

結合情感詞與階層式關注網路辨識股市文章之維度型情感

何彥南、吳政隆

東吳大學巨量資料管理學院

E-mail: 06170214@gm.scu.edu.tw, jlwu@gm.scu.edu.tw

摘要

自然語言處理與機器學習的技術蓬勃發展，許多分析者透過機器學習來辨識文章的語意，進而分析出文章背後所表達的情感。也因相關技術的精進，人們不必再透過人工方式去閱讀大量文章，可以讓機器自動獲取想要的資訊，並輔助人們決策。在股票市場上，許多研究也指出有效的股票市場情感資訊，可以提升整體股票預測能力，也因為如此，對於股市新聞文章的情感分析研究也相繼出現。過去，研究多著重在如何分辨股票文章所表達的情感傾向，試圖抽取此資訊並用於預測股價。就目前而言，情感分析技術多是著重在情緒正負面(Valence)，而本研究想要在股票情感分析中導入情感激動程度(Arousal)。情感激動程度的辨識能夠反映該新聞內容觸發投資行為的強弱，情感激動程度可以更準確判斷交易的重要性。例如，表現正向(上漲)趨勢且平穩波動的內容，此時平穩波動仍不需要積極採取購入行為，反之亦然。本研究提出的人工智慧預測模型，能針對股票市場文章進行趨勢與波動的辨識，此股票市場文章辨識模型稱之為趨勢波動模型(Valence-Arousal, VA)，透過所提出的情感模型，可以精準的辨識股票市場文章的價格趨勢與交易即時性。

關鍵字：情感分析、股市新聞、階層式關注網路、VA 模型

一、研究動機與目的

在股市的領域，股價的變動難以預測，在一般的情況下投資者們的信心會反映在股價上，簡單來說多數人看好時則股票上漲，反之有可能下跌，其中存在許多複雜因素。股市新聞是一個接收資訊的重要管道，新聞內容會影響投資者對股市的信心與投資意願，這將會影響投資者的信心間接影響股票價格的走勢。正面的新聞內容可能會呈現出上漲趨勢，投資者將接受到可購入股票資訊，而負面新聞內容也會產生出售股票的資訊。

然而這些新聞文章的數量龐大，每個人的時間是有限的，要投資者每天去閱讀大量的股市新聞並接收重要信息，所以能自動化的分析文章的語意及資訊，並呈獻給投資者，是格外重要的，而情緒分析技術就是一種可以將大量的文本資料中，抽取並解析情感資訊的技術，可以有效的將情感資料的資訊量化，藉此輔助決策。

在情感分析中過去研究著重在股票文章的趨勢辨識，即文章的正負面情感傾向(Valence)，股票文章在 Valence 的應用中確實可以有效的幫助股票市場的預測，但在情感分析中的激動程度(Arousal)尚未在股票文章被探討，因此，本研究將 Arousal 概念引入股票文章之情感分析中。在股票文章分析的 Arousal 是一種表達股票市場投資行動的激動程度，即 Arousal 的高低代表進入市場的交易快慢，可以做為投資者採取投資行為的觸發程度標準，而這種同時考慮 Valence 與 Arousal 的模型稱是為 VA 模型。為了使預測股票新聞的 VA 狀態更加良好，本研究試圖加入情感字詞關注網路模型，使得 VA 分類的準確度有所提升。

在文章辨識上，HAN 模型是一個基礎且實用的模型，所以本研究使用 HAN 模型作為我們基本模型。在處理情感資訊的部分，使用 CVAW 和 NTUSD 等繁體中文的情感詞典去抽取文章中的情感詞，並建立情感關注模型(SAN)並與 HAN 做結合，會去探討在導入 SAN 模型後是否會比原先的 HAN 模型好，在本研究中還會比較不使用 HAN 模型時，分類結果會不會更好。

綜合上述的動機，本研究的問題將著重於建立股票市場文章的 VA 分類資料集建立及股市文章 VA 分類模型。以台灣股票市場作為研究標的物，進行股票文章的 Valence 與 Arousal 情感的資料標記。再以機器學習(Machine Learning, ML)中的深度學習(Deep Learning, DL)技術，建立一個股票文章深度學習 VA 預測模型，期望透過人工智慧的技術，進行自動化股票市場文章 VA 情感的學習與預測。

二、文獻回顧與探討

隨著網路的興起，改變了許多生態，股市的趨勢也漸漸受到網路上的各種消息影響，Agarwal 等人[1]針對 1992 至 2017 年所出版的論文，進行網路資訊擴散與股市反應相關的研究，其中分成社交媒體、網路新聞文章和投資者行為，可以得知網路上的資訊對股市是有一定影響力，而相關研究多數以網路社群與網路新聞為主流。近幾年自然語言處理(Natural Language Processing, NLP)的技術蓬勃發展，利用 NLP 技術分析網路文章的資訊，進而預測股市趨勢成為熱門議題[2-3]，另也有利 NLP 技術進行特定金融市場領域的新聞分析[4]。

近幾年情感分析(Sentiment Analysis)也是一

* 本研究接受科技部編號：MOST 108-2813-C-031-001-E 和 MOST 107-2218-E-031-002-MY2 研究計畫經費補助

個主要的熱門研究議題，有些研究是結合兩個情感詞典建構情感空間，並證實情感架構的導入可以提升模型的準確度[5]。以股市新聞與台灣股市市場的關係進行分析，提出了綜合新聞情感指數 ANSI，證明情感水準可用於投資組合決策[6]。另有研究指出，新聞文章包含的信息，會改變投資者的認知，並影響他們的投資決策，影響力會隨文章內容而異[7]。也有其他的研究是不同專家來定義數個情感關鍵字，以不同專家的觀點來獲得數個不同觀點的重要資訊，這將有助於提升股票市場的預測能力[8]。由此可知，藉由對股市新聞的自動化情感分析的分類，可輔助投資者進行相關的投資決策。

除了上述的情感狀態正負傾向(Valence)外，目前也有許多研究著重在情感激動(Arousal)的分析，Arousal 是探究情感的激動程度。例如，狂漲和上漲同樣是正面詞，但是他們是有程度差別的，「狂漲」這用詞背後隱含著高 Arousal，而一個字「上漲」，則是較低 Arousal。這種情感傾向-激動(Valence-Arousal, VA)模型已被廣泛使用在心理學相關的研究上。Yu 等人[9-11] 結合 VA 的概念建立中文維度型情感詞典。

機器學習技術近幾十年來相當熱門，許多研究也證實機器學習技術可以成功預測股票市場趨勢或價格變化[12-14]。多數研究利用機器學習技術進行財經新聞分析，自動化分析股市文章的情感程度，並用來預測未來股票價格的變動[15-16]。而然單方面的對文字做分析的效果有限，除了撰寫者的語意分析外，讀者的感受也是一個很重要因素，因此有些研究就針對讀者的情感去分析，像是 UAM 模型，可以捕捉讀者的情感反應[17]。另有研究表情符號(emojis)的情感表達可有效的獲得投資者的情感狀態，投資者在閱讀新聞時的實際感覺也是很重要，並結合深度學習技術進行讀者情感分類訓練，其結果顯示深度學習與表情符號可以有效預測讀者的情感狀態[18]。

除了情感特徵的擴張之外，近年來針對文章語意萃取技術多數採用深度學習技術，這樣可以保留文章的原始資料進行估計，意味著模型的預測與特徵萃取都是經由一個大型的神經網路完成，不需要再經由特徵工程或是降維的方式來獲得文字特徵。而一個有如人類閱讀文章的理解過程的模型將有助於萃取有用的特徵，Yang 等人[19]所提出的階層式關注網路(Hierarchical Attention Networks, HAN)用來做文章的分類，此 HAN 模型是以分層針對句子的字進行關注，再者針對文章的句子進行專注，透過 word level 與 sentence level 的分層估計後，相較於單純的循環神經網路(Recurrent Neural Networks, RNN)，在 6 個大型的文字分類問題上都獲得卓越的改善。

根據上述的文獻回顧，本研究以股票文章 VA 作為預測目標，有別於過去研究只在股票文章的情

感正負傾向進行分析與預測，本研究同時考量文章所呈現的情感激動程度，是一種在兩個維度上的情感分類，並以深度學習技術中的 HAN 模型作為基礎模型，進而對股票文章進行文字特徵萃取，這樣的模型將能更有效的掌握文章特性與資訊，對股市文章的 VA 分類會有所改善。

三、研究方法及步驟

本研究由兩個重要的部分進行研究，第一為資料的收集、標記與分類，第二則是預測模型建立與驗證。以下將針對這兩部分的研究進行詳細說明。

第一節 股票市場文章收集、標記與分類

在資料收集方面，以股票市場網路新聞為主，從鉅亨網¹ 收集相關台灣股市新聞，此網站提供許多與股票市場相關的資訊，也是目前較為完善的股票市場分析網站。

在 VA 標記方面，本研究採用 5 位投資者分別獨立標記，每筆新聞資料皆依標記者給予 Valence 和 Arousal 兩個程度值，標記值的範圍介於 1 分至 5 分，在 Valence 的 1 分表示負向情感(下跌趨勢)，5 分表示正向情感(上漲趨勢)。而在 Arousal，1 分表示激動程度低(不採取交易活動)，5 分表示激動程度高(採取交易活動)。標記後移除偏離值，即將五個標記值中排除最大和最小值，剩餘的標記值平均作為每篇新聞的 VA 數值。取得數值後，標記值以 3 作為類別區分標準，在 Valence 為 3 分(含 3 分)以下屬於下跌趨勢，反之亦然。而 Arousal 為 3 分(含 3 分)以下屬於不採取交易活動，反之亦然。當五筆標記的分數有極端的分歧，本研究會將此筆新聞排除。

舉例說明，當投資者閱讀一篇股市新聞後，覺得充分滿足上漲趨勢的語意，且必須立馬進行購入交易，VA 值將分別(V=5,A=5)，此時該新聞屬於 Type 1。相對來說，Type 3 則是充分滿足上漲趨勢，但沒那麼急迫，所以可以稍微觀望一下再進行投資。也就是說加入 Arousal 這個變數後，使電腦可以區分以上這兩種上漲狀態，相較於只使用 Valence，其結果可以表留更多資訊。關於四種類別的資訊可以看下圖：

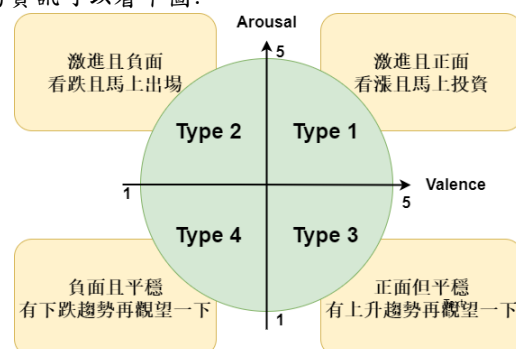


圖 1 股票新聞 VA 情感類型示意圖

¹ 鉅亨網提供財經新聞。https://news.cnyes.com/

第二節 深度 VA 辨識模型建立

本研究的深度 VA 辨識模型採用深度學習技術來建立，是一個階層式關注網路於 VA 情感辨識，整體的架構採用 Yang 等人[19]用所提出的階層式關注網路為主基礎，HAN 是用於萃取整篇新聞文章的重要特徵，另外，本研究也另外加入情感關注網路(Sentiment Attention Networks, SAN)進行重要特徵萃取。由圖 2 可以得知，首先每篇股市文章透過開源的 jieba² 所提供的斷詞系統進行斷詞。再者，斷詞後的詞將所與情感詞典進行比對，並獲得包含情感詞的字。所有斷詞後的字經過階層式關注網路進行估計，並獲得屬於該文章的上下文向量。該文章所包含有情感字經過情感關注網路進行估計，並獲得屬於該文章的在情感詞的向量。最後一個步驟則是用來估計文章分類，主要是以階層式關注網路的上下文向量和情感關注網路的向量來進行估計。

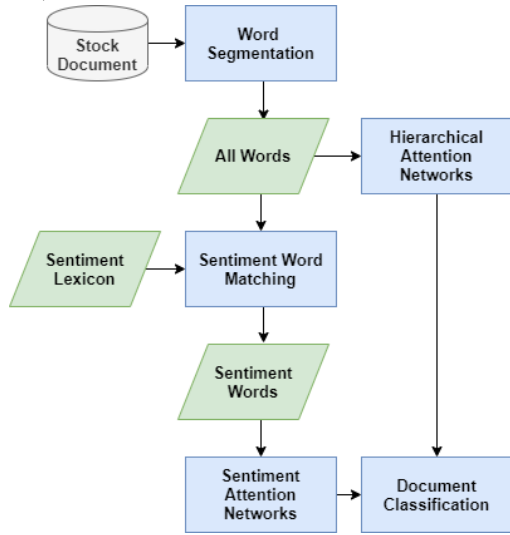


圖 2 VA 情感辨識模型流程圖

以下將針對 HAN、SAN 及文章分類訓練進行詳細說明。

1. 階層式關注網路於股票文章

階層式關注網路是一種分層估計法，模仿人的思考模式，像是當人要了解一篇文章時，會先找出每一句的重要字詞，再找出重要的句子，才能判斷整個文章的意思。此模型主要是估計一個句子的字向量，並估計不同字有不同的權重值，接著是估計一篇文章的句子向量，並估計不同句子有不同的權重值。詳細的估計過程將由以下步驟獲得：

1.1 字詞編碼器與關注

假設一個句子的字 $w_{ij}, j \in [0, N]$ ，首先要從嵌入矩陣(Embeddings Matrix)中獲得每個字嵌入。然後使用 Bidirectional GRU 來獲得每個字的隱藏狀

態，Bidirectional GRU 包含前饋式估計 $\overrightarrow{\text{GRU}}$ 與後饋式估計 $\overleftarrow{\text{GRU}}$ ，所以每個字經過兩個 GRU 估計後可獲得一個字的隱藏詞狀態 $h_{ij} = [\vec{h}_{ij}, \overleftarrow{h}_{ij}]$ 。而 h_{ij} 將可由以下公式獲得：

$$x_{ij} = W_e \varphi(w_{ij}), j \in [1, N], \quad (1)$$

$$\vec{h}_{ij} = \overrightarrow{\text{GRU}}(x_{ij}), j \in [1, N], \quad (2)$$

$$\overleftarrow{h}_{ij} = \overleftarrow{\text{GRU}}(x_{ij}) j \in [1, N], \quad (3)$$

其中， W_e 是可訓練的網路參數。而 GRU 則是使用標準的 GRU 網路模型。 N 是一個具句子的字數。接著要關注每個字的重要程度，而不是每個字都具有相同的重要程度。這意味著並不是每個字對分類任務有幫助，所以需要讓電腦知道一句話中哪些字較重要。注意力機制就可以用來估計其重要性，一個句子的上下文向量 s_i ，將可由以下公式獲得：

$$u_{ij} = \tanh(W_w h_{ij} + b_w) \quad (4)$$

$$\alpha_{ij} = \frac{\exp(u_{ij})}{\sum_j \exp(u_{ij})} \quad (5)$$

$$s_i = \sum_j \alpha_{ij} h_{ij} \quad (6)$$

其中， W_w 和 b_w 是可訓練的網路參數。首先用一層的神經網路估計每個字的重要性 u_{ij} ，再透過 softmax 函數對一個句子的所有字進行機率轉換，並獲得每個字的重要程度機率值 α_{ij} 。最後使用權重加總法計算一個句子的向量 s_i 。

1.2 句子編碼器與關注

上述已經可以獲得每個句子的向量 s_i ，而一篇股市文章的上下文向量將使用雙向 GRU 模型與關注機制獲得，即 $h_i = [\vec{h}_i, \overleftarrow{h}_i]$ ，其公式如下：

$$\vec{h}_i = \overrightarrow{\text{GRU}}(s_i), i \in [1, L], \quad (7)$$

$$\overleftarrow{h}_i = \overleftarrow{\text{GRU}}(s_i) i \in [1, L], \quad (8)$$

其中，GRU 仍是使用標準的 GRU 網路模型。接著利用關注機制來關注每個句子的重要性，一篇股市文章的情感向量 v^D ，其公式如下：

$$u_i = \tanh(W_s h_i + b_s) \quad (9)$$

$$\alpha_i = \frac{\exp(u_i)}{\sum_i \exp(u_i)} \quad (10)$$

$$v^D = \sum_i \alpha_i h_i \quad (11)$$

² Jieba 中文斷詞系統。 <https://github.com/fxsjy/jieba>

其中， W_s 和 b_s 是可訓練的網路參數。此 v^D 是一個經過階層是關注機制後的文章向量，此向量除了估計了每個字的重要程度，也同時估計了每個句子的重要程度，在這樣的估計下將有助於後續的文章分類。

2. 情感關注網路於情感詞彙 - Sentiment Attention Network (SAN)

除了上述的文章經由 HAN 模型估計外，其文章 VA 分類與文章內所使用的情感字詞有所，因此本研究導入情感詞彙進行關注情感相關資訊。根據 NTUSD 和 CVAW 正負面詞依序將每篇新聞的正、負面詞分別抽出，作為每篇文章的正、負情感詞輸入特徵，透過 SAN 模型估計情感詞隱藏特徵後，再與原本的 HAN 模型輸出的隱藏特徵作結合，希望藉由導入 SAN 模型，使整個模型可以提升在情緒分類相關任務的效果。下面依序介紹本研究提出來三種 SAN 為基礎的模型。

2.1 情感網路階層關注模型- Sentiment HAN (SHAN)

圖 3 為 SHAN 模型的基本流程，可以分成兩部分，Part 1 是萃取正負情感詞的隱藏特徵，Part 2 則是合併 HAN 的隱藏與 SAN 模型的隱藏特徵。在合併部分，採用 concatenation 方式，將 HAN 隱藏特徵與 SAN 隱藏特徵合併起來，作為 SHAN 模型，此模型也是本研究提出的主要情感模型。

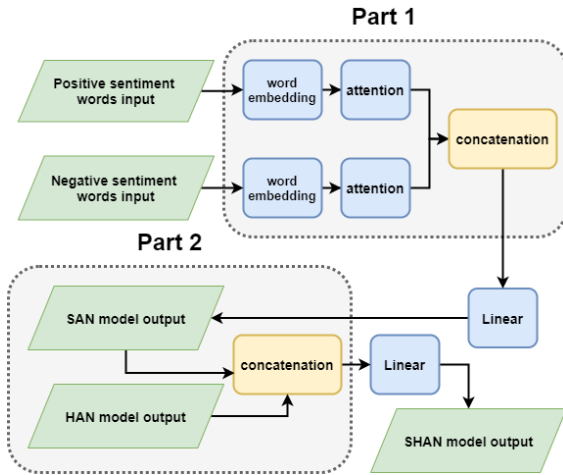


圖 3 情感網路階層關注模型流程圖

詳細 Part 1 的估算公式如下，分別提取正負情感詞 word embedding x_k^P 和 x_k^N ，每個正向詞索引 w_k^P 與負向詞索引 w_k^N ，可由情感字典 SD 獲得，以及透過 attention 機制進行估算隱藏特徵。在 Word embedding 的部分，一篇股市文章的正負向情感詞的詞遷入可有以下公式獲得：

$$x_k^P = W_e \phi(w_k^P), w_k^P \in SD \quad (12)$$

$$x_k^N = W_e \phi(w_k^N), w_k^N \in SD \quad (13)$$

其中， x_k^P 第 k 的正向情感詞的詞向量； x_k^N

第 l 的負向情感詞的詞向量。 ϕ 是 one-hot encoding 的轉換操作。而 W_e 是與公式(1)的 W_e 共享網路參數。關於 attention 機制，可由公式(9)至公式(11)相同，最後可以獲得正負向情感詞隱藏特徵，即分別為 v^P 與 v^N ，其中，attention 機制的網路參數與 HAN 為獨立學習的網路參數。最後 v^S 是由正負向情感隱藏特徵合併而得，並透過一層線性層估計，其中 W_s 和 b_s 是可訓練參數，即：

$$v^{PN} = \text{concat}(v^P, v^N) \quad (14)$$

$$v^S = W_s v^{PN} + b_s \quad (15)$$

在 Part 2 直接將情感網路輸出的隱藏特徵 v^S 與 HAN 網路的輸出 v^D 合併，中間經過一層線性層估算，最後輸出預測的類別 y 。其中 W_f 和 b_f 是可訓練參數，公式如下：

$$v = \text{concat}(v^S, v^D) \quad (16)$$

$$y = W_f v + b_f \quad (17)$$

2.2 GRU 情感網路階層關注模型 - GRU Sentiment HAN (GSHAN)

此模型是以上述的 SHAN 為基底，在處理情感特徵時加入了 GRU，希望可以透過 GRU 增加分類的準確度，由圖 3 可以看到在原本的 SHAN 模型的 Part 1 中加入了一層 GRU，Part 2 與 SHAN 一樣。

在 GSHAN 模型中，在正負向情感詞透過 attention 估計前，先透過一層 GRU 估計，正負情感詞將各別使用雙向 GRU 模型獲得隱藏特徵，即 h_k^P 和 h_k^N ，其公式如下：

$$h_k^P = [\overrightarrow{\text{GRU}}(x_k^P); \overleftarrow{\text{GRU}}(x_k^P)], \quad (18)$$

$$h_k^N = [\overrightarrow{\text{GRU}}(x_k^N); \overleftarrow{\text{GRU}}(x_k^N)], \quad (19)$$

其中，GRU 仍是使用標準的 GRU 網路模型。接著利用關注機制分別來正負向關注情感詞隱藏特徵(h_k^P 和 h_k^N)的重要性，可由公式(9)至公式(11)獲得，即 v^P 和 v^N 。後續可由公式(14)至公式(17)獲得最後的預測值 y 。

2.3 雙層情感網路階層關注模型 - Double Sentiment HAN (DSHAN)

此模型是以上述的 SHAN 模型為基底，雙層的意思是在模型中多加了一層 attention 層用來估計 HAN 和 SAN 的重要性，所以此模型稱做雙層情感網路階層關注模型。Part 1 與 SHAN 的 Part 一樣，而新加入的一層 attention 是在 Part 2，即將 HAN output 和 SAN output 結合時，採用 attention 進行估算。Attention 機制可由公式 (9)至公式(11)獲得。

第三節 文章分類訓練

由上述說明可以得知，最後每篇文章會得到一個 1×4 向量 Y ，本模型的預測目標為多類別分類。為了預測四種分類，本研究採用 Cross Entropy 計算訓練誤差，其誤差公式如下：

$$\text{loss}(Y, \text{class}) = -\log\left(\frac{\exp(Y[\text{class}])}{\sum_j \exp(Y[j])}\right) \quad (20)$$

其中， Y 表示第 d 篇的股市文章的類別標籤。而在收斂部分，根據誤差 loss 反饋傳遞計算梯度(Gradient)，是一種梯度下降(Gradient Descent)的收斂方式，並以隨機最佳化的 Adam 最佳化器進行網路參數收斂。上述模型能夠階層式萃取文章與情感詞間的隱藏特徵，為了開發此一模型，本研究採用 Pytorch³框架來建構深度學習模型。

四、實驗結果

第一節 資料集

本研究的新聞資料集是從鉅亨網抓取 2018 年下半年的股市新聞資料，其中將 7~10 月作為訓練資料，而 11~12 月則作為測試資料。

1. 資料集分布統計

本研究將資料切成三部分，分別為訓練資料集(Train)、驗證資料集(Validation)和測試資料集(Test)，訓練集有 2,628 筆，驗證資料集有 498 筆，測試資料集有 462 筆。本研究採用使用 Jieba 作為斷詞工具，並採用元智大學的中文維度型情感詞典(CVAW)⁴及台灣大學的情感詞庫(NTUSD)⁵，其中 CVAW 情感詞集包含 2,043 個正向詞與 3,369 個負向詞，而 NTUSD 情感詞即包含 2,813 個正向詞與 8277 個負向詞。由表 1 可以看到每筆新聞平均有 5.2 句，每句平均有 7.2 個詞左右，每篇新聞平均有 37 個詞。在情感詞的方面，每篇正面情感詞平均有 1.6 左右，每篇負面情感詞則平均不超過一個。

表 1 資料集基本資料

	Train set	Validation set	Test set
No. of document	2,628	498	462
Average sentence	5.28	5.20	5.20
Average word / per sentence	7.19	7.22	7.25
Average keyword	37.94	37.54	37.74
Average positive words	1.60	1.66	1.48
Average negative words	0.73	0.78	0.70

2. 情感詞的資料分佈

以合併台灣大學和元智大學的情感詞典做篩選，挑出每一篇的正負面詞。在這個資料集，多數資料都不包含情感字，3588 中有 2079 篇沒有負面詞，有 1008 篇沒有正面詞。其中正面詞較常出現。正面最多有一篇 10 個，負面最多則一篇 9 個。

3. 各預測類別資料分佈

其中 Type 1 佔了 51%，也就是說資料集有許多資料都落於正面且激進這類，Type 2 有 14%，Type 3 有 27%，而 Type 4 屬於負面但平穩得新聞只有 8%。可以知道目前資料的類別分佈不平衡，這也增加了分類模型在分類任務上的難度。

第二節 參數設計

本研究將實驗結果分別依據不同的參數分組，以各組驗證資料集的最高 F1 score 值作為挑選模型基準，其測試資料即的 F1 score 為最後模型效能比較用，並以此分數做為排名依據，使用此方法可以排除模型 overfitting 的狀況。基本的參數包含 RNN type、learning rate、hidden size 三組參數做比較，其中 RNN type 包含 LSTM 和 GRU，learning rate 分別測試 0.0001、0.001、0.01，Hidden size 分別測試 50、100、200。此外還會比較台灣大學與元智大學的情感詞對模型的改善程度。另實驗單純使用 SAN 且不合併 HAN 模型的效果比較。

第三節 分類模型說明

本研究共提出了七個不同的模型，用在對股市新聞的 VA 分類問題，單純使用 HAN 的模型作為 baseline 模型，依序介紹如下：

- **HAN**: 此模型為單純 Hierarchical Attention Networks，本研究將 HAN 套用到這次的 VA 分類上，並將其結果做為本實驗的比較標準，與本研究的數種 SAN 結合 HAN 的相關模型做比較。
- **SHAN**: 此模型為本研究所提出的 SAN 模型與 HAN 模型結合，也是本研究所提出的主要模型。
- **GSHAN**: 在 SHAN 的基礎下，加入一層 GRU 來估計正負向情感詞的隱藏特徵，此模型也是本研究的目標模型之一。
- **DSHAN**: 在 SHAN 的基礎下，在 SAN 與 HAN 的隱藏特徵透過 attention 層進行估算，希望透過模型自動調整 HAN 與 SAN 分別要使用多少資訊來判斷結果，此模型也是本研究提出的目標模型之一。
- **SAN**: 為了驗證 HAN 模型在此分類任務還是很重要的，即排除 HAN 模型直接使用 SAN 做分類，此模型為本研究不使用 HAN

³ Pytorch: a deep learning framework URL: <https://pytorch.org/>

⁴ Chinese Valence-Arousal Words (CVAW) is an affective lexicon containing 3,552 single words。

<http://nlp.innobic.yzu.edu.tw/resources/cvaw.html>

⁵ NTU Sentiment Dictionary(NTUSD)。

<http://academiasinicanlab.github.io/#resources>

的對照組。

- **GSAN**: 在 GSHAN 的基礎上，將 HAN 模型排除，做為不使用 HAN 的對照組之一。
- **DSAN**: 在 DSHAN 的基礎上，將 HAN 模型排除，做為不使用 HAN 的對照組之一。

第四節 不同 RNN 類型的效能比較

首先比較不同的 RNN 類型對 VA 分類結果的影響，即比較 GRU 和 LSTM 兩個演算法，四組模型採用不同的 RNN 模型，由表 2 可以看出採用 GRU 類型的 RNN，其 Val 型都比使用 LSTM 的模型表現較好，由此可知 GRU 在本研究的分類效果較優秀。

表 2 使用不同 RNN 種類的效能比較

Model	RNN type	Val loss	Val F1 score
HAN	GRU	0.85	0.66
	LSTM	0.94	0.60
SHAN	GRU	0.80	0.66
	LSTM	0.88	0.65
DSHAN	GRU	0.96	0.64
	LSTM	0.97	0.60
GSHAN	GRU	0.82	0.67
	LSTM	0.91	0.65

第五節 不同學習率的效能比較

針對不同的學習率作為比較，由表 3 可得知，其中表現最好的也是 GRU 情感網路階層關注模型(GSHAN)，但比較意外的是 0.001，它在 4 組模型中皆表現最好，也就是說在同個模型的情況下學習率為 0.001 是可以獲得最好的分類效果。

表 3 使用不同學習率的效能比較

Model	Learning rate	Val loss	Val F1 score
HAN	0.0001	0.96	0.58
	0.001	0.85	0.66
	0.01	1.04	0.62
SHAN	0.0001	0.93	0.60
	0.001	0.80	0.66
	0.01	1.07	0.63
DSHAN	0.0001	1.03	0.58
	0.001	0.96	0.64
	0.01	0.97	0.60
GSHAN	0.0001	0.83	0.65
	0.001	0.82	0.67
	0.01	1.07	0.61

第六節 不同隱藏層大小的效能比較

針對 hidden size 做效能比較，共有 50、100、200 這種大小。由表 4 可以得知，表現最好的是 GSHAN 模型，所使用 hidden size 為 50。在 4 個模型中有 3 個使用 50 的 hidden size 可以獲得相當好的表現，由此可知，hidden size 不是越高越好，而在本實驗 50 則是一個較好的選擇。

第七節 採用不同情感詞典的效能比較

為了驗證情緒字典的有效性，上述實驗結果皆採用合併 CVAW 和 NTUSD 的情緒字典。此部分將針對有使用情感詞的三個模型進行探討，使用不同的情感詞典影響分類結果的情況。實驗設計別是單純採用 CVAW 或 NTUSD，以及兩者合併使用(Both)，由表 5 可以看到最好的實驗結果是使用 NTUSD 的 GSHAN 模型相對較佳。比較特別的是個模型的組內差異小，所以基本上不同情感詞典在同個模型的變化不大，可以發現三個模型中有兩個模型採用 NTUSD 效果會是最佳的，可以合理推測篩選出的情感字不一定是越多越好。

第八節 未使用 HAN 模型的效能比較

這個部分，為了驗證情感字能否達到的一定程度的效果，也就是完全不使用 HAN 模型。由表 6 可以發現未採用 HAN 模型時，其效果變的非常差，將近有 0.1 的差距，也就是說本研究提出的 SHAN 與 GSHAN 兩個模型，分類效果比單純使用 HAN 模型較好，但不能沒有 HAN 模型對文章進行特徵萃取。因此情感詞加入股市新聞的 VA 分類可以有效提高準確度，也就是說不能完全捨去非情感詞的新聞內容，這樣會導致資訊量過少，因為每篇股市新聞文章的情感字詞的使用率偏低。

表 4 使用不同隱藏大小的效能比較

Model	Hidden size	Val loss	Val F1 score
HAN	50	0.83	0.65
	100	0.85	0.66
	200	0.94	0.61
SHAN	50	0.80	0.66
	100	0.86	0.64
	200	0.92	0.63
DSHAN	50	0.96	0.64
	100	0.99	0.58
	200	0.98	0.58
GSHAN	50	0.82	0.67
	100	0.91	0.65
	200	0.83	0.65

表 5 使用不同情感辭典的效能比較

Model	Sentiment dictionary	Val loss	Val F1 score
SHAN	Both	0.81	0.64
	CVAW	0.80	0.65
	NTUSD	0.80	0.66
DSHAN	Both	0.96	0.64
	CVAW	1.00	0.59
	NTUSD	0.96	0.63
GSHAN	Both	0.79	0.66
	CVAW	0.79	0.66
	NTUSD	0.82	0.67

表 6 HAN model 與 SAN model 比較表

Model	Val loss	Val F1 score
SHAN	0.80	0.66
SAN	1.00	0.59
GSHAN	0.82	0.67
GSAN	0.98	0.60

第九節 各模型比較

此部分將針對整理上述的實驗結果進行模型間的彙整與比較。從各組模型中挑出 Val F1 score 最佳的各模型參數組合，做為代表該模型內的最好結果，再依各模型的 Test F1 分數進行比較，由表 7 可以發現其中 SHAN 最為出色，然而在 Val F1 分數表現最佳是 GSHAN 模型，但該模型的 Test F1 分數並未獲得最佳效果，但是該模型的 Test loss 是最低的。可以發現本研究提出的三種結合 SAN 的模型，有兩組效果明顯比只使用 HAN 模型來的好，其中 DSHAN 模型雖然最終 Test F1 分數輸給 HAN 模型，但是它的 Test loss 也是比 HAN 模型來的低，依此結果來說，本研究所提出的情感網路模型是可以有效提高股票新聞 VA 分類準確度。

表 7 所有模型比較表

Model	RNN type	Hidden size	Learning rate	Sentiment dictionary	Test F1 score	Rank
SHAN	GRU	50	0.001	NTUSD	0.70	1
GSHAN	GRU	50	0.001	NTUSD	0.66	2
HAN	GRU	100	0.001	None	0.65	3
DSHAN	GRU	50	0.001	Both	0.63	4
SAN	LSTM	100	0.0001	Both	0.59	5
GSAN	GRU	200	0.0001	Both	0.58	6
DSAN	LSTM	200	0.01	Both	0.58	7

第十節 綜合討論

由上述的實驗數據結果，本研究成果將分成三個層面進行綜合分析，分別是參數調整、情感詞庫比較、不同的模型的效果。

1. 模型超參數分析

根據所有實驗結果可以發現，在使用 SHAN 模型的情況下，採用 GRU 模型、0.001 的學習率、50 的 hidden size，獲得最佳的分類效果。

2. 情感字詞典分析

本研究測試台灣大學和元智大學所提供的情感詞典，從本研究的實驗結果發現，台灣大學的詞典效果比元智大學的詞典優秀。另外還可以發現情感詞的多寡對預測的影響相對較小，主要原因是本研究中情感詞佔的比例相對較少，所以最終分類的效果主要還是模型本身的學習能力為主。情感字詞的導入可以作為輔助效果，有效的提升分類模型

3. 預測模型效能分析

所有模型分成三組，從上面的分析結果整理成圖 4，有合併 SAN 模型與 HAN 模型的三種模型 (SHAN、GSHAN 和 DSHAN)，雖然其中 DSHAN 的效果比 HAN 差，但是其他兩個 SHAN 模型都勝過單純的 HAN 模型，也就是說加入情感網路關注模型是可以有效的提升 HAN 在預測股市新聞 VA 類別的準確度。另外我們還可以發現，當完全不使用 HAN 的模型 (SAN 和 GSAN) 時效果最差，可以證實單純使用情感詞做預測，是無法達到現有的分類效果，所以 HAN 模型所扮演的文章解析腳色，是預測模型的非常必要的一部分。

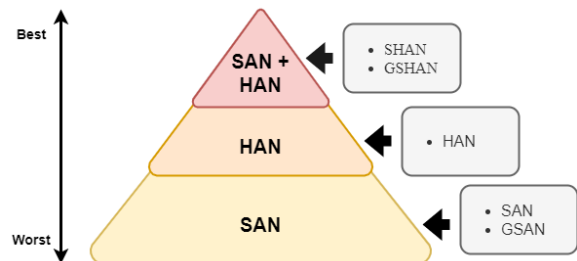


圖 4 各模型效能綜合整理與比較

五、結論與未來展望

在股市新聞中隱含許多資訊，許多投資者接收到相關的訊息進而影響到投資決策，進而影響到股價的波動，本研究使用非結構化的新聞，抽取有用的相關資訊辨識股市文章情緒表達。其所導入的 Arousal & Valence 概念將有助於標記股市新聞文章的情感導向，本研究透過深度學習的技術成功將股市新聞進行 Arousal & Valence 分類，相較於單純的正負向情緒分析，此方法可以更有效的表示股市新聞中的情感強度的資訊。本研究的 HAN 系列模型，結合 CVAW 和 NTUSD 提供的詞典，有效的將文章中重要的情感詞抽取出來當作情感特徵，並成功導入情感網路關注模型 (SAN) 提升情感文章分類的準確度。

從實驗結果中可以看到本研究提出的三種與

HAN 結合的情感網路關注模型，其中有兩組比單純使用 HAN 做分類的效果更好，也就是說使用情感網路關注模型(SAN)來輔助 HAN 模型，是真的可以使整個模型的效果更好。為了做對照，本研究將三組情感關注模型抽取出來，不使用 HAN 模型，也就是說整個分類的過程只使用情感詞和 SAN 模型，可以看到效果明顯下降許多，所以我們可以知道在 VA 分類的任務中，情感詞固然重要，但是文章中其他的用詞也是很重要的。

因為目前研究的部分只針對文本分類部分，外來可以將研究成果實際應用，做出一個可以真的提供輔助投資者資訊的服務。此外未來可以把本研究的情感網路關注模型套用在不同的文本分類情境下，像是情感詞比較豐富的歌詞，其效果將可比新聞文章更適用，因為本次實驗的股市新聞資料，多數股市文章使用情感詞的數量相對較少。所以，本研究未來預計著重在特徵向量的精進，從隨機生成的 word embeddings，延伸導入 Word2vec、GloVe、BERT 等預訓練 word embeddings，進而加強文字資訊的萃取與辨識，提高整體模型預測的準度。

參考文獻

- [1]. Shweta Agarwal, Shailendra Kumar, Utkarsh Goel, "Stock market response to information diffusion through internet sources: A literature review," *International Journal of Information Management*, vol. 45, pp. 118-131, 2019.
- [2]. Weiling Chen, Chai Kiat Yeo, Chiew Tong Lau, Bu Sung Lee, "Leveraging social media news to predict stock index movement using RNN-boost," *Data & Knowledge Engineering*, vol. 118, pp. 14-24, 2018.
- [3]. Yefeng Ruan, Arjan Durrresi, Lina Alfantoukh, "Using Twitter trust network for stock market analysis," *Knowledge-Based Systems*, vol. 145, pp. 207-218, 2018.
- [4]. Stephen Kelly, Khurshid Ahmad, "Estimating the impact of domain-specific news sentiment on financial assets," *Knowledge-Based Systems*, vol. 150, pp. 116-126, 2018.
- [5]. Xiaodong Li, Haoran Xie, Li Chen, Jianping Wang, Xiaotie Deng, "News impact on stock price return via sentiment analysis," *Knowledge-Based Systems*, vol. 69, pp. 14-23, 2014.
- [6]. Yu-Chen Wei, Yang-Cheng Lu, Jen-Nan Chen, Yen-Ju Hsu, "Informativeness of the market news sentiment in the Taiwan stock market," *The North American Journal of Economics and Finance*, vol. 39, pp. 158-181, 2017.
- [7]. Qing Li, TieJun Wang, Ping Li, Ling Liu, Qixu Gong, Yuanzhu Chen, "The effect of news and public mood on stock movements," *Information Sciences*, vol. 278, pp. 826-840, 2014.
- [8]. Jheng-Long Wu, Liang-Chih Yu, Pei-Chann Chang, "An intelligent stock trading system using comprehensive features," *Applied Soft Computing*, vol. 23, pp. 39-50, 2014.
- [9]. Liang-Chih Yu, Lung-Hao Lee, Shuai Hao, Jin Wang, Yunchao He, Jun Hu, K. Robert Lai, and Xuejie Zhang, "Building Chinese affective resources in valence-arousal dimensions," In: *Proceedings of NAACL/HLT-16*, pp.540-545, 2016.
- [10]. Liang-Chih Yu, Lung-Hao Lee, Kam-Fai Wong, "Overview of the IALP 2016 shared task on dimensional sentiment analysis for Chinese words," In *Proceedings of IALP-16*, pp.156-160, 2016.
- [11]. Liang-Chih Yu, Lung-Hao Lee, Jin Wang, Kam-Fai Wong, "IJCNLP-2017 task 2: Dimensional sentiment analysis for Chinese phrases," In: *Proceedings of IJCNLP-17, Shared Tasks*, pp. 9-16, 2017.
- [12]. Suryoday Basak, Saibal Kar, Snehanishu Saha, Luckyson Khaidem, Sudeepa Roy Dey, "Predicting the direction of stock market prices using tree-based classifiers," *The North American Journal of Economics and Finance*, vol. 47, pp. 552-567, 2019.
- [13]. Bruno Miranda Henrique, Vinicius Amorim Sobreiro, Herbert Kimura, "Stock price prediction using support vector regression on daily and up to the minute prices," *The Journal of Finance and Data Science*, vol. 4(3), pp. 183-201, 2018.
- [14]. Mu-Hui Yu, Jheng-Long Wu, "CEAM: A novel approach using cycle embeddings with attention mechanism for stock price prediction," In: *Proceedings of BigComp2019 – IWSF*, 2019.
- [15]. Wen Long, Linqiu Song, Yingjie Tian, "A new graphic kernel method of stock price trend prediction based on financial news semantic and structural similarity," *Expert Systems with Applications*, vol. 118, pp. 411-424, 2019.
- [16]. Bin Weng, Lin Lu, Xing Wang, Fadel M. Megahed, Waldyn Martinez, "Predicting short-term stock prices using ensemble methods and online data sources," *Expert Systems with Applications*, vol. 112, pp. 258-273, 2018.
- [17]. Weiming Liang, Haoran Xie, Yanghui Rao, Raymond Y.K. Lau, Fu Lee Wang, "Universal affective model for Readers' emotion classification over short texts," *Expert Systems with Applications*, vol. 114, pp. 322-333, 2018.
- [18]. Nader Mahmoudi, Paul Docherty, Pablo Moscato, "Deep neural networks understand investors better," *Decision Support Systems*, vol. 112, pp. 23-34, 2018.
- [19]. Zichao Yang, Diyi Yang, Chris Dyer, Xiaodong He, Alex Smola, Eduard Hovy, "Hierarchical attention networks for document classification," In: *Proceedings of the NAACL-HLT 2016*, 2016.