此外，為了讓情緒較為明顯的向量擁有較大的影響力，我們對負面情緒和正面情緒向量賦予不同的權重，如公式 (4) 所示。

$$e_t = e^{Positive} \times \frac{n_t^{Positive}}{n_t^{Positive} + n_t^{Negative}} + e^{Negative} \times \frac{n_t^{Negative}}{n_t^{Positive} + n_t^{Negative}} \quad (4)$$

$e^{Positive}$ 代表正面情緒向量，$e^{Negative}$ 代表負面情緒向量，$n_t^{Positive}$ 為某檔股票在第 $t$ 天的正面情緒發文數。$n_t^{Negative}$ 為某檔股票在第 t 天的負面情緒發文數。通過此公式進行加權相加，我們可以得出該檔股票在第 $t$ 天的情緒特徵向量 $e_t$。

## 3.2 架構

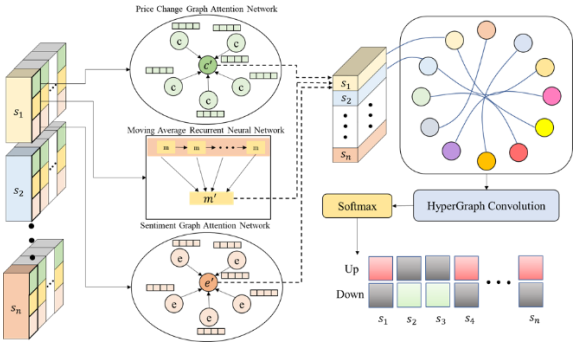



圖 2. CME-HG 模型架構圖

如圖 2 所示，我們提出了一種名為 CME-HG 的預測模型。並定義每檔股票 $s \in S = \{s_1, s_2, \ldots, s_n\}$，其中，對於每檔股票 $s_i$，所有天數的特徵按以每 $T$ 天劃分為一個資料樣本。每個資料樣本包含三種類型的特徵：

1. 價差特徵 $C_t = \{c_{t-T+1}, c_{t-T+2}, \ldots, c_t\}$：特徵由每日的價格變動計算得出，旨在捕捉價格波動情況，進而預測未來趨勢。
2. 移動平均特徵 $M_t = \{m_{t-T+1}, m_{t-T+2}, \ldots, m_t\}$ 此特徵通過移動平均來平滑價格變化，以觀察短期和長期趨勢變化。
3. 情緒特徵 $E_t = \{e_{t-T+1}, e_{t-T+2}, \ldots, e_t\}$：此特徵來自於社交媒體情緒的變動。藉以捕捉情緒變化對股價的影響。

這三種特徵通過各自專屬的模型訓練方式進行處理，然後再進行融合，以提升預測的準確性。

### 3.2.1 個股的價格變化特徵

我們定義了一個「Price Change-to-Price Change」同構圖 $G = (V, E)$，其中 $V$ 和 $E$ 分別代表圖的節點與邊。節點 $V$ 由第 $t$ 天的價差特徵 $c_t$ 組成。而價差特徵 $c_t$ 包含四個部分：$[c_t^o, c_t^h, c_t^c, c_t^l]$，分別代表開盤價、最高價、收盤價和最低價的價差，這些組成作為節點的初始向量。為了得到過去一段時間的平均價差值，我們將過去 $T$ 天內的價差特徵中的各項分別相加並取平均，得到一個平均價差特徵向量 $c_{t+1}$ 作為新的節點初始向量。這樣的平均處理能夠平滑化數值波動，使數據更具穩定性，如公式 (5) 所示。

$$c_{t+1} = \frac{\sum_{i=t-T+1}^{t} [c_i^o, c_i^h, c_i^c, c_i^l]}{T} \quad (5)$$

每一天的價差特徵向量包括四個值，分別是根據開盤價計算得到的 $c^o$、最高價計算得到的 $c^h$、收盤價計算得到的 $c^c$ 以及最低價計算得到的 $c^l$。我們將過去 $T$ 天中這四個價差值分別相加並取平均，從而得出第 $t$+1 天的平均價差特徵向量 $c_{t+1}$。

邊 $E$ 由 $E_{c_i c_j}$ 表示，具體來說，$E_{c_t c_{t+1}}$ 表示第 $t$ 天的價差特徵 $c_t$ 與第 $t$+1 天的平均價差特徵向量 $c_{t+1}$ 之間的關係。這用來表示過去的價格波動可能會影響到未來投資者的決策。因此，我們將過去 $T$ 天的價差特徵與第 $t$+1 天的價差特徵逐一建立邊，如圖 3 所示。

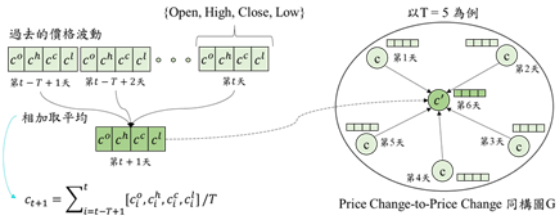



圖 3. Price Change-to-Price Change 同構圖 G 說明

我們將有向圖 $G$ 作為輸入，並使用圖注意力機制網路模型（Graph Attention Network, GAT），通過引入注意力機制（attention mechanism）。對於圖中的一對節點 $(i, j)$，，我們使用自注意力機制來學習注意力係數 $e_{ij}$，該係數表示節點 $i$ 對節點 $j$ 的重要性，如公式 (6) 所示

$$e_{ij} = attention(Wc_i, Wc_j), c_i, c_j \in C_t \quad (6)$$

我們將 $c_i$ 與 $c_j$ 進行線性轉換，然後將兩者串接起來，接著乘以上轉置的權重向量 $a^T$。隨後，結果會輸入至 LeakyReLU 激活函數，並通過 softmax 函數進行正規化，從而得到注意

力權重。$N_i$代表節點 $i$ 的鄰居節點，這裡指的是該節點前五天的節點（即 $i$-5, …, $i$-1）。當計算出每一對節點之間的權重 $\alpha_{i,j}$ 後，我們將所有鄰居節點的特徵乘上相應的權重，再加總得到更新後的節點特徵表示 $c'_{t+1}$，如公式 (7) 所示。

$$c'_{t+1} = \sigma(\sum_{j0 \in \mathcal{N}_i} \alpha_{i,j} W c_j) \tag{7}$$

最後，我們採用多頭注意力機制來學習更穩定的嵌入表示（embedding representation）。如公式 (8) 所示，我們將公式 (7) 的轉換過程執行 $k$ 次，並將每次輸出的表示 $c'_{t+1}$ 串接起來，從而獲得最終的輸出表示 $c'_{t+1}$。這樣的多頭注意力機制能夠捕捉更多樣的關係和特徵，從而提高模型的表現穩定性和泛化能力。

$$c'_{t+1} = \Big\Vert_{k=1}^{k} \sigma(\sum_{j \in \mathcal{N}_i} \alpha_{i,j}^{k} W^k c_j) \tag{8}$$

### 3.2.2 個股移動平均序列式特徵

我們將過去 $T$ 天的移動平均特徵作為輸入，並使用門控循環單元（GRU）來提取移動平均線的時間序列特徵。第 $t$ 天的 GRU 輸出表示如公式 (9) 所示。這樣可以有效捕捉移動平均線在時間序列中的動態變化，進而提升對股價趨勢的預測能力。

$$h_t = \mathrm{GRU}(m_t, h_{t-1}) \tag{9}$$

我們定義第 $t$ 天的移動平均特徵為$m_t$，由$m_t^5, m_t^{10}, m_t^{20}, m_t^{60}$組成，其中 $m_t^5$代表五日線，$m_t^{10}$代表十日線，$m_t^{20}$代表二十日線，$m_t^{60}$代表六十日線。為了更有效地捕捉不同天數移動平均線對股價趨勢的影響，我們使用時間注意力機制來學習不同天數之間的移動平均特徵對預測結果的影響權重，並對 GRU 的所有隱藏層特徵進行聚合。權重的計算方式如公式 (10) 所示。

$$\alpha_i = \frac{\exp(h_i^T W \bar{h}_z)}{\sum_{i=1}^{T} \exp(h_i^T W \bar{h}_z)} \tag{10}$$

其中，$\bar{h}_z$代表 GRU 的過去隱藏層表示，$\alpha_i$ 為第 $i$ 天的注意力權重，$W$ 為一個可學習的參數矩陣。移動平均線將被輸入至 GRU 和時間注意力機制，具體結構如圖 4 所示。這樣的設計能夠充分利用時間序列的特徵，進一步提高模型對股價趨勢預測的準確性。

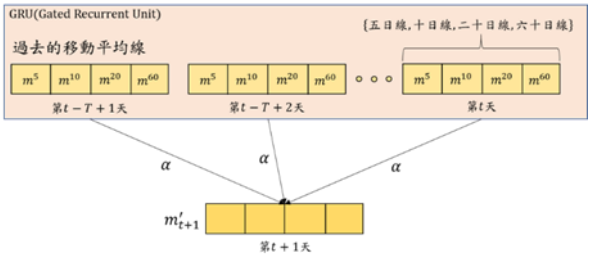



圖 4. 移動平均線輸入 GRU 訓練示意圖

### 3.2.3 個股評論情緒特徵學習

情緒特徵的處理方法與價差特徵的訓練方式相似。我們定義 sentiment-to-sentiment 同構圖 $G$=($V$,$E$)，其中 $V$ 和 $E$ 分別代表圖的節點和邊。節點 $V$ 由第 $t$ 天的情緒特徵 $e_t$組成，而情緒特徵 $e_t$由使用 word2vec 模型轉換而成的向量組成，這些向量作為節點的初始表示。為了獲得過去一段時間的平均情緒特徵向量，我們將過去 $T$ 天的情緒特徵向量$e_i$相加並取平均，從而得到第 $t$+1 天的平均情緒特徵向量 $e_{t+1}$，該向量將作為節點的初始向量，如公式 (11) 所示。

$$e_{t+1} = \frac{\sum_{i=t-T+1}^{t} e_i}{T} \tag{11}$$

邊 $E$ 由 $E_{e_i e_j}$表示，其中$E_{e_t e_{t+1}}$描述第 $t$ 天的情緒特徵$e_{t+1}$與第 $t$+1 天的平均情緒特徵向量 $e_{t+1}$ 之間的關係，如圖 5 所示。

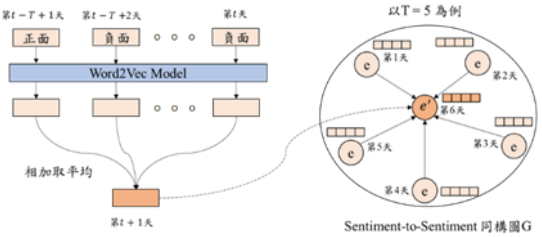



圖 5. Sentiment-to-Sentiment 同構圖 G 說明

### 3.2.4 特徵融合圖神經網路

定義特徵融合圖 $G$=($V$,$E$)，其中 $V$ 和 $E$ 分別作為圖的節點與邊。節點 $V$ 由第 $t$+1 天的價差特徵$c'_{t+1}$、移動平均特徵 $m'_{t+1}$、情緒特徵 $e'_{t+1}$組成，此外還包含一個融合節點 $f_{t+1}$。該融合節點$f_{t+1}$是通過將上述三個特徵向量相加得到的。邊 $E$ 則連接這三個特徵向量節點與融合節點$f_{t+1}$，用以表示不同性質的特徵與融合節點之間的關係，如圖 6 所示。

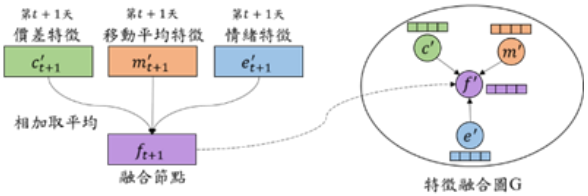


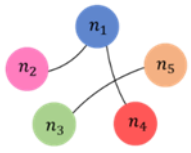

圖 6. 特徵融合圖說明

接著，我們將特徵融合圖 G 輸入至圖注意力網路（Graph Attention Network, GAT）進行訓練。訓練完成後，從網路中提取融合向量 $f'_{t+1}$，該向量作為每檔股票在不同性質特徵融合後的股票表示形式。

### 3.2.5 不同股票間關係特徵

我們建構 stock-to-stock 的同構圖，其中節點為每檔股票的表示形式，邊則依據 Feng 等人 (Feng et al., 2019) 提出兩種關係來構建不同股票之間的關聯，第一種關係是**行業關係**，根據全球行業分類標準 (Global Industry Classification Standard, GICS)，我們將屬於同一行業的股票進行連邊。在我們研究的 85 檔股票中，共構建了 16 種行業關聯。例如，Computer Software 這一行業包含 Google 和 Facebook，這兩家公司作為同一行業的股票因此相連。第二種關係是基於 **Wikidata** 提供的股票資訊。例如，Alphabet Inc. 是 Google LLC 的母公司，而 Microsoft 與 Branded Entertainment Network 由共同的董事 Bill Gates 連接。在 85 檔股票中，共構建了 70 種這樣的關聯。通過這兩種關係，我們最終構建了 86 種股票之間的關聯，具體示意圖如圖 7 所示。

e1: Computer Software這一行業
e2: Alphabet Inc.是Google LLC的母公司
e3: Microsoft及Branded Entertainment Network則有共同的董事Bill Gates




| | $e_1$ | $e_2$ | $e_3$ |
|---|---|---|---|
| $n_1$(google) | 1 | 1 | 0 |
| $n_2$(Alphabet) | 0 | 1 | 0 |
| $n_3$(Branded) | 0 | 0 | 1 |
| $n_4$(Facebook) | 1 | 0 | 0 |
| $n_6$(Microsoft) | 0 | 0 | 1 |

圖 7. HyperGraph 建圖示意圖

由於我們所建構的關聯來自不同概念，包括同一行業的關聯、母公司與子公司的關聯、以及同一董事的關聯，這些關聯的複雜性較高。基於此，我們採用 Feng 等人 (Feng et al., 2019) 提出的 HyperGraph Neural Networks 來處理這種更為複雜的高階關聯性，並學習股票的表示形式。如公式 (12) 所示，$H$ 代表 85 檔股票所構建的 86 種關聯的矩陣，$f^{(l)}$ 為第 $l$ 層的股票表示形式：

$$f^{(l+1)} = \sigma\left(D_v^{-\frac{1}{2}} H W D_e^{-1} H^T D_v^{-\frac{1}{2}} f^{(l)} \Theta^{(l)}\right) \quad (12)$$

經過超圖神經網路的訓練，我們獲得每檔股票的二維向量表示，並通過 softmax 函數對結果進行正規化，從而得到預測結果 $p_t^k$，表示第 $k$ 檔股票在第 $t$ 天的預測結果。為了計算分類結果的誤差，我們使用交叉熵損失函數 (cross entropy)，如公式 (13) 所示，其中 $n$ 為股票的數量，$y_t^k$ 為第 $k$ 檔股票在第 $t$ 天的實際漲跌結果：

$$\mathcal{L} = \sum_{k=1}^{n} -[y_t^k * \log(p_t^k) + (1 - y_t^k) * \log(1 - p_t^k)] \quad (13)$$

最後，我們對所有股票在所有測試天數內的預測結果進行加總並取平均，從而得到最終的預測結果。

## 4 實驗

我們將在以下章節詳細描述實驗參數與結果。

### 4.1 資料集

本文使用了由 Yum 等人 (Xu & Cohen, 2018) 在 2018 年提供的結合股票與推文進行股票預測的數據集。該數據集包含 88 檔具有高討論度的股票，均來自美國的 S&P 500 和 NASDAQ 指數，數據範圍為 2014 年 1 月 1 日至 2016 年 1 月 1 日。扣除非交易日後，總共有 504 天的交易記錄，88 檔股票共計 41,990 筆數據，其中股價上漲的筆數為 21,679，股價下跌的筆數為 20,311，漲跌比例分別為 52% 和 48%。此外，該數據集中還包含了 2014 年 1 月 1 日至 2016 年 1 月 1 日期間提及這些股票的推文，共計收集了 106,338 篇推文。由於 BABA (Alibaba Group Holding)、AGFS (AgroFresh Solutions Inc.) 和 GMRE (Global Medical REIT) 這三檔股票沒有提供相關推文，因此我們將其從數據集中移除。數據集的劃分比例為 7:1:2，具體的日期範圍如表 1 所示。

| | Training | Validation | Testing |
|---|---|---|---|
| 日期 | 2014/1/1-<br>2015/7/31 | 2015/8/1-<br>2015/9/30 | 2015/10/1-<br>2016/1/ |

表 1.資料集日期

### 4.2 實驗參數

實驗環境基於 Ubuntu 18.04 操作系統，運行於具有 32GB 記憶體的主機上，顯卡為 GeForce RTX 2080 Ti（11GB）。實驗使用 Python 3.8.3 和 PyTorch 1.10.2+cu102 進行。實驗中，價差特徵與情緒特徵的計算均使用 GAT，隱藏層向量的維度為 4，Attention head 設為 1。Word2vec 的 embedding 維度設為 300。移動平均特徵則是通過 GRU 模型計算，隱藏層向量的維度同樣為 4。CME-HG 模型使用的損失函數為 Cross Entropy Loss，dropout 設定為 0.38，Optimizer 為 Adam(Diederik, 2014)，Learning Rate 設為$5 \times 10^{-5}$，最大訓練 epoch 數為 200。

### 4.3 評估指標

實驗評估指標採用兩種評估方式，分別為 Accuracy 與 F1-Score。這兩項指標能夠有效衡量模型在分類任務中的表現，其中 Accuracy 用於評估模型的整體正確率，F1-Score 則綜合考慮了 Precision 與 Recall。

### 4.4 比較模型

為了驗證所提出模型的效能，我們參考 MAN-SF (Sawhney et al., 2021) 的方法，使用以下模型作為比較對象：

- ARIMA (Autoregressive Integrated Moving Average model) (Brown, 2004)：：也稱為整合移動平均自我迴歸模型，這是一種時間序列預測方法，使用歷史股價進行預測。
- Selvin et al. (Yang et al., 2021)：分別使用三種深度學習模型 RNN、CNN 和 LSTM。我們選擇表現最佳的 LSTM 作為比較模型。
- RandForest (Breiman, 2001)：將推文文本經由 word2vec 轉換後，使用 Random Forests classifier 進行訓練與分類。
- TSLDA (Nguyen & Shirai, 2015)：使用情緒分析和主題模型來分析社群媒體言論，並進行股票趨勢預測。
- HAN (Hu et al., 2018)：利用分層注意力機制，對單日與多日的 tweets 文本進行編碼，以預測股票趨勢。
- StockNet (Xu & Cohen, 2018)：使用股價與 tweet 文本特徵，通過分層注意力對單日與多日的 tweet 進行編碼，並使用序列式神經網路模型處理價格特徵。
- Chen et al. (Chen et al., 2018)：該模型使用歷史股價作為節點表示形式，並通過 GCN 整合公司之間的關係進行股票趨勢預測。
- HATS (Kim et al., 2019a)：該模型使用多檔股票的歷史價格特徵作為節點表示形式，並通過圖注意力機制結合每檔股票之間的關係進行股票趨勢預測。

### 4.5 實驗結果

實驗結果如表 2 所示。表格中展示了不同模型的比較結果，包括準確率（Accuracy）、F1-Score 等性能指標。

| Method | Model | Accuracy | F1-score |
|---|---|---|---|
| Regression | ARIMA | 0.514 | 0.513 |
| RNN | Selvin et al | 0.530 | 0.529 |
| Social Media | RandFores | 0.531 | 0.527 |
| | TSLDA | 0.541 | 0.539 |
| | HAN | 0.576 | 0.572 |
| RNN+ Attention | StockNet | 0.550 | 0.546 |
| Graph | Chen et al | 0.532 | 0.530 |
| | HATS | 0.562 | 0.559 |
| Our Method | **CME-HG** | **0.604** | **0.583** |

表 2.實驗結果比較

與傳統的回歸模型相比，我們提出的 CME-HG 模型在 Accuracy 和 F1-score 上分別高出約 7% 和 9%。相比於序列式神經網路模型的方法，CME-HG 模型的效果分別提升了約 5% 和 7%。針對使用社群平台言論作為特徵的三種方法，CME-HG 模型相較於 RandForest 與 TSLDA 分別高出約 7% 和 5%，並領先 HAN 模型約 1% 和 2%。而相比於 StockNet 模型（同時結合股價與社群言論特徵，並使用序列式神經網路和注意力機制），CME-HG 模型的準確率和 F1-score 分別高出約 5% 和 3%。此外，針對使用圖神經網路來學習股票間關係的兩種模型，CME-HG 模型比 Chen 等人的研究結果高出約 7% 和 5%，領先 HATS 模型約 2% 和 4%。實驗結果表明，我們的 CME-HG 模型在股票趨勢預測中表現出卓越的效果。

## 5 分析和討論

我們在此章節分析了不同的特徵，和短、中、長期資料對模型的影響，以及對個案的分析。

### 5.1 不同特徵的比較

為了更清楚地了解哪些特徵對於股價漲跌的影響，我們進行了單一特徵和兩兩特徵組合的模型訓練，並對比分析了三種特徵對模型表現的重要性。實驗結果如表 3 所示。

| 平均特徵 | 價差特徵 | 情緒特徵 | Accuracy | F1-score |
|---|---|---|---|---|
| o | x | x | 0.53 | 0.606 |
| x | o | x | 0.581 | 0.606 |
| x | x | o | 0.523 | **0.641** |
| o | o | x | **0.604** | 0.576 |
| o | x | o | 0.531 | 0.622 |
| x | o | o | 0.602 | 0.563 |
| o | o | o | **0.604** | 0.583 |

表 3.不同特徵的效能比較

從這項實驗結果中，我們觀察到兩個特別值得注意的現象。首先，當只使用情緒特徵作為唯一輸入時，模型的 Accuracy 表現最低，僅為 0.523，而 F1-Score 卻達到了最高，為 0.641。為了進一步了解這一現象，我們計算了實驗中的 Recall 和 Precision，分別為 1 和 0.522。這表明模型在預測中所有結果均被判定為「漲」，顯示單單依賴情緒特徵並不能有效地預測股價趨勢，導致分類結果嚴重失衡。其次，透過不同特徵組合的實驗結果，我們發現如果只使用價差特徵（如第二項實驗），或者價差特徵與其他特徵搭配使用（如第四項和第六項實驗），模型的預測表現均較高。這說明價差特徵對於模型的準確預測具有較強的影響力。最後，從第七項實驗可以看到，當三種不同性質的特徵都作為輸入特徵時，模型表現最佳，表明多樣化的特徵融合能夠最大限度地提升模型的預測效果。

### 5.2 短、中、長期資料的比較

我們嘗試將資料按不同的天數進行取樣，分別是每 1 天（單日）、每 2 天、每 3 天、每 5 天、每 6 天、每 7 天、每 8 天、每 9 天、每 10 天（雙周）和每 20 天（單月）作為資料樣本，以探討模型在學習短期、中期和長期資料時，對預測表現的影響。實驗結果如圖 8 所示，展示了不同時間取樣間隔下模型的預測效果。

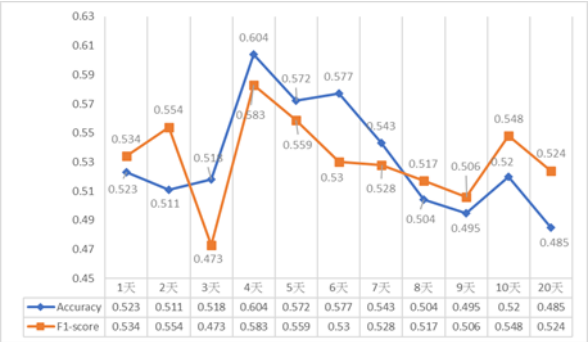



圖 8. 不同時間長度之實驗結果

從實驗結果得知，當模型在訓練中使用 4 天的資料進行學習時，預測表現最佳。隨著取樣天數的增加，模型表現逐漸下降。這是因為隨著時間跨度的增加，中、長期的數據特徵對股價的影響程度降低。此外，社群發文中的情緒隨著時間推移變得模糊，增加了噪聲和干擾，進而導致模型效果下降。

### 5.3 CASE STUDY

由 CME-HG 模型預測的 85 檔股票中，每檔股票的預測表現不一。我們針對其中預測表現最差的股票——準確率僅為 0.48 的 AAPL（APPLE）進行了深入分析，如圖 9 所示。此圖為 AAPL（APPLE）從 2015 年 10 月 1 日到 2016 年 1 月 1 日的收盤價折線圖。圖中的 x 軸為日期，y 軸為收盤價。紅色點代表模型在當天的預測結果與實際結果相同，即模型預測正確；藍色點代表模型在當天的預測結果與實際結果不同，也就是模型預測錯誤。可以看出，AAPL 股票在 2015 年 11 月 3 日之後，股價從 122 美元快速下跌至約 106 美元，導致模型在這段期間的預測表現較差。經過檢視推文內容後發現，這段期間股價波動較大，發文者對於股價走勢的看法存在多空分歧，這種情緒不一致性導致了模型預測的偏差，從而顯著降低了準確率。這表明在股價波動劇烈且市場情緒不穩定的情況下，模型的預測能力會受到較大影響。

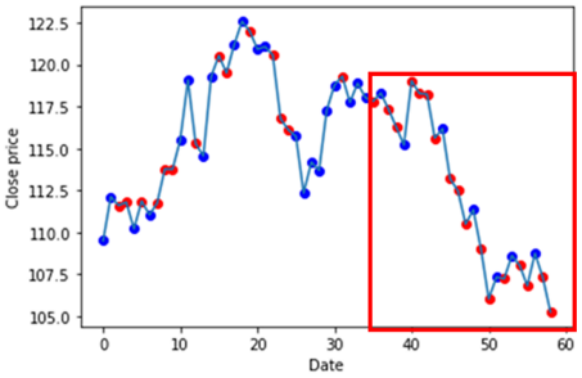



圖 9. AAPL(APPLE)收盤價折線圖(2015/10/01-2016/01/01)

## 6 結論與未來展望

上市公司股票趨勢受到歷史價格、技術分析、網路言論和股票之間複雜關聯性的影響。本文提出了一種結合歷史股價與社群情緒的多模態整合模型，用於預測股票的漲跌走勢。該模型利用三種不同的特徵來預測股票的漲跌，包括基於收盤價計算的移動平均線、基於開盤價、最高價、最低價與收盤價計算的股價變動，以及社群言論的情緒分析。移動平均線使用 GRU 提取序列特徵，短期股價變化藉由圖注意力機制獲取價差特徵，社群情緒特徵也通過圖注意力機制得出。我們將三種不同性質的特徵融合為股票的表示形式，最終通過超圖神經網路預測股票的漲跌。

實驗結果顯示，我們所提出的模型在結合多種不同性質的特徵並考慮股票之間的關聯後，Accuracy 達到 0.604。相比先前的方法，我們的模型更有效地捕捉了股票的漲跌趨勢。並且透過實驗可以發現，在這項預測股票趨勢的任務中，價差特徵對模型表現的影響最大，其次是移動平均特徵，最後是情緒特徵。

未來的研究將著重於三個方面的改進。首先，在特徵選取方面，我們計劃引入財經新聞的語義分析作為額外特徵，以進一步提高模型在股票趨勢預測中的準確率。其次，針對情緒特徵，由於本研究中未能精確分類出針對個股的正面與負面情緒，導致情緒特徵中存在噪聲，我們計劃針對這一問題進行更精細的處理，以提升模型的預測準確性。最後，面對模型潛在的過度擬合以及在現實世界中的適用性，我們將考慮使用 regularization techniques 和更多的真實市場分析來進一步改善。

## Acknowledgments

The authors would like to thank the supports from the National Science and Technology Council, Taiwan under the grant numbers: NSTC113-2221-E-027-096, and NSTC113-2634-F-027-001-MBK.

## References

Amin, M. S., Ayon, E. H., Ghosh, B. P., MD, M. S. C., Bhuiyan, M. S., Jewel, R. M., & Linkon, A. A. (2024). Harmonizing Macro-Financial Factors and Twitter Sentiment Analysis in Forecasting Stock Market Trends. *Journal of Computer Science and Technology Studies*, *6*(1), 58-67.

Asgarov, A. (2023). Predicting Financial Market Trends using Time Series Analysis and Natural Language Processing. *arXiv preprint arXiv:2309.00136*.

Behrendt, S., Schmidt, A. J. J. o. B., & Finance. (2018). The Twitter myth revisited: Intraday investor sentiment, Twitter activity and individual-level stock return volatility. *96*, 355-367.

Ben Cheikh, S., Amiri, H., & Loukil, N. J. I. J. o. S. E. (2024). Social media investors' sentiment as stock market performance predictor. *51*(6), 713-724.

Breiman, L. (2001). Random forests. *Machine learning*, *45*, 5-32.

Brown, R. G. (2004). *Smoothing, forecasting and prediction of discrete time series*. Courier Corporation.

Chen, Y., Wei, Z., & Huang, X. (2018). Incorporating corporation relationship via graph convolutional neural networks for stock price prediction. Proceedings of the 27th ACM international conference on information and knowledge management,

Diederik, P. K. J. (2014). Adam: A method for stochastic optimization.

Duan, Y., Wang, L., Zhang, Q., & Li, J. (2022). Factorvae: A probabilistic dynamic factor model based on variational autoencoder for predicting cross-sectional stock returns. Proceedings of the AAAI Conference on Artificial Intelligence,

Fan, J., & Shen, Y. (2024). StockMixer: A Simple Yet Strong MLP-Based Architecture for Stock Price Forecasting. Proceedings of the AAAI Conference on Artificial Intelligence,

Feng, F., He, X., Wang, X., Luo, C., Liu, Y., & Chua, T.-S. (2019). Temporal relational ranking for stock prediction. *ACM Transactions on Information Systems (TOIS)*, *37*(2), 1-30.

Feng, Y., You, H., Zhang, Z., Ji, R., & Gao, Y. (2019). Hypergraph neural networks. Proceedings of the AAAI conference on artificial intelligence,

Hu, Z., Liu, W., Bian, J., Liu, X., & Liu, T.-Y. (2018). Listening to chaotic whispers: A deep learning framework for news-oriented stock trend prediction. Proceedings of the eleventh ACM international conference on web search and data mining,

Kim, R., So, C. H., Jeong, M., Lee, S., Kim, J., & Kang, J. (2019a). Hats: A hierarchical graph attention network for stock movement prediction. *arXiv preprint arXiv:1908.07999*.

Kim, R., So, C. H., Jeong, M., Lee, S., Kim, J., & Kang, J. J. a. p. a. (2019b). Hats: A hierarchical graph attention network for stock movement prediction.

Li, T., Liu, Z., Shen, Y., Wang, X., Chen, H., & Huang, S. (2024). MASTER: Market-Guided Stock Transformer for Stock Price Forecasting. Proceedings of the AAAI Conference on Artificial Intelligence,

Li, W., Bao, R., Harimoto, K., Chen, D., Xu, J., & Su, Q. (2021). Modeling the stock relation with graph network for overnight stock movement prediction. Proceedings of the twenty-ninth international conference on international joint conferences on artificial intelligence,

Lu, W., Li, J., Wang, J., Qin, L. J. N. C., & Applications. (2021). A CNN-BiLSTM-AM method for stock price prediction. *33*(10), 4741-4753.

Mikolov, T. J. a. p. a. (2013). Efficient estimation of word representations in vector space.

Nguyen, T. H., & Shirai, K. (2015). Topic modeling based sentiment analysis on social media for stock market prediction. Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume 1: Long Papers),

Nuñez-Mora, J. A., Mendoza-Urdiales, R. A. J. S. N. A., & Mining. (2023). Social sentiment and impact in US equity market: an automated approach. *13*(1), 111.

Ranjith, J. (2024). Enhancing Stock Market Trend Prediction Using Explainable Artificial Intelligence and Multi-source Data. *Fusion: Practice and Applications*, *16*(2), 178-178-189.

Sawhney, R., Agarwal, S., Wadhwa, A., Derr, T., & Shah, R. R. (2021). Stock selection via spatiotemporal hypergraph attention network: A learning to rank approach. Proceedings of the AAAI Conference on Artificial Intelligence,

Sawhney, R., Agarwal, S., Wadhwa, A., & Shah, R. (2020). Deep attentive learning for stock movement prediction from social media text and company correlations. Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP),

Selvin, S., Vinayakumar, R., Gopalakrishnan, E., Menon, V. K., & Soman, K. (2017). Stock price prediction using LSTM, RNN and CNN-sliding window model. 2017 international conference on advances in computing, communications and informatics (icacci),

Tabari, N., Biswas, P., Praneeth, B., Seyeditabari, A., Hadzikadic, M., & Zadrozny, W. (2018). Causality analysis of Twitter sentiments and stock market returns. Proceedings of the first workshop on economics and natural language processing,

Wang, S., Bai, Y., Ji, T., Fu, K., Wang, L., & Lu, C.-T. (2023). Stock Movement and Volatility Prediction from Tweets, Macroeconomic Factors and Historical Prices. 2023 IEEE International Conference on Big Data (BigData),

Xu, Y., & Cohen, S. B. (2018). Stock movement prediction from tweets and historical prices. Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers),

Yang, H., Zeng, B., Xu, M., & Wang, T. (2021). Back to reality: Leveraging pattern-driven modeling to enable affordable sentiment dependency learning. *arXiv preprint arXiv:2110.08604*.

Zhang, L., Aggarwal, C., & Qi, G.-J. (2017). Stock price prediction via discovering multi-frequency trading patterns. Proceedings of the 23rd ACM SIGKDD international conference on knowledge discovery and data mining,