

以注入情緒構面的雙向長短期記憶模型應用於社群媒體產品評論之情感分析

Using Valence and Arousal-infused Bi-LSTM for Sentiment Analysis in Social Media Product Reviews

Yu-Ya Cheng¹, Wen-Chao Yeh², Yan-Ming Chen³, Yung-Chun Chang^{*}

¹Professional Master Program in Data Science, Taipei Medical University, Taipei, Taiwan

²Institute of Information Systems and Applications, National Tsing Hua University, Hsinchu, Taiwan

^{3, *} Graduate Institute of Data Science, Taipei Medical University, Taipei, Taiwan

^{1, 3, *} {i906108009, m946109002, changyc}@tmu.edu.tw

²wyeh@m109.nthu.edu.tw

摘要

現今是網際網路普及的時代，社群媒體平台擔任私人企業與大眾間相互溝通的橋樑角色。本研究之目的是透過解析不同領域的產品評論資料獲取用戶對於產品的心得。我們提出一個具有更細膩詞彙情感資訊的 BiLSTM (Bi-directional Long-Short Term Memory) 模型，不只能預測文字中最小詞素「詞彙」的情緒構面 Valence & Arousal，也能加入詞彙間的依存關係。實驗結果顯示，本研究在預測詞彙之 Valence & Arousal 可以達到良好的效能。而且，經由融合 VA (Valence & Arousal) 及詞彙間依存關係至 BiLSTM 模型更可以對於社群文字情感分析做出優異表現，驗證此模型對於產品評論留言的情感預測準確性。

Abstract

With the popularity of the current Internet age, online social platforms have provided a bridge for communication between private companies, public organizations, and the public. The purpose of this research is to understand the user's experience of the product by analyzing product review data in different fields. We propose a BiLSTM-based neural network which infused rich emotional information. In addition to consider Valence and Arousal which is the smallest morpheme of emotional information, the dependence relationship between texts is also integrated into the deep learning model to analyze the sentiment. The experimental results show

that this research can achieve good performance in predicting the vocabulary Valence and Arousal. In addition, the integration of VA and dependency information into the BiLSTM model can have excellent performance for social text sentiment analysis, which verifies that this model is effective in emotion recognition of social medial short text.

關鍵字：情感分析、情緒構面、社群媒體、產品評論

Keywords: Sentiment Analysis, Valence & Arousal, Social Media

1 研究動機與目的

隨著人們越來越依賴不斷發展的網路科技，各種便利的網路購物平台及社群平台如雨後春筍般興盛，不只轉換人們散佈及獲取資訊的管道，也改變現代人的生活型態。例如：可以從 Google 地圖中看到不同地標的評論，而這些評論也確實影響大多數人在行動前會先參考其他顧客的意見作為考量。但是，累積越來越多的評論卻使人們無法迅速的閱讀到關鍵資訊。情感分析技術能夠對這些資料進行有效地分析與挖掘，並識別出情感傾向，成為自然語言處理 (Natural Language Processing, NLP) 領域中最活躍的研究主題之一。

分析產品評論是瞭解群眾意見相當重要的一環，透過分析留言內容可掌握評論的情感傾向與趨勢。情感分析已有許多相當先進的研究，因此本論文著重於目前的先進研究上做出突破，計畫應用預測詞彙的情緒構面

(Valence-Arousal) 來提取出社群媒體上的輿論情感。除了分析句子的詞彙情感外，句子中各詞彙間的依存關係也是組成句子的重要參考依據。充分考量詞彙間的關聯性可獲得更細膩的分析資料，本研究在多種領域的產品評論資料上使用有別於以往的特徵方法進行情感分析任務。

2 相關研究探討

隨著近幾年網路普及與社群網站平台的盛行，網際網路已經成為分享訊息的重要場域。人們透過諸如 Facebook、Instagram 等社交平台發表個人意見或討論當前的公眾議題。如何從大量的評論留言抓取重點與關鍵表達，更成為資料科學的熱門研究主題。近年來，透過情感分析 (sentiment analysis) 技術從社群媒體文本中檢測和提取主觀信息 (例如：觀點和態度) 在眾多不同領域的應用相當廣泛 (Liu, B., 2012)，尤其在政治 (Shakeel and Karim, 2020)、產品 (Zhang et al., 2012) 與電影 (Yenter and Verma, 2017) 等領域上都有傑出的表現。

在語法的架構單位上，Turney (2002) 提出以文檔級別的情感分類任務方法，接續有其他研究提出以更小單位對此進行分類任務，至今發展成可以句子級別 (Hu et al., 2004)、短語級別 (Agarwal et al., 2009) 與詞彙級別 (Sayeed, Asad, et al., 2012) 進行分類任務。本研究將採用最小語素的詞彙級別進行情感預測實驗，並將詞彙情感預測延伸為每則留言的情感分類任務，進而以此探討產品評論的情感層級。本論文提出有別以往的方法，從預測詞彙情感的結果去追溯原評論內容，並探討群眾意見之情感面向。

情感分析的研究領域基本作法是先將情感區隔表示成數個類別 (例如：正面、負面)，但僅以情感類別來分類將無法呈現各種情感的高低程度。情緒構面 (Valence-Arousal, VA) (Yu, Liang-Chih, et al., 2015) 為使用雙維度來表示情感的類別及程度，藉由預測詞彙的 Valence & Arousal 來清楚辨識該詞彙的情感類型及呈現出來的強弱程度。其中，Valence 為表示情感程度分級，以 1 至 9 的連續數值代表

情緒由負面到正面。Arousal 為情緒激動程度，由 1 至 9 的連續數值代表，其數值由高到低分別代表情緒由激動到平靜。Valence-Arousal 能清楚辨識詞彙的極性分數也可以有效提升情感分析的效能 (Chang et al., 2019; Cheng, Yu-Ya, et al., 2021)，因此本研究將情緒構面融合至雙向長短期記憶神經網路於分析文本情感，並應用在中文評論留言上以證明其有效性。

此外，依存句法分析 (Dependency Parsing) 是將句子解析成一棵依存句法樹，描述句子中詞彙與詞彙間的依存關係。本研究採用哈爾濱工業大學的 LTP 自然語言處理套件 PyLtp 進行文本依存句法關係探討。透過分析語言單位內成分間的依存關係來揭示其句法結構：文本句法結構為詞彙與詞彙間的修飾關係，兩個詞彙間連接一個依存關係，而依存關係包含許多類型。例如：主謂關係 (SBV)、動賓關係 (VOB) 等。近年來有許多相關研究使用到 LTP 的依存句法分析皆獲得良好的效能 (Zhai, Pengjun, et al., 2020; Qidi, Jiao., 2021)，證明加入句子結構特徵可增強語意信息的表達能力，進而幫助提升模型學習能力。

3 研究方法

首先將經由社群媒體產生的大量非結構化輿論文字資料施行預處理，進而預測詞彙之情緒構面向量，基於 BiLSTM 模型建構情感分類之預測機制來推論出文字之最小語素「詞彙」的情感資訊：Valence 與 Arousal，再將之融合進深度學習模型中，更細膩地分析社群文本之情感表現。

本節將依序介紹 VA 預測方法以及將 VA 當作特徵來預測多個產品領域的情感分類任務。

3.1 情緒構面預測

計畫透過更細粒度的情感分析預測詞彙的情感極性分數，並加入特徵信息提升模型推論表現。以下將依序介紹所使用到的特徵：

1. POS: 先以正體中文語料訓練出來的 MONPA¹ (Hsieh, Yu-Lun, et al., 2017) 斷詞套件取得每條句子中各個中文詞彙的詞性標註，再將這些詞性轉

¹ <https://github.com/monpa-team/monpa>

化為具 50 維的表示向量，作為後續預測詞彙情感的重要特徵。

2. NTUSD: 參考台大中文情感極性詞典 (NTUSD)² 將文本詞彙進行情感正負面分類，例如：正面中文詞彙包含「開心、大方、公平」等；負面中文詞彙包含「反感、可怕、失望」等。若該詞彙出現在正面詞典列表，類別表示為 1，詞彙位於負面詞典，則將類別表示為 -1。當詞彙未被 NTUSD 詞典收錄，則使用餘弦相似度 (cosine similarity) 對其進行排序，藉此挑選出前五個在 NTUSD 有收錄最相近的詞彙，並以多數決選擇所屬類別 (Chang, Yung-Chun, et al., 2019)，以此詞彙分類方法取得情緒特徵。
3. E-HowNet: E-HowNet 是基於 HowNet 拓展出的知識本體，依據詞彙的語義將其轉換為概念 (Chen, Wei-Te, et al., 2010)，例如：「一去不返」的概念是「消失」、「一心一意」的概念是「誠心」，透過 E-HowNet 將相同概念的 VA 值取平均，當作該詞彙的語義特徵。如未能找到相同概念的詞彙，就利用餘弦相似度 (cosine similarity) 找到最相近的概念，並取其 VA 平均值。透過以上方法可以得到兩維的詞彙語義特徵。

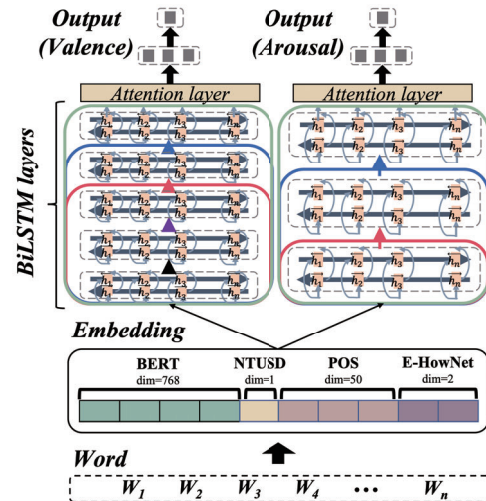


圖 1. 基於深度神經網路之 V & A 預測模型

本研究採用由 Devlin et al. (2018) 提出的 BERT (Bidirectional Encoder Representations from Transformers) 從文本資料中提取語言特徵向量表示作為嵌入向量表示，將其與上述經由 POS、NTUSD 與 E-HowNet 等方法取得之特徵向量做串聯得到該詞彙完整向量表示，如圖 1 的 Embedding 所示，共 821 維。

雙向長短期記憶 (BiLSTM) 除了解決較長距離的依賴性外，藉由組合前向和後向的 LSTM 可以更好捕獲雙向語意特徵，疊加 Attention Layer 對輸入的每個部分賦予不同權重 (Vaswani, Ashish, et al., 2017)。近年來，集成學習方法已被廣泛用於提高分類性能 (Moreno, Jose G., 2020; Gomes, Heitor Murilo, et al., 2017; Dong, Xibin, et al., 2020)，因此本研究將使用集成學習的方法透過組合不同級別的模式來得到更好的表達情感輸出。圖 1 是集成學習模型預測 V & A 的架構。

² <http://academiasinicanlplab.github.io/>

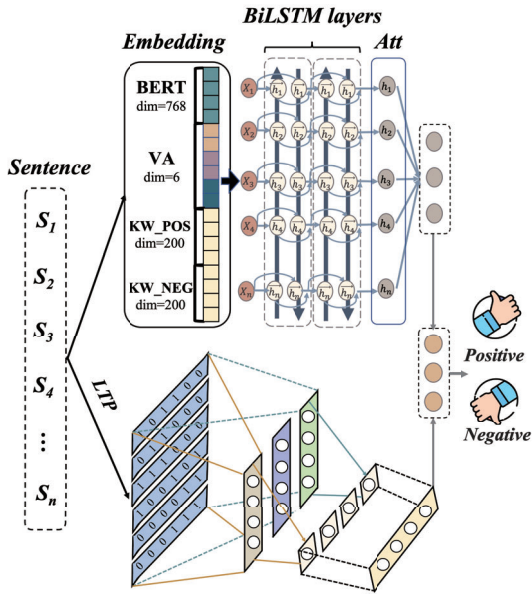


圖 2. 社群留言情感預測模型 (BLV-depCNN)

3.2 社群留言情感預測

圖 2 為基於情緒構面之社群留言情感分析模型架構圖，將上下兩部分的預測輸出合併成最終分類結果。圖 2 上半部分，先將社群留言以 MONPA 套件斷詞並移除 Stop Words，再以 3.1 節所述取得所有詞彙的 Valence 與 Arousal，結合句子語義特徵後丟進深度學習模型，得到該留言的情感傾向，依序介紹所使用的特徵：

1. 情緒關鍵詞特徵：本研究透過對數似然比 (Log Likelihood Ratio, LLR) 擷取出正面和負面的關鍵詞彙。公式 (1) 可獲取假設的可能性，為有利於選擇關鍵單詞特徵的方法。其 PS 表示訓練樣本(句子)中肯定句子的集合； $N(PS)$ 和 $N(\neg PS)$ 表示正樣本和負樣本的數量； $N(w\wedge PS)$ 表示包含正面詞 w 的正面句子的數量。經最大似然估計來獲得概率 $p(w)$ 、 $p(w|PS)$ 和 $p(w|\neg PS)$ ，具有較高 LLR 的詞彙視為與特定情感具有較高的聯繫。選出正負面各 200 個關鍵詞進行特徵表示，分別將原始句子的詞彙對應到正負面關鍵詞列表中，並將所對應到的詞彙以權重數字取代，其他則以「0」表示，因此會形成正面 200 維與負面 200 維，共 400 維的特徵向量。

$$-2 \log \left[\frac{p(w)^{N(w\wedge PS)} (1-p(w))^{N(PS)-N(w\wedge PS)}}{p(w|PS)^{N(w\wedge PS)} (1-(w|PS))^{N(PS)-N(w\wedge PS)}} \right] \times \left[\frac{p(w)^{N(w\wedge \neg PS)} (1-p(w))^{N(\neg PS)-N(w\wedge \neg PS)}}{p(w|\neg PS)^{N(w\wedge \neg PS)} (1-(w|\neg PS))^{N(\neg PS)-N(w\wedge \neg PS)}} \right] \quad (1)$$

2. 情緒構面特徵：透過上述 LLR 關鍵詞去對應原文的 VA，只選取權重最高的前三名，將其 Valence 和 Arousal 作為特徵表示。若不足三組則選擇原文中極性最高的 Valence 和 Arousal 補足到三組，因此共可獲得共 6 維的特徵表示。將 BERT 文字向量表示法與上述的 LLR 與 VA 特徵向量合併後獲得 1,174 維的向量，投入 BiLSTM + Attention Layer 的推論模型。

3. 依存關係特徵：圖 2 下半部分為加入句子語義特徵。使用 LTP 套件中的依存句法分析 (Dependency Parsing) 獲得句子中各詞彙的依存關係，如圖 3 所示。箭頭連接兩詞彙表示含有依存關係，紅色字體顯示兩者的依存關係類型，將兩詞彙間的關係轉換為矩陣形式。如圖 3，詞彙間有關係以「1」表示，無依存關係則以「0」表示。考量每句話的詞彙數量可能不相等，將所有矩陣長度都轉換為最長句子的長度。例如，此資料集最長句子長度為 145，就將所有矩陣大小轉換成 [145x145] 後投入卷積神經網絡 (Convolutional Neural Network, CNN) 模型。

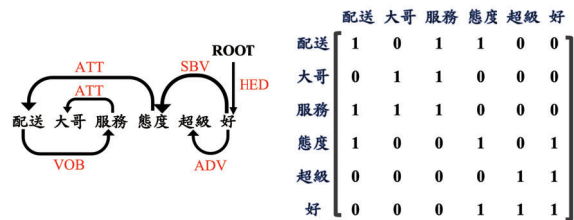


圖 3. 依存分析樹與句子關係矩陣之範例

最後，結合 BERT 文字向量表示法、VA 以及 LLR 的 BiLSTM 加上 Attention Layer 之推論模型的輸出與考慮詞彙間依存關係的矩陣特

徵放入 CNN 模型，使用 concat layer 進行合併。concat layer 的功能為將兩個通道 (Channel) 進行拼接，這有利於在向量表示上具有更多特徵並提高分類性能，經過兩層全連接層 (Dense Layer) 取得二元分類預測結果。命名為 BLV-depCNN 模型，用來分析文本情感傾向。

4 實驗結果與討論

實驗分成兩部分，第一部分運用 VA 預測模型取得以多維度連續數值方式來精確表達情感的詞彙 VA。第二部分則是延用第一部分的結果當成特徵進行情感分類任務。實驗結果皆優於以往的方法，確認本模型架構的價值。

4.1 情緒構面預測之效能評估

資料集來自於 2017 年 IJCNLP 的共享任務：中文短語的維度情感分析 (Dimensional Sentiment Analysis for Chinese Phrases, DSA_P)，稱為 CVAW 3.0。訓練資料有 2,802 筆，測試資料為 750 筆，共 3,552 筆具有 VA 等級分數的中文維度型情感詞彙，所有中文詞彙皆為人工標註 Valence 和 Arousal 的數值。將所有矩陣大小轉換成 [145x145] 後投入超參數設置 batch_size 為 64、optimizer 為 Adam、dropout rate 為 0.2。

本實驗之評估指標比照 DSA_P 預測任務之平均絕對誤差 (Mean Absolute Error, MAE) 及皮爾森相關係數 (Pearson Correlation Coefficient, PCC) 進行評估，如以下公式：

$$\text{MAE: } \text{MAE} = \frac{1}{n} \sum_{i=1}^n |A_i - P_i| \quad (2)$$

$$\text{PCC: } r = \frac{1}{n-1} \sum_{i=1}^n \left(\frac{A_i - \bar{A}}{\sigma_A} \right) \left(\frac{P_i - \bar{P}}{\sigma_P} \right) \quad (3)$$

其中，n 代表測試資料個數， A_i 表示人工註釋的 V 值和 A 值， P_i 表示預測出的 V 值和 A 值， \bar{A} 和 \bar{P} 分別為 A 和 P 的算術平均數， σ_A 和 σ_P 分別為 A 和 P 的標準差。

本研究使用三種不同的嵌入方法，將 BERT 文字向量表示法與 Word2Vec 和單字向量嵌入 (Character Embedding, CE) 相比，分別加入深度神經網路 (Deep Neural Network, DNN) 後的結果如表 1 所示。BERT 在 Valence 與 Arousal 的平均絕對誤差 (MAE) 與皮爾遜相關係數 (PCC) 相較於其他嵌入方法明顯獲得最好的效

能。因為 Word2Vec 的每個詞彙表示一個固定的向量，CE 則是每個單詞表示一個固定的向量，兩者皆與單詞出現的上下文無關，而 BERT 生成的向量特徵表示是有考慮上下文的關係及位置關係，因此所生成的向量表示較具有優勢，從而提高模型的效能。

表 1. 嵌入方法之性能比較

Embedding	Valence		Arousal	
	MAE	PCC	MAE	PCC
CE	0.824	0.710	0.936	0.535
Word2Vec	0.947	0.662	1.194	0.138
BERT	0.687	0.831	0.878	0.634

表 2 為本研究方法 (Our method 欄位) 與 VA 預測任務前三名 (Wu, Chuhan, et al., 2017; Zhou, Xin, et al., 2017; Li, Peng-Hsuan, 2017) 以及疊加不同特徵的效能比較。本研究方法與基於 BERT 疊加多個特徵向量 (BERT+POS 與 BERT+POS+NTUSD) 可以發現考慮越多的特徵可有效提高模型預測能力，凸顯出在 NTUSD 的特徵中將詞彙分成正負兩類可有效幫助模型判斷 Valence 值可能落點區間，而 E-HowNet 的特徵原理是將相似概念的詞彙獲得一樣的特徵表示，可增加模型學習能力，因此可說明考慮越多基於詞彙特徵可有效提升模型效能。

表 2. 嵌入方法之性能比較

Group	Valence		Arousal	
	MAE	PCC	MAE	PCC
THU_NGN	0.509	0.908	0.864	0.686
Our method	0.543	0.887	0.855	0.689
AL_I_NLP	0.545	0.892	0.857	0.678
CKIP	0.602	0.858	0.949	0.576
BERT + POS + NTUSD	0.605	0.869	0.875	0.617
BERT + POS	0.646	0.837	0.883	0.615

本論文所提出之方法在預測中文單詞的情感維度方面表現出眾，媲美排名最高的 THU_NGN。由於本研究充分考慮詞彙間的特性，疊加三種不同特徵以增加模型訓練，且使用考慮單詞向量也考慮上下文及位置關係的 BERT，而能提升模型效能。對於 Valence 的預測結果中 MAE 為 0.543，PCC 增至 88.7%，

這表明本模型預測與正確值之間具有非常高的相關性；對於 Arousal 的預測結果中 MAE 大幅降低至 0.855，PCC 增至 68.9%，證明本模型之優勢。

4.2 社群文本情感預測之效能評估

第二個實驗將利用 VA 預測模型的輸出，集成到混合的深度神經網路模型，對社交媒體評論的整體情緒進行分類。資料集來自於 2014 年 NLPC 競賽的共享任務「Sentiment Classification with Deep Learning」，以下簡稱為 SCDL。該資料集由中文及英文的產品評論網站上收集得出，此數據來自多個產品領域，包含書籍、DVD 和電子產品等多個領域。資料標籤區分為正面與負面兩個類別，各有 6250 筆。訓練集包含 5,000 筆正面資料以及 5,000 筆負面資料。測試資料包含 1,250 筆正面資料以及 1,250 筆負面資料。為了證明所提出模型的泛化能力，我們選用電子商務服務評論數據集 (ECSR) 以及外送平台用戶評論數據集 (URDP) 進行比較實驗。ECSR 數據集包括從電子商務網站收集的電視產品和發行服務的評論，在該數據集中每個評論都帶有一個情感標籤：正面或負面，數據總共包含 4,212 條評論，其中 1883 條為正面，2329 條為負面。URDP 數據集為某外送平台收集的用戶評論，每則評論皆帶有一個情感標籤：正面或負面，數據集總共包含 6000 筆對於外送餐點、外送員以及平台使用相關評論，正面與負面筆數皆為 3000 筆，是相當平衡的資料。

情感分類任務將使用 precision、recall 和 F₁-score 進行評估。本論文將與 SCDL 數據集的最新系統 (NNLM) (Wang, Yuan, et al., 2014) 以及近年來使用相同資料集的 CNN-SVM (Cao, Yuhui, 2015) 方法進行比較，在 ECSR 和 URDP 數據集上使用 Naïve Bayes Classifier、XGBoost 等機器學習模型以及 BiLSTM 作為 Baseline 方法進行基準比較，結果如表 3 所示。

Bi-LSTM 模型較機器學習方法可獲得更好的效能，因為通過雙向處理文本來學習過去和將來的上下文信息。此外，並非所有單詞對文本的情感分析做出相同的貢獻，注意力機制 (Attention Mechanism) 能夠根據單詞註釋權重對句子含義的重要性來對它們進行改組，可以發現基於注意力的 Bi-LSTM(Bi-

LSTM+Att) 比沒有注意力機制的 Bi-LSTM(Bi-LSTM) 效能更高，這是因為注意力機制 (Attention Mechanism) 可以從眾多信息中選擇出對當前任務目標更關鍵的信息，進而增強模型學習的能力。從結果得出在每個類別中本研究之方法 (BLV-depCNN) 表現皆較優，在 SCDL、ECSR 和 URDP 數據集上，BLV-depCNN 模型可以分別達到 80%、94% 和 86% 的 F₁-score。使用注意力機制將情感的 Valence 與 Arousal 注入到 Bi-LSTM，也考慮了單詞與單詞間的關係，並利用 CNN 模型完整的學習文本間的單詞的兩兩關聯性，增強了情感信息和語義信息的表達能力，從而有效地增強正確識別商品評論情緒的能力。

表 1. 嵌入方法之性能比較

Data	Method	Positive	Negative
		Precision, Recall, F ₁ -score	
NLPC 2014	Naïve Bayes	0.723/0.741/0.732	0.732/0.713/0.722
	XGBoost	0.754/0.754/0.754	0.754/0.754/0.754
	Bi-LSTM	0.748/0.729/0.738	0.736/0.754/0.745
	Bi-LSTM+Att	0.741/0.772/0.756	0.762/0.729/0.745
	NNLM	0.758/0.789/0.773	0.780/0.748/0.764
	CNN-SVM	0.766/0.806/0.785	0.795/0.754/0.774
ECSR	BLV-depCNN	0.814/0.796/0.804	0.786/0.813/0.799
	Naïve Bayes	0.772/0.788/0.780	0.826/0.811/0.818
	XGBoost	0.861/0.857/0.858	0.873/0.792/0.831
	Bi-LSTM	0.858/0.824/0.841	0.862/0.890/0.876
URDP	Bi-LSTM+Att	0.852/0.842/0.847	0.873/0.881/0.877
	BLV-depCNN	0.946/0.925/0.935	0.937/0.954/0.945
	Naïve Bayes	0.861/0.783/0.827	0.797/0.867/0.822
	XGBoost	0.864/0.753/0.804	0.775/0.879/0.823
URDP	Bi-LSTM	0.864/0.779/0.819	0.802/0.874/0.836
	Bi-LSTM+Att	0.847/0.841/0.843	0.841/0.849/0.844
	BLV-depCNN	0.881/0.833/0.856	0.840/0.891/0.864

根據以上實驗結果，本研究之方法確實可以藉由提供更多詳細的情感知識來提高情感

分類器的分類有效性，在不同類型的情感分類數據集中也表現出色的性能。

5 結論與未來展望

隨著越來越多人在線上購物，顧客與購物網站之間的交流也越趨頻繁，2020年台灣網路資訊中心(TWNIC)¹⁰網路報告統計當年度59.6%的民眾有網購的經驗。顧客留下包含帶有不同情感色彩和個人語義信息的產品評論。大量產品評論已成為潛在客戶的關鍵信息來源，面對來自網際網路上大量的資訊，很難快速掌握到產品評論的關鍵重點。所以，本研究利用社群媒體上的短文本，提出一個具有細粒度情感特徵的深度神經網路模型，同時考慮詞彙情感特徵與句子結構特徵，利用LTP套件的依存句法分析短文本中詞彙間的依存特性，並整合BERT文字向量表示法以及VA預測模型訓練出具有價值的情感類別預測模型並應用於電子商務服務平台、外送平台、書籍、DVD和電子產品等領域皆獲得良好的效能。證明情感特徵可有效幫助模型學習，也顯示出本研究之貢獻。

如何妥善處理大量資料並有效地轉為知識是相當重要的議題，產品評論分析的目的在於萃取出消費者對於產品的評價，以獲得重要的資訊，不僅對於潛在消費者、企業、製造商與零售商等都是相當有價值的資訊。在產品評論分析中將眾多評論進行情感分類，包含二元分類(Binary Classification)、多元分類(Multi-Class Classification)以及基於觀點的情感分析(Aspect Based Sentiment Analysis, ABSA)。未來，可開發一個多種分析評論系統幫助企業快速掌握用戶對於產品的回饋，提供評論情感傾向以及資訊視覺化。節省閱讀與統整人力時間，快速且有效的分析萃取重要的顧客知識與商品的優缺點以提供企業有價值的資訊，針對網路平台上豐富的留言意見可以幫助企業改善產品品質、提供後續產品設計依據、規劃行銷策略以及促銷方案等，以達到更好的業績。

References

- Liu, B. 2012. Sentiment analysis and opinion mining. *Synthesis lectures on human language technologies*, 5(1), 1-167.
- Shakeel, M. H., and Karim, A. 2020. Adapting deep learning for sentiment classification of code-switched informal short text. In *Proceedings of the 35th Annual ACM Symposium on Applied Computing*, pages 903-906.
- Zhang, W., Xu, H., and Wan, W. 2012. Weakness Finder: Find product weakness from Chinese reviews by using aspects-based sentiment analysis. *Expert Systems with Applications*, 39(11), 10283-10291.
- Yenter, A., and Verma, A. 2017. Deep CNN-LSTM with combined kernels from multiple branches for IMDb review sentiment analysis. In *2017 IEEE 8th Annual Ubiquitous Computing, Electronics and Mobile Communication Conference (UEMCON)*, pages 540-546.
- Turney, P. D. 2002. Thumbs up or thumbs down? Semantic orientation applied to unsupervised classification of reviews. *arXiv preprint cs/0212032*.
- Hu, M., and Liu, B. 2004. Mining and summarizing customer reviews. In *Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 168-177.
- Agarwal, A., Biadys, F., and Mckeown, K. 2009. Contextual phrase-level polarity analysis using lexical affect scoring and syntactic n-grams. In *Proceedings of the 12th Conference of the European Chapter of the ACL (EACL 2009)*, pages 24-32.
- Sayeed, A., Boyd-Graber, J., Rusk, B., and Weinberg, A. 2012. Grammatical structures for word-level sentiment detection. In *Proceedings of the 2012 Conference of the North American Chapter of the Association for computational Linguistics: Human language technologies*, pages 667-676.
- Yu, L. C., Wang, J., Lai, K. R., and Zhang, X. J. 2015. Predicting valence-arousal ratings of words using a weighted graph method. In *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume 2: Short Papers)*, pages 788-793.
- Chang, Y. C., Yeh, W. C., Hsing, Y. C., and Wang, C. A. 2019. Refined distributed emotion vector representation for social media sentiment analysis. *Plos one*, 14(10), e0223317.
- Cheng, Y. Y., Chen, Y. M., Yeh, W. C., and Chang, Y. C. 2021. Valence and Arousal-Infused Bi-

³ <https://report.twnic.tw/2020/index.html>

- Directional LSTM for Sentiment Analysis of Government Social Media Management. *Applied Sciences*, 11(2), 880.
- Zhai, P., Huang, X., Zhang, B., and Fang, Y. 2020. Relation extraction based on fusion dependency parsing from chinese EMRs. *Scientific Programming*.
- Qidi, J. 2021. Research on Topic Mining of Medical Surgical Mask Reviews Sold on E-commerce Platform. In *Journal of Physics: Conference Series* (Vol. 1820, No. 1, p. 012182). IOP Publishing.
- Hsieh, Y. L., Chang, Y. C., Huang, Y. J., Yeh, S. H., Chen, C. H., and Hsu, W. L. 2017. MONPA: Multi-objective named-entity and part-of-speech annotator for Chinese using recurrent neural network. In *Proceedings of the Eighth International Joint Conference on Natural Language Processing (Volume 2: Short Papers)*, pages 80-85.
- Chang, Y. C., Yeh, W. C., Hsing, Y. C., & Wang, C. A. 2019. Refined distributed emotion vector representation for social media sentiment analysis. *PLoS One*, 14(10), e0223317.
- Chen, W. T., Lin, S. C., Huang, S. L., Chung, Y. S., and Chen, K. J. 2010. E-HowNet and automatic construction of a lexical ontology. In *Coling 2010: Demonstrations*, pages 45-48.
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... and Polosukhin, I. 2017. Attention is all you need. *arXiv preprint arXiv:1706.03762*.
- Moreno, J. G., Boros, E., and Doucet, A. 2020. TLR at the NTCIR-15 FinNum-2 Task: Improving Text Classifiers for Numeral Attachment in Financial Social Data. In *Proceedings of the 15th NTCIR Conference on Evaluation of Information Access Technologies, Tokyo Japan*, pages 8-11.
- Gomes, H. M., Barddal, J. P., Enembreck, F., and Bifet, A. 2017. A survey on ensemble learning for data stream classification. *ACM Computing Surveys (CSUR)*, 50(2), 1-36.
- Dong, Xibin, et al. 2020. A survey on ensemble learning. *Frontiers of Computer Science* 14.2, pages 241-258.
- Wu, Chuhan, et al. 2017. Thu_ngn at ijcnlp-2017 task 2: Dimensional sentiment analysis for chinese phrases with deep lstm. In *Proceedings of the IJCNLP 2017, Shared Tasks*, pages 47-52.
- Zhou, Xin, et al. 2017. Alibaba at IJCNLP-2017 Task 2: A Boosted Deep System for Dimensional Sentiment Analysis of Chinese Phrases. In *Proceedings of the IJCNLP 2017, Shared Tasks*, pages 100-104.
- Li, Peng-Hsuan, Wei-Yun Ma, and Hsin-Yang Wang. 2017. CKIP at IJCNLP-2017 Task 2: Neural Valence-Arousal Prediction for Phrases. In *Proceedings of the IJCNLP 2017, Shared Tasks*, pages 89-94.
- Wang, Yuan, et al. 2014. Word vector modeling for sentiment analysis of product reviews. *CCF International Conference on Natural Language Processing and Chinese Computing*. Springer, Berlin, Heidelberg.
- Cao, Yuhui, Ruifeng Xu, and Tao Chen. 2015. Combining convolutional neural network and support vector machine for sentiment classification. *Chinese national conference on social media processing*. Springer, Singapore.