

## 以社群媒體語言建構深度學習模型：

### 以「校正回歸」為例

# Constructing a Deep Learning Model Using Language in Social Media: The Case Study of ‘Retrospective Adjustment’

段人鳳\*、邱淑怡+、劉慧雯#

Ren-feng Duann, Shu-I Chiu, and Hui-Wen Liu

#### 摘要

本研究以台灣新冠肺炎期間首度出現的「校正回歸」一詞相關的臉書貼文為語料，進行人工情感分析與模型預測。我們對 6,917 筆語料進行人工標記，並將這些標記完成的語料分成 70% 和 30%，以 BERT-Chinese 之預訓練模型(pre-trained model)利用 70% 的語料進行微調機制(fine tune)，再以微調後的模型預測剩餘的 30% 語料，並加以比對人工標記和模型預測的結果，試圖從語言特徵找出兩者間差異的可能原因。研究結果顯示，在人工標註為中立的貼文中，模型有較好的預測能力，正確率達 0.81；而人工標註為正向和負向的貼文中，模型的預測能力較差，分別為 0.64 和 0.63。進一步觀察人工標註和模型預測的差異，人工標註為負向而模型預測為正向的有 0.23，乃所有錯誤之最，其次為

---

\* 國立臺東大學通識教育中心

Center for General Education, National Taitung University.

E-mail: rduann@nttu.edu.tw

+ 國立政治大學資訊科學系（通訊作者）

Department of Computer Science, National Chengchi University. No. 64, Sec. 2, Zhina Road, Wenshan District, Taipei City, Taiwan 116011, R.O.C.

Tel: 886-(0)2-29393091 ext. 88112. Fax: 886-(0)2-22341494

E-mail: sichiu@nccu.edu.tw

# 國立政治大學新聞系

Department of Journalism, National Chengchi University.

E-mail: huiwen@nccu.edu.tw

人工標註為正向而模型預測為中立的貼文，0.22。我們逐筆檢視這兩大類貼文，歸納出 7 類負向情感的語言特徵及 4 類正向情感的語言特徵。在檢視語言特徵時，研究者亦發現，由於本文所搜集之語料具有高度的公共性與政治性，僅討論貼文內容有時不易判斷意義，還需考慮貼文者身份，此亦可能影響了機器預測的正確率。我們主張，社群媒體的語言有別於當下模型訓練使用的資料集，且貼文者常常使用表情符號或標點符號來表達情感，未來我們將發展適合台灣的社群媒體語意的預測模型，以期提升模型預測的正確率。

### Abstract

This research, which used Facebook posts related to the term “retrospective adjustment” in Taiwan as the corpus, manually coded the sentiments of 6,917 posts. Randomly dividing the dataset into two subsets for training (70%) and testing (30%) and using the Chinese pre-trained BERT model as the foundation, we trained and fine-tuned the model with the training dataset and ran the fine-tuned model to predict the sentiments in the test dataset. We then compared the results of the manual coding and model prediction to explain the differences from the perspective of linguistic features. The results indicated that the model performed better for the posts manually coded as “neutral,” with an accuracy of 0.81, while the accuracies of model prediction were only 0.64 and 0.63 for the posts manually coded as “positive” and “negative,” respectively. Regarding inaccuracy, the posts manually coded as “negative” but predicted by the model as “positive” and those manually coded as “positive” but predicted by the model as “neutral” ranked the highest (0.23) and the second highest (0.22), respectively. Examining the linguistic features of the two groups of posts, we identified seven categories of linguistic features that, we claim, led to “negative” coding and four categories that led to “positive” coding. Moreover, both groups contained posts that could not be coded accurately without knowledge of the news and the Facebook account owners’ political/social inclinations, which was attributed to the posts’ high relatedness to the general public and the politics of Taiwan. Considering that the language used in social media is different from the language employed to train current models, and that Facebook users frequently use punctuation marks and emoticons to express their moods, we argue that there is a need to develop a model for social media.

**關鍵詞：**社群媒體、深度學習、校正回歸、情感分析、自然語言處理

**Keywords:** Social Media, Deep Learning, Retrospective Adjustment, Sentiment Analysis, Natural Language Processing

## 1. 緒論 (Introduction)

社群媒體(social media)的發展,自 2006 年起,進入一個全新的樣貌。包括 Plurk、Facebook, 以及 Twitter 在內,透過發展出能夠由使用者自行建立的「朋友清單」(friends' list),以及更為簡便的貼文介面,人們透過社群媒體發表意見,成為 21 世紀重要的媒體活動。2007 年,第一代 iPhone 手機問世,1992 年即已在商展上展出的智慧型手機從此進入商轉。隨著網路系統建置覆蓋率的提升,行動通訊徹底改變媒介使用的時間性,人與人之間的意見交換更為迅速。同時,也因為貼文門檻的大幅降低,使得社群媒體上刊載了大量由一般人表達的意見。Keen (2007)認為這些由「業餘者教派」產製並分享的意見會扼殺我們的文化價值,主張不該放任「任何人」對任何主題隨意發表評論。

然而,不論論者是否同意,網際網路基礎架構,加上降到極低的參與門檻,再加上行動通訊載具三方的共同發展,促成的全民貼文、全民分享行為,已經是不爭的事實。光是社群媒體龍頭臉書(Facebook)在 2021 年 1 月底發佈的報告就指出,即使在劍橋分析(Cambridge Analytic)公司爆發個資外洩事件之後,臉書的社會聲望下挫,但 2020 年 12 月底,每日仍有 18.4 億個活躍帳號;尤其,最愛使用臉書的地區,分別是拉丁美洲與東南亞(Facebook, 2021)。

在這個大規模的媒體使用活動中,產生了大量「使用者產生內容(user-generated content, UGC)」。這些內容透過朋友清單迅速地將個別意見傳送給特定社群網絡,形成類公共論壇(public sphere)。是以,自從 2010 年之後,掌握並分析社群媒體上的材料,就成為了解一般人意見、態度,乃至於輿論風向的關鍵。由於 UGC 的生產快速而量大,因此,研究者必須透過資訊工具,才有可能迅速掌握內容,同時也才能進行資料清洗與分析等工作。其中,最常見的就是語料分析方法。

本研究以新冠肺炎期間,環繞著指揮中心於 2021 年 5 月 22 日開始使用的「校正回歸」一詞所產製的臉書貼文為語料,先進行人工情感分析,再將完成人工標記之語料切分為 70%和 30%,前者為訓練資料集(training dataset),後者為測試資料集(test dataset)。以前者訓練 BERT-Chinese 模型,再以該預訓練模型預測後者的情感,並比對人工標註和模型預測的結果,以期從人工標註和模型預測的差異找出不同情感向度(負向和正向)的語言特徵,以為後續模型發展提出我們的建議。

本研究主要回答下列兩個問題:

- 一、人工標註和模型預測之差異為何?
- 二、負向/正向情感的語言特徵為何?

本文第 2 章回顧自然語言處理與社群媒體語言特徵等相關研究;第 3 章說明本研究採用的語料、標記方法與使用的模型;第 4 章陳述語料分析結果與討論,第 5 章則說明本研究的結論、限制與未來方向。

## 2. 文獻回顧 (Literature Review)

### 2.1 自然語言處理 (Natural Language Processing, NLP)

自然語言處理是人工智慧(artificial intelligence)與語言學的結合領域，此領域探討如何處理及運用自然語言。自然語言處理分為認知、理解、生成三個步驟：認知和理解是讓電腦把輸入的語言變成有意義的符號與關係，再依據目的進行處理，最後利用演算法生成最後的結果。近年來電腦硬體設備效能不斷地提升，2014 年由英國 Google DeepMind 開發的 AlphaGo 人工智慧圍棋軟體，它由 Silver *et al.* (2016)訓練，為第一個無需讓子之下擊敗圍棋職業九段棋士的電腦圍棋程式。AlphaGo 採用兩個深度學習(deep learning)的類神經網路(neural network)及搜尋機制來選擇落子，並進一步使用強化學習(reinforcement learning)加以改善。在這股風潮引領之下，深度學習再度成為熱門話題，透過類神經網路的多層架構的學習機制，讓模型更接近人類大腦可以自發學習進行訓練。這樣的技術也應用於文字的語意分析(semantic analysis)，運用大量閱讀文本資料後，可進行語意分析將文本資料區分類型、對文本資料進行摘錄、或進行文本內容的預測，例如：讓電腦閱讀大量的金庸武俠小說後，模型便可寫出另一本武俠小說。

然而並非每位研究者都能取得大量文本進行模型建構，因此，Devlin *et al.* (2018)提出自然語言處理的預訓練的技術，由 Google 於 2018 年發布的基於變換器的雙向編碼器表示技術(bidirectional encoder representations from transformers, BERT)，這是一種預訓練語言的方法，Devlin *et al.*以大量文本語料庫（如：維基百科）訓練一個通用的語言理解的模型。BERT 的模型一開始以英文文本進行預訓練，之後由 Cui *et al.* (2019)等人發展出以中文訓練的 BERT-Chinese，他們採用大量的中文文本建立模型，提出的模型能處理 NLP 的中文句子和文本。本研究將利用 BERT-Chinese 預訓練的模型，再以我們的資料訓練模型，讓預訓練模型進行微調(fine-tuning)機制，建構分析臉書貼文情感之模型。

情感分析向來受到自然語言處理學者關注，社群媒體的貼文提供了這些學者豐富的語料。Roberts *et al.* (2012)探討 Twitter 上推文的情感分類，他們蒐集了 14 個主題的推文，如情人節、2010 世界盃、2012 美國總統大選、琳賽蘿涵…等，將憤怒、厭惡、恐懼、喜悅、愛、悲傷和驚訝等七種情緒，分別歸類為正向、負向及未預期(unexpected)三大類型<sup>1</sup>。Nissim & Patti (2017) 認為語意層面至關重要，他們回顧了自然語言學界發展之情感分析的語意資源：除了帶有正／負向標記或情緒標記的詞典，也包含在概念層次的情感／情緒分析上，呈現與多字詞語(multiword expressions)相關的語意、概念和情感訊息。他們認為，單字層次的分析顯然已不足以判定發言人的情感，因此出現將分布語意學(distributional semantics)－在相同前後文出現的字詞應有類似的意義－作為情感分析的判斷依據。Tang *et al.* (2014)提出神經網路來學習能將正／負情感編入字串(n-gram)的特定情感詞嵌入(sentiment-specific word embedding, SSWE)，並以推特貼文訓練機器。以屬

---

<sup>1</sup> Roberts *et al.* (2012)的正向類別為喜悅和愛，負向類別為憤怒、厭惡、悲傷和恐懼，未預期類別則包含恐懼與驚訝。

性為基礎的情感分析(aspect-based sentiment analysis)需要辨識特定實體，以及實體與事件之間的關聯，Socher *et al.* (2013)推出了建置於文法結構之上的回歸神經網路模型(Recursive Neural Tensor Network)，並以包含 20 多萬短語的細緻情感標籤的情感結構樹資料庫(Sentiment Treebank)來訓練這個模型，Socher *et al.*主張這個模型較以往方法能更正確地判斷單句的正／負情感，並能更正確預測短語的細緻情感標籤。

在探討 COVID-19 和社群媒體的互動方面，Wang *et al.* (2020) 以新浪－微博(Sina Weibo)上與 COVID-19 相關貼文為資料集，採用無監督 BERT 模型對情感類別（正向、中立和負向）進行分類，並使用「詞頻－逆文檔頻率(term frequency-inverse document frequency, TF-IDF)」模型匯總貼文的主題，再進行趨勢分析和主題分析，以識別負向情緒的特徵。Lu *et al.* (2021) 也依據微博上 COVID-19 主題的貼文進行情感分類，有別於 Wang *et al.* (2020)，Lu *et al.* 將情感分為對某特定對象的「批判」與「支持」兩類。利用 BERT 模型微調的技術，他們探討在社群媒體上公眾情緒(public sentiment)如何隨著 COVID-19 的傳播而演變，進而預測貼文的情緒分類。Singh *et al.* (2021)則探討 COVID-19 對社會生活影響的情緒分析，因疫情而出現的社交距離，促使人們迅速地轉向在社群媒體上表達意見與渴望，該論文以 BERT 模型對 Twitter 上的推文進行情感分析，並將情感分類為正向、負向和中立，以探究人們的心理狀態。

Jain *et al.* (2022)以 BERT 模型為基礎，加入卷積神經網路(convolutional neural networks, CNNs)，提出新的模型「BERT 擴張卷積神經網路(BERT dilated convolutional neural networks, BERT-DCNN)」，該模型包含情感知識庫，建構概念性層級(concept-level)的情感分析，他們分析美國國內及國際航空公司的消費者評論，評論內容亦分為正向、負向及中立三種類別。

上述文獻顯示，無論是社群媒體的推文或是消費者評論，多數研究採用正向、負向及中立作為語意分析的分類，本研究亦將使用正向、負向及中立作為情感類別。

## 2.2 社群媒體的語言特徵 (Linguistic Features on Social Media)

基於使用者產生內容的特性，臉書上的互動吸引了語言學－特別是言談分析－學者的目光。Tannen (2013)探討電子媒體上的語言使用，她主張媒體的選擇本身便傳達了「後設訊息」，即「『言談的主題即對話者之間的關係(‘the subject of discourse is the relationship between the speakers’)』…說話者如何傳達訊息，而聽話者如何解讀訊息」(p. 101)。Tannen 主張包含臉書、即時通訊等新媒體上的互動與口語對話非常雷同，使用者之大寫字母、重複驚嘆號或問號、重複單字或母音字母等「熱忱標記(enthusiasm marker)」，類似口語對話中的提高語調、增強情感，以及拉長音節等情緒表達；使用者轉貼連結或簡短回覆，類似口語表達中說話者不直接傳達的訊息；回應速度則類似口語交談中的對話步調與停頓…等，這些特徵皆凸顯新媒體上的溝通與口語對話極為相似。Maíz-Arévalo (2015)研究了英文與西班牙文臉書上朋友間的嘲弄(jocular mockery)，發現兩國的臉書使用者最常使用拼字變化（如表情符號、重複字母、重複驚嘆號或問號、模擬他國語言腔調等）來嘲弄貼文作者，也透過「公式化詞彙（如驚嘆詞、習慣用語等）」、反諷、轉移話題，誇

張詞彙等來傳達嘲弄。Tabe (2016)探討喀麥隆境內英語使用者在臉書、即時通訊等社群媒體上的幽默，她歸納出 11 類幽默的表達：(1) 違逆 Grice 準則、(2) 拼字變異（如重複字母或單字、玩弄同音字等）、(3) 音效（如狀聲詞、頭韻或尾韻、共鳴等）、(4) 重複標點符號、(5) 幽默軼事（如新聞或故事）、(6) 笑聲、(7) 玩弄身分（將傳統命名轉移到線上互動、調整真實姓名使之看起來新潮，或縮寫姓名）、(8) 表情符號（包含現成的表情符號和特殊標點符號或字母）、(9) 誇飾、(10) 隱喻和(11) 其他（如無規則可循的符號組合）等。Ye (2019)則聚焦中文社群媒體上，「被」字句的創新用法及其傳達的諷刺意味，透過語言、文化和社會分析，作者指出：傳統被字句可不提及施事者，及其凸顯受事者受負面影響之特質，讓使用者能嘲弄未被指名的施事者／權威者，並表達使用者對施事者／權威者的負面態度。

學者也採批判言談分析的角度探討臉書互動如何形塑、維持、強化內部團體(in-group)和對抗外部團體(out-group)。Morin & Flynn (2014)探討美國的新興政黨－茶黨－支持者如何在 2010 年「催票周末(“Get Out The Vote” weekend)」期間運用臉書來建構、維持身分認同。他們研究了三位雖隸屬於共和黨、但強調茶黨身分的參議員候選人的官方臉書頁，並且發現支持者的回饋可分為兩大主題：攻擊與鼓勵，即攻擊威脅，並鼓勵與候選人團結一致。攻擊的語言包含負面評價用語、貼標籤（或譯為「咒罵法」）、誇飾法等，這些都是藉由攻擊外部團體來堅定內部團體的身分認同；鼓勵的語言包含支持者使用準社會接觸(para-social contact)與候選人互動，例如在臉書上直接與候選人對話，使用 Tannan 提出的熱忱標記，否認選情落後，為候選人辯護等。Al-Tahmazi (2015)探討伊拉克臉書上對政治人物與事件的（再）定位如何使政治討論兩極化。透過分析評論者（即臉書使用者）在臉書上如何將政治人物與其作為正當化／去正當化，評論者將政治討論從議題導向轉變為人物導向，同時，評論者也在與其他評論者的互動中，建構了自己的社會－政治認同，並根據言論，將自己與其他評論者放入對立的線上社群，即內部團體和外部團體；換句話說，臉書使用者之間的線上互動為實際生活的延伸，他們臉書上的政治認同和他們實際生活上的種族、教派或文化群體的屬性交互影響，形成互為敵對的政治陣線。Chibwe & Ureke (2016)聚焦辛巴威境內兩敵對陣營在 2013 年的選舉中如何攻擊彼此，探討臉書自由來去且匿名的特性如何讓兩敵對陣營攻擊彼此，且導致兩極化，他們主張，即使臉書提供了另類公領域，這個公領域也可能被缺乏理性的辯論粉碎。

### 3. 研究方法 (Methodology)

#### 3.1 資料搜集與標記 (Data Collection and Coding)

為了在一定範圍內針對社群媒體使用者的語言使用型態進行較為細緻的討論，同時考慮解釋本身的有效性以及可概推能力，本研究選擇臉書為搜集語料之平台。我們利用臉書官方資料蒐集器 CrowdTangle，擷取 2021 年 5 月 22-25 日（共計 4 日）公開資訊中使用或提及「校正回歸」字樣的貼文（包括：粉絲專頁、公開社團與個人帳號的公開貼文）共計 6,917 筆，交由三位標記者標註「正向」(Positive, 以下簡稱 P)、「負向」(Negative,

以下簡稱 N)，以及「中立」(Neutral, 以下簡稱 Neu)三種記號之一。這三位標記者中，第一、二位為傳播背景，第三位為語言學背景。

本文所稱標記，是針對 CrowdTangle 擷取檔案中，登記為 text/message 的欄位進行判讀。我們的標記分為兩階段，第一階段為三位標記者分開標記，我們參考新聞學的「5W1H」(who, when, where, what, why, how)表示事實資訊的主要面向，若貼文僅呈現 5W1H，我們便標記為中立<sup>2</sup>；我們亦參考陳韋帆、古倫維(2018)發展的中文情感語意分析套件，判斷詞彙的情感向度，若貼文表達正向情感者（如：讚揚、感謝、希望、改善、積極作為…等），我們標註為正向，反之則標註為負向（如：指責、驚嚇、失望、恐懼、擔憂…等）。第二階段，則由第三位標記者審視三位標記者的結果，並給予貼文最終判斷。我們的情感標記採多數決，當兩位或三位標記結果相同時，即確定該貼文的標註結果。

本研究以「校正回歸」為個案，除了因為研究進行期間正值 COVID-19 疫情在台灣首度爆發大規模本土感染，中央疫情指揮中心(CDC)宣佈全國進入三級警戒狀態，相當程度限制外出的情況下，正是社會人心惶惶，一般人對疫情資訊有著高度需求的時刻之外，我們也觀察到「校正回歸」一詞在 5 月 22 日 CDC 例行記者會上首度從衛福部長陳時中嘴裡說出時，舉國嘩然。當天之後，不論是電視上的政論性節目、社群媒體上的意見領袖(key opinion leader)，或者一般庶民，都出現了許多對此詞彙的闡述、評論與解讀。這顯示，這個詞彙雖有學理上的定義，但在市民社會中，卻呈現多樣多元的闡述意義。不僅如此，由於臺灣社會剛經歷 2018 年底九合一大選、2020 年初總統大選、2021 年公民投票等事件，整個社會處於高度社會政治化的狀態中。這個狀態使得「疫情治理」這個公共衛生事件，也成為政治攻防的關鍵場域。由此，我們可以預見中央疫情指揮中心的資訊，也容易政治化。在此時對「校正回歸」這個關係染疫人數的詞彙進行討論，處理的也就不只是一個統計詞彙，更是一個雜揉了政治與社會輿論氣候的詞彙。藉由對這個詞彙的資料搜集、標記與自然語言處理，我們可以透過在社會生活中呈現多義狀態的詞彙，一窺人工標記／解讀，與自然語言處理「解讀」同一批測試資料集的異同，並進一步提出對社群媒體使用者語言實踐和（繁體中文）自然語言處理的調校方式。

### 3.2 自然語言處理模型 (The NLP Model)

本研究基於自然語言處理方式，採用遷移學習的技術，以 BERT-Chinese 作為模型架構，該模型利用 Wikipedia 上（包含繁體中文及簡體中文）的文章，有 2.1 萬個詞彙，參數數量約 97M (Cui *et al.*, 2019)。本研究藉由訓練和微調技術建構屬於臉書議題的語意模型，

---

<sup>2</sup> 在新聞學上，以「5W1H」(who, when, where, what, why, how)表示事實資訊的主要面向，在純淨新聞寫作中，聚焦這六個面向，可以簡短但精準傳達新聞事件的核心內容，提供讀者必要的資訊。因此，5W1H 乃是新聞必要內容。本文將僅涉及 5W1H 的貼文視為中立，因為這表示貼文有、且僅有傳遞資訊的能力。同時，既有的文獻顯示，若欲發展機器下標、撰寫導言等智能模型，也經常以 5W1H 做為模仿的架構（參見：鍾文翔，2018）；除了因為範疇相對明確，如此訓練出的模型，在生成語句或段落時，更加接近人所撰寫的文字。

最後透過我們的模型進行預測分類，將這些內容進行語意分類建構正向、負向及中立的類別。在 6,917 筆完成人工標記的語料中，先排除三位標記者互不同意的筆數，將語料隨機切分為 70% 和 30%，前者作為訓練資料集(training dataset)，後者做為測試資料集(test dataset)。換句話說，我們的模型以 BERT-Chinese 預訓練模型為基礎，運用我們人工標註的 70% 進行該預訓練模型的微調(fine tune)機制，接著，再將剩餘的 30% 交由模型預測情感分類，再對比人工標註與模型預測的結果，以評估我們的模型效果。

#### 4. 資料分析結果 (Results)

##### 4.1 標記結果比對 (The Comparison between the Manual Coding and Model prediction)

在測試資料集中，我們將人工標註的結果當作基準真值(ground truth)，與模型預測的類別進行比較（詳見圖 1）。圖 1 顯示各分類之下的貼文數量，其中，藍色（深色）表人工標註的結果，橙色（淺色）表模型預測的結果。對於負向(N)貼文可以看出人工標註的數量大於模型預測的結果，然而，正向(P)和中立(Neu)的貼文數量反而是模型預測的數量大於人工標註的數量；整體而言，無論是人工標註或模型預測的結果，皆是中立(Neu)貼文數量最多。

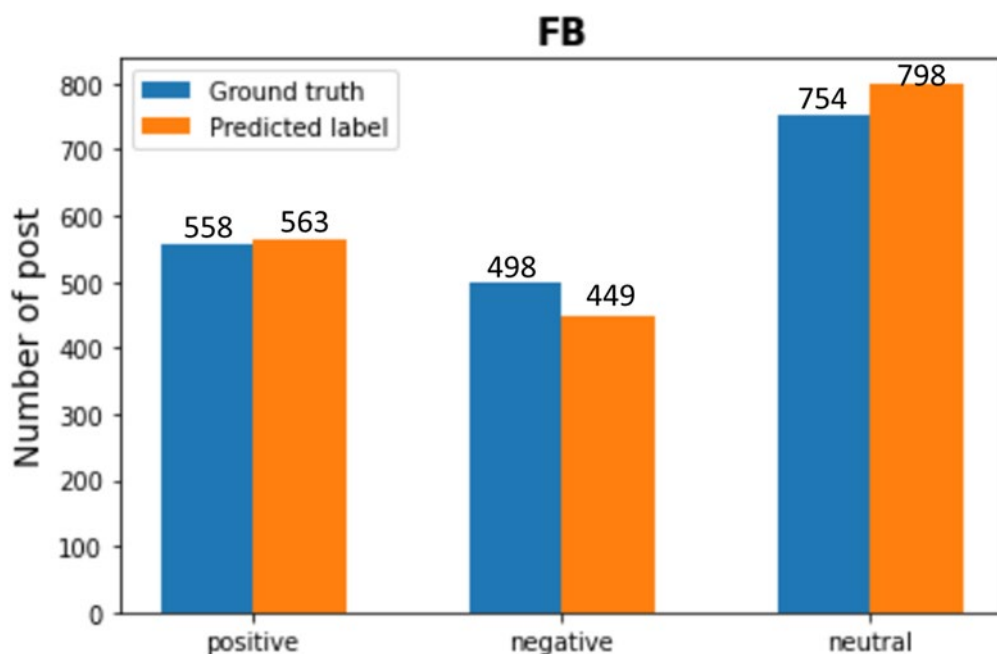


圖 1. 人工標註與模型預測於各類別貼文的數量

[Figure 1. The number of manually coded posts and model predicted posts



透過圖 1，我們發現人工標註為正向與模型預測為正向的貼文差異最少，人工標記為正向共 558 筆貼文，模型預測為正向共 563 筆貼文，差異數量對應人工標註的比例僅 0.9%<sup>3</sup>，顯示正向貼文預測效果最好；而人工標註為負向與模型預測為負向的貼文差異最多，差異數量對應人工標註的比例達 9.8%。為進一步探討這些差異的問題，我們利用混淆矩陣(confusion matrix)來評估語意分析的模型（詳見圖 2）。圖 2 縱軸為人工標記，橫軸表模型預測，軸線上的 P 表示正向、N 表示負向，而 Neu 表示中立。透過混淆矩陣可以看出人工標註和模型預測於各類別貼文的正確率及錯誤率。

就正確率而言，圖 2 顯示，在人工標註為正向貼文中，模型預測為正向的比例為 0.64<sup>4</sup>；人工標記為負向的貼文中，模型預測也為負向者為 0.63；人工標記為中立的貼文中，模型預測也為中立者達 0.81，為這三種情感分類正確率最高的數值。

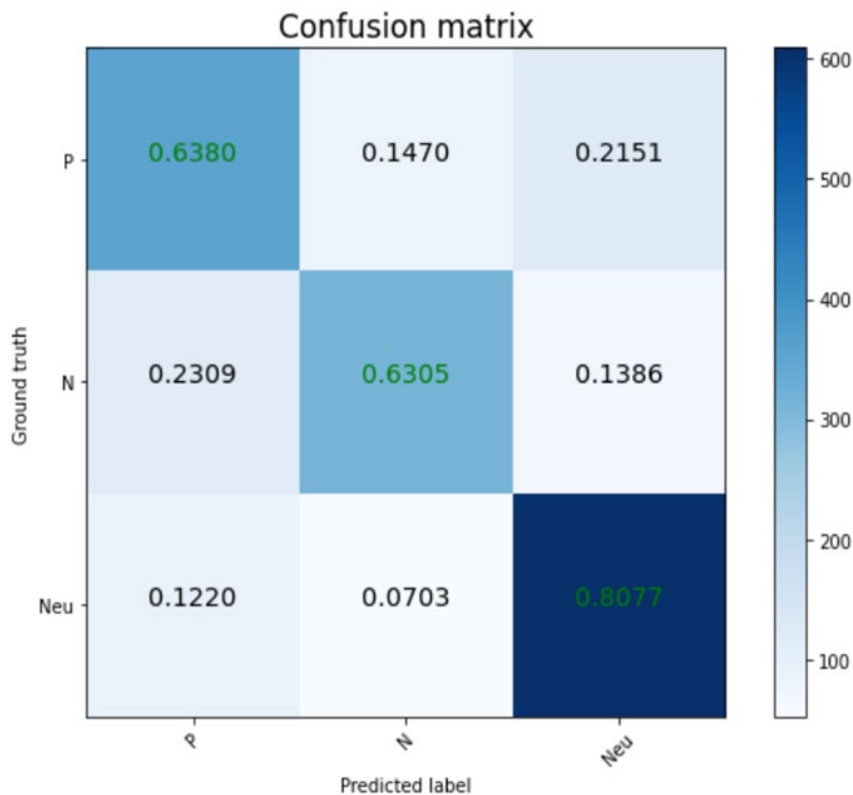


圖 2. 混淆矩陣  
[Figure 2. Confusion matrix]

<sup>3</sup> 該比例計算方式為： $(563-558)/558=0.896\%\approx 0.9\%$ ，人工標註為負向與模型預測為負向的貼文差異比例計算方式亦同，即 $(498-449)/498=9.83\%\approx 9.8\%$

<sup>4</sup> 圖 2 的數字我們四捨五入取至小數點後第二位，因此，人工標記為正向而模型預測為正向的 0.6380 我們四捨五入為 0.64，其他數字亦同。

就錯誤率而言，人工標記為正向的貼文中，模型預測為負向或中立的比例分別為 0.15 和 0.22；而人工標記為負向的貼文中，模型預測為正向或中立的比例分別為 0.23 和 0.14；人工標記為中立的貼文中，模型預測為正向或負向的比例分別為 0.12 和 0.07。錯誤率中，人工標記為負向而模型預測為正向者最高，其比例為 0.23；錯誤率最低者為人工標記為中立而模型預測為負向，其比例為 0.07。

綜上所述，人工標記為負向而模型預測也為負向的正確率最低，僅 0.63，其次為人工標記為正向且模型預測也為正向者，0.64，這顯示該模型對於負向和正向情感貼文的判斷能力較差。人工標記為負向的貼文中，有 0.23 模型預測為正向貼文，為所有錯誤率最高者；錯誤率次高者為人工標記為正向而模型預測為中立的 0.22，這兩個現象顯示，人工判斷屬於負向或正向情感的貼文，預訓練模型卻預測為正向或中立情感。這將讓我們進一步探討箇中原因。

## 4.2 語言特徵 (The Linguistic Features)

圖 2 顯示，人工標記為負向的貼文中，模型預測為正向（以下簡稱「人工負－模型正」）的比例為 0.23，為所有錯誤率最高者，其次為人工標記為正向而模型預測為中立（以下簡稱「人工正－模型中立」）的 0.22。我們因而聚焦於這兩個部分。排除三人標記不一致的筆數，以及第三位標記者再次確認內容所傳達的訊息後認為應調整情感分類的貼文，我們檢視並分類負向貼文和正向貼文的語言特徵，以利為模型後續發展。

我們逐筆檢視，試圖透過人工標註和模型預測之間的差異來歸納特定情感的語言特徵。我們一方面認同 Nissim & Patti (2017) 的看法，認為詞彙層次的分析已不足以判斷貼文的情感，而應納入跨詞訊息（如句法）、非詞訊息（如標點符號和表情符號）及言談訊息（如類比）等；另一方面，我們主張應考量跨欄訊息（如貼文者身分）以及時事知識，方能較完整取得貼文的情感樣貌。

我們雖然設定「校正回歸」為關鍵詞，但基於臉書使用者不同目的和特性，這些語料大部分聚焦於校正回歸相關新聞，但仍有一部分圍繞其他主題，例如股市、旅遊、購物、美食…等，我們一併納入討論。我們先分別討論負向情感和正向情感之語言特徵，然後再比較兩種情感語言特徵的異同。

### 4.2.1 負向情感語言特徵 (Linguistic Features of Negative Polarity)

我們將負向情感的語言分為（甲）－（庚）類：（甲）貼文使用負向評價語言，（乙）貼文使用標點符號、表情符號或語尾助詞傳達負向情感，（丙）貼文使用類比，（丁）貼文訴諸其他國家，（戊）貼文使用反語或反串文，（己）貼文引述典故、文學作品、電視劇、電影等，（庚）貼文變造口號。此外，我們也發現有些貼文需要跨欄知識或對時事的了解才能正確判斷情感，我們將之歸類為（辛）。我們對這 8 類分別說明、舉例如下，並以底線標示屬於該類型的語言特徵。

（甲）貼文使用負向評價語言：這一類貼文使用負向評價語言（施孟賢等，2021）表達負向情感，參考施孟賢等(2021)，我們認為正向語言可以是單詞、片語或子句。由於負向語言有不同的形態，我們又分為下列（甲1）－（甲6）等次類：

（甲1）使用指責、批評、否定或不信任的語言：對象可能是政府、政治人物或某些政治人物的支持者，如【例1】明確批評指揮中心創造新名詞的作法可能使民眾更加驚恐：

【例1】#點新聞 都已經疫情這麼嚴重了，指揮中心還在那邊 #創造新名詞，這樣會讓民眾更恐慌吧！

（甲2）引述、討論名人／專家：這類貼文多半引述名人／專家批判性的言論，或報導名人／專家的作為導致負向結果，或以主題標籤「#」標註名人／專家，如【例2】引述媒體人李艷秋批評指揮中心作為的言論，且使用負面言說動詞「開嗆」；【例3】則直接使用「被罵翻」來說明陳時中提出「校正回歸」一詞所遭遇到的反彈：

【例2】李艷秋開嗆：辛苦陳時中了，每天記者會還要提供娛樂表演，反正大家防疫靠自己，就當宅在家的消遣吧！[...]<sup>5</sup>

【例3】[...]幕後 | 陳時中「校正回歸」被罵翻 [...]

（甲3）貼文質疑政府、政治人物或某些政治人物的支持者：包含一般問句（如【例4】）和修辭性問句（【例5】）：

【例4】[...]重點是，除了 321 例外又有 400 例，而這 400 例是因為檢驗的時間差造成的，但為何不直接每天滾動式調整而是一次公佈調整？

【例5】數字可以校正回歸，但這世界還能校正、回歸到疫情前的生活樣態嗎？

（甲4）貼文使用貼標籤法（或譯為「咒罵法」）：如同施孟賢等(2021)主張，標籤可以是人物的標籤，如【例6】之「巨嬰」，將小聖蚊和藍白粉（即國民黨與民眾黨的支持者）直接稱呼為「巨嬰」，表示把這些人「將原本自己該負的責任，移轉給國家承擔」（呂秋遠，2018）；標籤也可以是政策或事件的標籤，如【例7】，指涉鼓吹政府施打疫苗的倡議為「疫苗業務大會」：

<sup>5</sup> [...]表示由本文作者刪節的貼文部分；若是貼文者自行使用的刪節號，則以…呈現。

【例 6】我真的不希望台灣巨嬰太多 小聖蚊跟藍白粉這些咖小 就跟確診人數一樣、拜託不要再增加了。

【例 7】今天應該不會有人再去討論校正回歸 我想今天最大的擂台賽應該是「滯台中國人的疫苗業務大會」

(甲 5) 貼文直接使用表達負向情感的文字：包含痛苦、失望、絕望、恐懼、驚嚇、擔憂...等，如【例 8】：

【例 8】恐怖！連 8 日破百！本土確診+321

(甲 6) 貼文表達不理解或猜測：貼文使用表達無法理解或猜測的文字，如【例 9】中的「滿臉問號」：

【例 9】今天學到一個新的詞「校正回歸」，然後大家看著指揮官滿臉問號。

(乙) 貼文使用標點符號、表情符號或語尾助詞傳達負向情感：這一類的貼文，類似 Tannen (2013) 提出的「熱忱標記」，即：使用（有時重複出現的）標點符號、表情符號和語尾助詞等，來強化無奈、驚嚇、崩潰或不信任...等負向情感。標點符號包含刪節號(...)、波浪紋(~)和驚嘆號(!)，【例 10】在表達負向情感的文字「懶得回應」後使用刪節號表達欲言又止的負面情緒，可能是無奈；表情符號包含 QQ、Σ(°Д°;)、😭<sup>6</sup>...等，【例 11】在發現台東縣也出現確診案例後，使用表示流淚的 QQ 加強無奈或絕望等情緒；語尾助詞包含「啊(阿)」、「啦」、「吧」、「耶」等，【例 12】若單純看「台北市的量也太多了」可能難以斷定情緒方向，但在同時使用語尾助詞「吧」和刪節號後，便引發出無奈的負面情感。

【例 10】小編懶得回應了…… 反正大家都知道會有這種「現象」就好了……  
P.s. 1.如果真的「來不及」就開始好好講「來不及」，並且有因應作為，而非搞了好幾天，再來一個「校正」，並且甩鍋給地方。[...]

【例 11】台東也淪陷了 QQ

<sup>6</sup> 手機／Google 表情包的表情符號無法於 Word 檔案中顯示，我們以註腳說明，這個符號是哭臉。

【例 12】#懶編：台北市的量也太多了吧....

(丙) 貼文使用類比：貼文使用隱喻或「情節／劇本(scenario)」(Musolff, 2004, 2006)<sup>7</sup>，表達對於防疫措施或「校正回歸」一詞的負向情感，如「塞車」、「作帳」、「捅一刀」…等，這些隱喻或情節／劇本皆帶有負向含意：【例 13】中的「內外帳」隱喻，來自公司內部帳冊和對外公告帳冊的不一致性，指涉指揮中心對國人隱藏真實確診數字的作為；【例 14】則使用「黑洞」表示大眾無從得知真實的確診數字，指責地方政府防疫採檢通報量的不實；【例 15】使用路人指揮消防員滅火的情節／劇本，來表達對某政黨支持者的不滿。

【例 13】#這是內外帳的差別 台灣統計新冠肺炎確診人數，22 日本土病例 321 例，境外傳入 2 例，校正回歸 400 例。向世衛通報直接大方寫上「+723 人」。23 日對內宣佈：本土病例 287 例，境外傳入 3 例，校正回歸 170 例。向世衛通報，是「+460 人」。

【例 14】今天應該關注的重點還是新北的確診數吧，侯友宜轄下的新北不只是防疫漏洞，根本就防疫黑洞了

【例 15】要搞鬼或想算帳，請疫情結束再來，你們現在就像是火神的眼淚裡面一樣，兩個消防員因為專業吵就算了，至少我們知道他們都在努力想辦法解決，在想辦法找出大方向最好的方法。結果路人還跳出來想指揮消防員滅火，要朝哪噴水才正確，也只是想讓人家以為你很厲害而已[...]

(丁) 貼文訴諸其他國家：這類貼文比較台灣和歐美日等國的防疫措施，從而表達對台灣防疫政策的不滿，如【例 16】；或如【例 17】說明他國校正數字並非將確診人數往上加，而是削減確診人數，以糾正台灣對於「校正回歸」的用法：

【例 16】楊艾俐：企業要買快篩有何罪？指揮中心不同意，太不合理了。在美國，甚至個人都可以到藥房去買，別說是企業，在西雅圖的微軟及其他高科技公司早在去年五月，就已經每天快篩要進去的員工，這才降低西雅圖的感染率。美國屠宰場一度傳出有很多確診，後來凡上工的都要在外面快篩，才能進去上班，Amazon

<sup>7</sup> Musolff (2004, 2006)主張，情節／劇本為不同場景(scene)或故事線的集結，並且該語言使用者可理解這個情節／劇本中的角色、故事線，以及結果，並可判斷情節／劇本成功與否、正常與否，及／或正當與否等。

也如是，警察、醫護人員都是每天測，現在美國已測試 4 億多劑，平均每人測試 2 次。台灣測 20 餘萬劑，平均 100 人才測試一劑，國王的新衣是該脫下了。

【例 17】和大家分享，其實法國昨天才因應「錯誤重複計算」，「校正」#消除了 35 萬例確診數喔。

(戊) 貼文使用反語或反串文：根據 Wilson (2017: 202) 的研究，反語(irony)傳統定義為「說一件事，但卻表達反義」。Wilson 認為反語應視為 Sperber & Wilson (1981) 提出的「回響說明(echoic account)」：反語使用者並非說出與字面意義相反的語言，而是針對某一個體或群體，反應出自己的想法，並嘲笑、蔑視這個想法。Wilson (2017: 202) 舉例說明，當一個人說「政客絕不會撒謊」時，他並非斷言政客有時／總是撒謊，而是針對人們對政客懷抱「絕不撒謊」的期待表示輕蔑與嘲諷。Wilson 因而認為，反語的重點並非言談的內容，而是對該內容表達的態度。如【例 18】「博大精深」原為褒義詞，義為「廣博深遠」（教育部重編國語辭典修訂本），但出現在「300+400 校正回歸…」之後，凸顯貼文者預期指揮中心對防疫和篩檢應有更佳作為，然而指揮中心卻提出「校正回歸」一詞的失望心態，因而使用反語來表達其態度。再者，這類貼文亦可視為反串文，即貼文者本身並非具有某種身分或立場，但卻假裝自己是該身分立場的人來發言，以達到某種目的。常見的反串包括各種常成為爭吵（戰文）的身分，例如政治立場（藍綠）、支持球隊、性別、學校、職業、南北地域、類組等等。（PTT 鄉民百科）。

【例 18】300 + 400 校正回歸... 再次說明中文的博大精深

此外，【例 19】中的「真給力」出現在「數字遊戲」之後，同時利用「校」可讀為ㄊ一ㄠˋ的特性，造出「笑症回歸」一詞，下接「看好戲」，這首打油詩的貼文者明顯不認同指揮中心提出「校正回歸」的做法，並以反語「真給力」加以嘲諷：

【例 19】校正回歸四百例，數字遊戲真給力，時中防疫不努力，笑症回歸看好戲。

(己) 貼文引述典故、文學作品、電視劇、電影等：這類貼文有時直接引述帶有負面意涵的典故或作品，如【例 20】中「國王的新衣」指涉指揮中心在防疫作為上的自我欺瞞，【例 21】的「沐猴而冠」則直指陳時中的牙醫專業無法勝任防疫指揮官一職；有時引述的作品或典故並無特別的正負向意涵，需要透過前後文（如，「笑噴」）來判斷，如【例 22】：

【例 20】[...]台灣測 20 餘萬劑，平均 100 人才測試一劑，國王的新衣是該脫下了。  
[...]

【例 21】[...]憑恃政治正確的政治人物，沐猴而冠，擔任防疫指揮官，絕無成功抗疫的可能。[...]

【例 22】台版 TENET 天能 校正回歸成流行語！😂 引爆網友笑噴🔥 有人說蓋牌就蓋牌嘛，XD！

（庚）貼文變造口號：疫情初始，政府推出「有政府，請安心」口號<sup>8</sup>，其後演變為「有政府，好安心」，而【例 23】結合「作帳」隱喻，變造了這個口號，來表達負向情感：

【例 23】果真有政府，會做帳，防疫成果好棒棒。

（辛）需要跨欄知識或時事認知才能正確判斷的貼文：這類貼文看似表達正向情感，實際上需要跨欄知識（如，貼文者或粉絲團的政黨立場）或對時事有所認識才能適切判斷貼文的情感。【例 24】單就貼文內容而言，看起來是善意提醒，然而貼文後方以主題標籤「#」標註台中市長盧秀燕、她的小編以及「烏龍」一詞。這篇貼文背景為，盧秀燕的小編在 5 月 22 日「校正回歸」一詞公布後，以盧秀燕的帳號發表「陳時中剛公布的 400 例 校正回歸 什麼鬼？」貼文，引發網友爭議。因此，這份看似善意提醒的留言，實為挖苦，而這一層的解讀，需要對於貼文者身分和時事的認識方能了解。

【例 24】小編...下次記得換帳號喔🙅（#蹲路編） #盧秀燕 #直播 #小編 #烏龍

以上為負向情感語言特徵，下個章節我們來探討正向情感語言特徵。

#### 4.2.2 正向情感語言特徵 (Linguistic Features of Positive Polarity)


我們將正向情感的語言分為 (1)–(4)類：(1) 貼文使用正向評價語言，(2) 貼文使用標點符號、表情符號或語尾助詞傳達正向情感，(3) 貼文使用類比，(4) 貼文訴諸其他國家。此外，我們也發現有些貼文出現上述語言特徵，但主要目的在促銷商品、服務或節目…等，即第(5)類以廣告為主要目的的貼文；另有如（辛）的第(6)類，雖然呈現(1)–(4)類語言

<sup>8</sup> 在蘇貞昌的 YouTube 頻道中，2020 年 1 月 29 日出現標題為「有政府，請安心」的影片，其中蘇貞昌說「有政府，會做事」以及「有政府，可以放心」，之後轉變為「有政府，好安心」的口號。

特徵，然而這類貼文實則並非傳達正向情感，需要跨欄知識或對時事的了解才能正確判斷。我們對這 6 類分別說明、舉例如下，並以底線標示屬於該類型的語言特徵。

**(1) 貼文使用正向評價語言：**這一類貼文使用正向評價語言（施孟賢等，2021）表達正向情感，參考施孟賢等(2021)，我們認為正向語言可以是單詞、片語或子句。由於正向語言有不同的形態，我們又分為下列(1A)-(1F)等次類：

**(1A) 使用呼籲團結、鼓勵、肯定積極行動的語言：**如【例 25】使用「我們一起確實做好防疫措施，好好照顧彼此」來呼籲人民團結對抗疫情：

【例 25】今日本土 287 例，境外移入 3 例  我們一起確實做好防疫措施，好好照顧彼此🙏❤️

**(1B) 引述、談論名人／專家：**引述名人／專家言論，或報導名人／專家的作為導致正向結果，或以主題標籤「#」標註名人／專家，如【例 26】說明政務委員唐鳳的作為，持續改善資料上傳的速度，有效解決以往資料上傳延遲所引發的問題。

【例 26】唐鳳又出手了！先前 500 筆資料要 2 小時、現在只要 10 分鐘，盼能降低地方醫療行政人員的負擔

**(1C) 感謝、體諒他人，對他人表達同理心：**用於感謝、體諒第一線醫護人員或為疫情提供協助的個人、公司或團體，如【例 27】之「再次感謝所有第一線醫護人員的努力！」，【例 28】則表達對勞工配合防疫作為的感謝：

【例 27】再強調一次，有真實的數據，才能進一步盤點各區情況，才能科學性地研議因應策略、及早部署，讓區域民眾安心。再次感謝所有第一線醫護人員的努力！

【例 28】有些在工地工作的勞工朋友，他們在工作上需要耗費極大的體力，口罩濕了又濕，換了又換，為了防疫，他們還是堅持戴上口罩，亦是最基層的無名英雄，謝謝你們的配合。

**(1D) 使用傳達正向情感的口號：**這一類的目的亦在呼籲人民團結，並確實做好防疫，但不同於(1A)，這一次類的語言多半為簡潔有力的短語，且一再出現，形成口號，如【例 29】的「台灣加油」：



【例 29】希望是往好的方向走，台灣加油👍👍

(1E) 表示自我肯定的語言：如同 Morin & Flynn (2014)和 Al-Tahmazi (2015)主張，臉書使用者會透過貼文來建構自我或內部團體，並使用正面自我(positive self-presentation, van Dijk, 2006)強化團結，如【例 30】，貼文者以確診數下降的趨勢肯定指揮中心使用「校正回歸」的作為：

【例 30】校正回歸後，我們的確症數趨勢真的逐漸往下了[…]

值得注意的是，(1E) 有時會貶抑他者，即藉由負面他者 (negative other-presentation, van Dijk, 2006)的對比，提升自我或內部團體的正向情感，如【例 31】直接批評韓國瑜的支持者不斷指責指揮中心蓋牌的言論：

【例 31】[…] 中央防疫一起來，成立國家篩檢隊，集合起民間業者力量。#當個堂堂正正的韓粉 #就繼續跳針蓋牌吧。。

(1F) 對於未來的期待與祝福：這個次類表達貼文者的期待與祝福，如【例 32】表達對於返回正常生活的期望。

【例 32】【#TOPick 新聞】希望台灣疫情早日緩和！

(2) 貼文使用標點符號、表情符號或語尾助詞傳達正向情感：這一類貼文使用（有時重複出現的）標點符號（驚嘆號、波浪符號）或表情符號、語尾助詞（「吧」「呀」「啦」…等）來表達驚喜、興奮，或善意提醒等正向情感。【例 33】以「各位都要乖乖待在家裡呀」提醒人民團結對抗疫情；【例 34】則是以「我們一起遵守防疫措施，度過疫情挑戰」加上祈禱和愛心表情符號，以及在「全面戴口罩 #務必配合實聯制，保護好彼此」之後，加入舉臂加油表情符號，呼籲人民團結對抗疫情。

【例 33】各位都要乖乖待在家裡呀（#Y 編）

【例 34】今日本土 334 例，境外移入 5 例 📄 我們一起遵守防疫措施，度過疫情挑戰 🙏❤️⁹[⋯] #全面戴口罩 #務必配合實聯制，保護好彼此 🙌🙌¹⁰

(3) **貼文使用類比：**正向情感貼文因引述指揮官的發言，使用了「塞車」、「蓋牌」和「作帳」等隱喻，這些隱喻大多帶有負面含意。【例 35】中的「塞車」，說明由於採檢量暴增，許多檢體來不及判讀，一周內便累積了一萬多件舊案延誤時間的情況，讓讀者透過車輛擁擠造成的交通阻塞，來理解採檢量過多而造成回報延誤的情況；「蓋牌」為博弈術語，表示隱藏牌面以迷惑對手；【例 36】的「作帳」意為改變數字以企圖美化帳目。為弱化這些隱喻的負面含意，正向貼文往往在其前後文中，出現修辭性問句（如【例 35】的「這樣的數字算是「#蓋牌」嗎？」）、反義詞（如【例 36】的「掀牌」）或否定詞（如【例 36】的「不是」）：

【例 35】#校正回歸 的問題：(20210522) 會有這樣的問題，因為 #北市 跟 #新北 #篩檢塞車 這兩地被校正的最多、其他地區數字是比較零星。塞車的問題會陸續解決，再行政上面簡化後，會快點出來，這樣子對疫情狀況會有幫助。[⋯]為什麼要採校正數字、主要是疫情評估的判斷、陽性率、趨勢的判斷，會影響疾管署的措施。這樣的數字算是「#蓋牌」嗎？把數字都秀給你了，這樣子叫做蓋牌.... 確診數字不是匯市、股市冷冰冰的數字，這些是確診數字，一個數字後面有一個人呢。

【例 36】《單日「721 例」解謎 新增 321 例、400 例「校正回歸」！陳時中：這是掀牌、不是作帳》

(4) **貼文訴諸其他國家：**如同負向情感語言中的（丁）類，正向情感也訴諸其他國家，特別是先進國家的作為。不同於（丁）的是，這類貼文的目的是為指揮中心使用「校正回歸」一詞辯護，稱「校正回歸」並非新創名詞，歐美日等先進國家和其他國家早已使用，指揮中心使用這個詞是為了讓確診數字合理，如【例 37】：

【例 37】不是蓋牌！ICU 醫生解釋，「校正回歸」是將處理好的舊通報案件，加回前面的日期，在歐美都有參照。

(5) **貼文以廣告為主要目的：**這類貼文可能一部分包含事實報導，也可能利用上述表達

⁹ 這三個表情符號分別是祈禱、方塊和愛心。

¹⁰ 這兩個符號皆為舉臂加油。

正向情感的語言，但通常主要目的為宣傳某團購、餐廳、節目、品牌、政治人物…等，如【例 38】出現第(1A)和(2)類，同時也出現百貨公司名稱進行宣傳：

【例 38】[...]#遠百信義 A13 #防疫日常 🍷 #全民防疫共同守護 ❤️ #安心賣場 🍷 #防疫人人有責

這一類貼文也有僅宣傳產品，只在文末使用主題標籤#標註「校正回歸」，如【例 39】：

【例 39】[...]🍷不以大集團式的華麗廣告行銷手法，而是回饋實實在在的風土。  
🍷裕森老師說酒莊的 Riesling 也深受似乎在這片有紅色石灰質土的園中活得很快活 沒有太多生長的壓力和困難 喝來溫暖而柔和 2 款經典新品以絕佳的品質立足澳洲酒海中 ❤️ 更是所有酒迷探索年度最佳酒莊的最佳捷徑！ ❤️ #酪洋國際酒業 #JAMESSUCKLING #JimBarryWines #TheArmagh #Shiraz #CabernetSauvignon #價錢超優惠 絕不 #校正回歸

(6) 需要跨欄知識或時事認知才能正確判斷的貼文：這類貼文看似表達正向情感，實際上需要跨欄知識（如，貼文者身分）或對時事有所認識才能適切判斷貼文的情感。【例 40】單就貼文內容而言，看起來是使用正向情感詞彙「創造」，然而貼文者為政黨傾向偏國民黨的「李姓中壢選民」，並引用聯合報的民意論壇，故而無法單就貼文內容判斷情感：

【例 40】#校正回歸 #創造新名詞 圖片來源：民意論壇：聯合報。世界日報。udn tv

在檢視貼文時，我們也發現，有些貼文僅使用單一類別或次類，如【例 41】僅使用（乙）刪節號表示貼文者的無奈；多數貼文則使用一個以上的類別或次類，【例 42】出現了(1B)和(3)：引述了專家（陳時中）的談話內容，而其談話內容則使用了類比，以「蓋牌」比喻隱瞞實際的確診人數，「掀牌」比喻揭露確診人數，「作帳」比喻美化確診人數的數字以圖粉飾太平：

【例 41】想篩檢，只能去排隊 ... #海角天編 【快點 TV】  
<https://gotv.ctitv.com.tw/2021/05/1778685.htm>

【例 42】陳時中表示：這是掀牌，怎會是蓋牌？ #校正回歸 #新冠肺炎 《單日「721 例」解謎 新增 321 例、400 例「校正回歸」！陳時中：這是掀牌、不是作帳》[…]

我們根據「人工負－模型正」和「人工正－模型中立」的語言特徵類別的筆數與比例製作表 1 和表 2：

**表 1. 負向語言特徵類別**

*[Table 1. Categories of linguistic features of negative polarity]*

負向語言特徵類別		筆數	比例 (%)
貼文使用負 向評價語言 筆數總和 101 比例 51.79%	(甲 1) 使用指責、批評、否定或不信任的語言	34	17.44
	(甲 2) 引述、談論名人／專家	11	5.64
	(甲 3) 貼文質疑政府、政治人物或某些政治人物的支持者	20	10.26
	(甲 4) 貼文使用貼標籤法	10	5.13
	(甲 5) 貼文直接使用表達負向情感的文字	21	10.77
	(甲 6) 貼文表達不理解或猜測	5	2.56
(乙) 貼文使用標點符號、表情符號或語尾助詞傳達負向情感		26	13.33
(丙) 貼文使用類比		12	6.15
(丁) 貼文訴諸其他國家		4	2.05
(戊) 貼文使用反語或反串文		36	18.46
(己) 貼文引述典故、文學作品、電視劇、電影等		5	2.56
(庚) 貼文變造口號		2	1.03
(辛) 需要跨欄知識或時事認知才能正確判斷的貼文		9	4.62
總和		195	100

表 2、正向語言特徵類別

[Table 2. Categories of linguistic features of positive polarity]

語言特徵類別		筆數	比例 (%)
貼文使用正向評價語 筆數總和 65 比例 54.62%	(1A) 使用呼籲團結、鼓勵、肯定積極行動的語言	27	22.69
	(1B) 引述、談論名人／專家	24	20.17
	(1C) 感謝、體諒他人，對他人表達同理心	3	2.52
	(1D) 使用傳達正向情感的口號	3	2.52
	(1E) 表示自我肯定的語言	3	2.52
	(1F) 對於未來的承諾、期待與祝福	5	4.20
(2) 貼文使用標點符號、表情符號或語尾助詞傳達正向情感		25	21.01
(3) 貼文使用類比		5	4.20
(4) 貼文訴諸其他國家		2	1.68
(5) 貼文以廣告為主要目的		20	16.81
(6) 需要跨欄知識或時事認知才能正確判斷的貼文		2	1.68
		119	100

表 1 顯示，第（甲）類「貼文使用負向評價語言」比例最高，共計 101 筆，佔所有筆數 51.79%，而（戊）類「貼文使用反語或反串文」比例次高，共 36 筆，佔 18.46%，第三高比例的語言特徵則為（乙）類「貼文使用標點符號、表情符號或語尾助詞傳達負向情感」，共 26 筆，佔 13.33%，但這三類貼文卻被模型視為正向情感，顯示模型的不足。表 2 顯示，第(1)類「貼文使用正向評價語言」比例最高，共計 65 筆，佔所有筆數 54.62%，次高為第(2)類「貼文使用標點符號、表情符號或語尾助詞傳達正向情感」共 25 筆，佔 21.01%，若排除貼文以廣告為主要目的，第三高則為第(3)類「貼文使用類比」，共 5 筆，佔 4.20%。

### 4.3 討論 (Discussion)

我們進一步討論負向情感和正向情感的語言特徵類別（含次類）的異同。就相異之處而言，除明顯表達負向或正向情感詞彙的使用（如（甲 1）和(1A)…等），負向情感使用了「貼標籤」、「反語或反串文」、「引述典故、文學作品、電視劇、電影等」以及「變造口號」等，而正向情感使用了「傳達正向情感的口號」。

就相同之處而言，兩種情感皆「引述、談論名人／專家」、「使用標點符號、表情符號或語尾助詞表達負向／正向情感」、「使用類比」、「訴諸其他國家」，以及「需要跨欄知識或時事認知才能正確判斷的貼文」，然而，貼文者的意圖使這些語言特徵產生相當大的差異，我們分別討論如下：

#### 4.3.1 引述、談論名人／專家 (Posts Citing or Discussing Celebrities/Experts)

雖然同樣「引述、談論名人／專家」，使用的詞彙和內容卻迥然不同：「人工負一模型正」的貼文者在引述名人／專家談論校正回歸時，選擇帶有負向情感的言說動詞，如「(開) 嗆」、「質疑」、「爆料」、「解密『校正回歸』內幕」，如上文【例 2】；若言說動詞為中性的「表示」或單純使用冒號，引述內容則傳達負向情感，如【例 43】的「自滿」、「膨風」和「疫情肆虐」，與【例 44】的「有何罪」和「不合理」：

【例 43】「書記長鄭麗文表示，國際媒體一致定調台灣是自滿加膨風，才造成疫情肆虐，證明民進黨若繼續大內宣，無法戰勝病毒，請蔡政府停止政治防疫，進入科學務實的防疫[...]」

【例 44】「[...] 楊艾俐：企業要買快篩有何罪？指揮中心不同意，太不合理了。」  
[...]

反觀「人工正一模型中立」的貼文在引述、談論名人時，言說動詞通常使用中性的「說」、「表示」、「指出」、「列出」、「解釋」、「說明」、「講解」、「公布」、「宣布」，甚至是帶有正向情感的「說情」，且引述內容通常在解釋緣由或澄清疑慮，或者帶有正向情感，如【例 45】；也有不使用言說動詞，僅使用引述內容傳達正向情感再加上主題標籤 #，如【例 46】

【例 45】市長柯文哲特別為大眾講解，採檢完到確診一般會有兩天的時間差，但最近檢體數量太多，「時間差」就會超過 2 天，因而出現校正回歸的情況。侯友宜指出，因為時間、資訊有落差，往往中央在確診個案篩檢過程中，有擁擠的狀況，時間有落差，光新北市最高達落差 6 天，為了防止造成防疫破口，中央防疫一起來，成立國家篩檢隊，集合起民間業者力量。

【例 46】校正回歸是要回歸到疫情的「真實面」，目前也積極解決問題中！ #校正回歸 #陳時中

#### 4.3.2 使用標點符號、表情符號或語尾助詞表達負向／正向情感 (Posts Showing Negative or Positive Polarity with Punctuation Marks, Emoticons, or Sentence Final Particles)

負向／正向情感的貼文皆使用標點符號、表情符號或語尾助詞，其中，表情符號包含英文字母結構，例如 QQ 或 XD，或特殊符號和標點符號結合之 Σ(°Д°;)，以及手機表情包

中的符號，如哭臉、舉臂加油等，以明確表達特定情感。語尾助詞因有多樣功能，例如，「啦」同時具有興奮和不耐等兩種情感，仍須視前後文而決定。

標點符號方面，兩種情感皆使用波浪紋和驚嘆號，但我們發現，刪節號(…)為負向情感獨有，我們主張貼文者使用刪節號，表達無奈、失望、難過、擔憂等負向情感。目前的模型並未考量標點符號、表情符號或語尾助詞等，未來在發展模型時建議納入。

#### 4.3.3 使用類比 (Posts Employing Analogy)

兩種情感使用的類比包含了隱喻及情節／劇本(scenario)，且大多帶有負面含意，例如「塞車」、「蓋牌」、「捅一刀」等。負向情感貼文除了上述隱喻，尚有「黑洞」、「作帳」、變造防疫指揮官名字的「時鐘一直都不準」，以及路人指揮消防員救火事宜之情節／劇本；正向情感貼文因引述陳時中的發言，無可避免地使用了「塞車」、「蓋牌」、「作帳」等隱喻，在使用這些帶有負面含意的隱喻時，正向情感貼文會使用修辭性問句、反義詞或否定詞以弱化其負面含意，而負向情感貼文則直接使用這些帶有負向情感的類比。

#### 4.3.4 訴諸其他國家 (Posts ASpealing to other Countries)

負向／正向情感皆訴諸其他國家，然而用法卻不同，負向情感的貼文者認為歐美日等國的防疫政策較能有效防疫，從而表達對台灣防疫政策的不滿，如前文的【例 16】，正向情感的貼文者則藉由說明其他國家也使用「校正回歸」來正當化台灣指揮中心使用「校正回歸」的舉措，如前文的【例 37】。

#### 4.3.5 需要跨欄知識或時事認知才能正確判斷的貼文 (Posts Requiring the Knowledge of News and of the Facebook Account Owner's Political/Social Inclination for Correct Prediction)

兩種情感皆有這類貼文，凸顯有些貼文無法單就內容來正確判斷貼文的情感面向，需要進一步了解貼文者身分、政黨傾向，有時甚至需要具備時事知識，方能正確判斷。在未來發展模型時，宜將貼文者的身分納入考量。

最後，我們也發現「貼文以廣告為主要目的」僅出現在「人工正一模型中立」的貼文中，且佔總所有語言特徵的 16.81%，顯示臉書的廣告文通常訴求正向情感。

### 5. 結論 (Conclusion)

本研究搜集台灣社群媒體上的使用者（包括社群媒體小編、一般社群媒體使用者）生產資料為分析對象，嘗試以既有之中文自然語言處理模型進行分析，並與人工標記結果進行比對，以了解模型預測與基準真值之間的落差，及其形成的可能原因。由此，本研究可以提出調校繁體中文自然語言處理模型面對社群媒體資料時的可能路徑。

我們的研究結果已回答了第 1 章提出的研究問題。首先，在人工標註和模型預測之差異方面，本研究的結果顯示，我們的模型採用 BERT-Chinese 預訓練模型作為模型參數

的初始值，再運用我們的訓練資料集進行模型訓練及微調機制，進而預測測試資料集貼文內容的情感傾向。對於人工標註為中立的貼文，模型有較好的預測能力，正確率高達 0.81；然而，人工標註為正向及負向的貼文，模型預測之正確率分別為 0.64 和 0.63，而其中在「人工負－模型正」的錯誤率最高，達 0.23，其次為「人工正－模型中立」的貼文，錯誤率達 0.22，可以看出當以社群媒體的貼文為語料時，人工標記為負向的貼文往往被模型判斷為正向貼文，而人工標記為正向的貼文，模型往往預測為中立。

為提供未來模型訓練的方向，我們逐筆分析「人工負－模型正」及「人工正－模型中立」的貼文，並分別歸納出 7 類負向情感和 4 類正向情感的語言特徵<sup>11</sup>，期能讓模型在處理繁體中文之社群媒體貼文時，提升預測的正確率。

本研究發現，若僅聚焦 text/message 欄位，模型很容易將反語或反串文標註為正向或中立（例如：「校正回歸，請造句」），我們因而主張，社群媒體的發文，除了對文本進行分析，亦需考慮發文者本身的政治與社會位置。如同語言學家分析詞彙時常考慮一詞彙前後的共現詞，對社群媒體的分析，也需考慮文本之外的其他線索。這種跨欄位的分析模式，是否有可能寫入自然語言學習的模型中，以便事先考慮其對分析結果的影響？例如，在訓練模型前，先將發文社團或個人的政治傾向加以分類，並在預訓練中將此設為變項，換言之，將貼文者的政黨屬性納入預訓練的一部分，是否能使模型更正確判斷該貼文的情感向度？

從上述的研究成果可以發現，由於社群媒體素材是由眾人一起生產出來的，因此相較於訓練素材取自字典、教科書、專業新聞報導的模型，本研究分析的語料具有文本較短、用字較為口語化（如：以「北車」指稱「台北車站」）、語句不一定完整（如：以空格隱藏明示意義）、重複音節、使用標點符號（如：本研究發現之刪節號為表達負向情感所獨有）或表情符號表達特殊情感、語詞意義層次變化多（如：隱喻、反語）等特質，這些皆可提供模型未來訓練的方向。換言之，社群媒體的語言有別於模型預訓練使用之書面語或正式語言，我們建議未來宜持續發展適合社群媒體語意的預測模型，並將上述語言特徵納入考量，在模型對社群媒體情感向度的判斷上，將助益良多。

本研究的限制，在於我們使用訓練資料集訓練後，模型在負向和正向的情感判斷上仍難符合人工標記的結果，我們認為這是訓練資料集不足所致，在有限的時間下，我們僅能完成這些語料的標記，特別是（戊）類「貼文使用反語或反串文」，模型預測的結果差強人意，這來自於反語和反串文本本身的複雜性：Wilson (2017)認為反語是否成立，取決於發言者是否抱持諷刺態度、情境／事件／表現是否違反人們期待的規範，以及發言者是否使用諷刺語氣，而反串文則是貼文者偽裝自己是某種身分立場的人來發言，以達到特定目的。這些條件，若抽離社會、政治脈絡，便難以正確判斷。若期待模型在預測這類貼文有良好表現，則需以大量資料訓練模型，以 NLP 被廣泛使用的 BERT 模型為例，它是由 Wikipedia (25 億字)及 Books Corpus (8 億字)等未經標籤的大量文本資料作為

---

<sup>11</sup> 相關臉書貼文及其語言特徵的類別標記公開於下列網址：  
[https://www.cs.nccu.edu.tw/~sichiu/Category\\_0509.xlsx](https://www.cs.nccu.edu.tw/~sichiu/Category_0509.xlsx)



預訓練模型的資料集；而 BERT-Chinese 的模型是採用約 50 萬字詞進行預訓練模型；換言之，需要以如此大量的資料進行訓練方得以建構預訓練模型。若要處理反語或反串文，亦需要大量的反語和反串文進行訓練，方得以建構能分辨這類語言使用之模型。

## 致謝 (Acknowledgements)

本論文承蒙科技部計畫 MOST 109-2410-H-004-MY2 補助，諸位匿名審查委員的指導，以及政大傳播學院碩士學位學程楊雅安、王孝成、陳姿樺三位同學協助編碼，特此致謝。

## 參考文獻 (References)

- Al-Tahmazi, T.H. (2015). The pursuit of power in Iraqi political discourse: Unpacking the construction of sociopolitical communities on Facebook. *Journal of Multicultural Discourse*, 10(2), 163-179. <https://doi.org/10.1080/17447143.2015.1042383>
- Chibuwe, A., & Ureke, O. (2016). 'Political gladiators' on Facebook in Zimbabwe: A discursive analysis of intra-Zimbabwe African National Union - PF cyber wars; Baba Jukwa versus Amai Jukwa. *Media, Culture & Society*, 38(8), 1247-1260. <https://doi.org/10.1177/0163443716671492>
- Cui, Y., Che, W., Liu, T., Qing, B., & Yang Z. (2019). Pre-training with whole word masking for Chinese BERT. In ArXiv preprint arXiv:1906.08101. <https://arxiv.org/abs/1906.08101>.
- Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2018). BERT: Pre-training of deep bidirectional transformers for language understanding. In ArXiv preprint arXiv:1810.04805. <https://arxiv.org/abs/1810.04805>
- Facebook (2021). *Facebook reports fourth quarter and full year 2020 results*. Retrieved from <https://investor.fb.com/investor-news/press-release-details/2021/Facebook-Reports-Fourth-Quarter-and-Full-Year-2020-Results/default.aspx>
- Jain, P. K., Quamer, W., Saravanan, V., & Pamula, R. (2022). Employing BERT-DCNN with sentic knowledge base for social media sentiment analysis. *Journal of Ambient Intelligence and Humanized Computing*, 2022, 1-13. <https://doi.org/10.1007/s12652-022-03698-z>
- Keen, A. (2007). *The Cult of the Amateur. How today's Internet Is Killing Our Culture*. Crown Business.
- Lu, Y., Pan, J., & Xu, Y. (2021). Public sentiment on Chinese social media during the emergence of COVID-19. *Journal of Quantitative Description: Digital Media*, 1, 1-47. <https://doi.org/10.51685/jqd.2021.013>
- Maíz-Arévalo, C. (2015). Jocular mockery in computer-mediated communication: A contrastive study of a Spanish and English Facebook community. *Journal of Politeness Research*, 11(2), 289-327. <https://doi.org/10.1515/pr-2015-0012>
- Morin, D.T., & Flynn, M.A. (2014). We are the Tea Party!: The use of Facebook as an online political forum for the construction and maintenance of in-group identification during the

- 'GOTV' weekend. *Communication Quarterly*, 62(1), 115-133.  
<https://doi.org/10.1080/01463373.2013.861500>
- Musolff, A. (2004). *Metaphor and Political Discourse: Analogical Reasoning in Debates about Europe*. Palgrave Macmillan.
- Musolff, A. (2006). Metaphor scenarios in public discourse. *Metaphor and Symbol*, 21(1), 23-38. [https://doi.org/10.1207/s15327868ms2101\\_2](https://doi.org/10.1207/s15327868ms2101_2)
- Nissim, M., & Patti, V. (2017). Semantic aspects in sentiment analysis. In Pozzi, F.A., Fersini, E. Messina, E., & Liu, B. (eds.), *Sentiment Analysis in Social Networks* (pp. 31-48). Elsevier Inc.
- Roberts, K., Roach, M. A., Johnson, J., Guthrie, J., & Harabagiu, S. (2012). Empatweet: Annotating and detecting emotions on twitter. In *Proceedings of the Eighth International Conference on Language Resources and Evaluation (LREC'12)*, 3806-3813.
- Silver, D., Huang, A., Maddison, C.J., Guez, A., Sifre, L., van den Driessche, G., Schrittwieser, J., Antonoglou, I., Panneershelvam, V., Lanctot, M., Dieleman, S., Grewe, D., Nham, J., Kalchbrenner, N., Sutskever, I., Lillicrap, T., Leach, M., Kavukcuoglu, K., Graepel, T., ... Hassabis, D. (2016). Mastering the game of Go with deep neural networks and tree search. *Nature*, 529, 484-489. <https://doi.org/10.1038/nature16961>
- Singh, M., Jakhar, A. K., & Pandey, S. (2021). Sentiment analysis on the impact of coronavirus in social life using the BERT model. *Social Network Analysis and Mining*, 11(1), 1-11. <https://doi.org/10.1007/s13278-021-00737-z>
- Socher, R., Perelygin, A., Wu, J.Y., Chuang, J., Manning, C.D., Ng, A.Y., & Potts, C. (2013). Recursive deep models for semantic compositionality over a sentiment treebank. In *Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing*, 1631-1642.
- Sperber, D., & Wilson, D. (1981). Irony and the use-mention distinction. In P. Cole (ed.) *Radical Pragmatics*. Academic Press. Reprinted in S. Davis (ed.) (1991) *Pragmatics: A Reader* (pp. 550-563). Oxford University Press.
- Tabe, C.A. (2016). Language and humor in Cameroon social media. In Taiwo, R., Odebunmi, A., & Adetunji, A. (eds.) *Analyzing Language and Humor in Online Communication* (pp. 131-163). IGI Global.
- Tang, D., Wei, F., Yang, N., Zhou, M., Liu, T., & Qin, B. (2014). Learning sentiment-specific word embedding for Twitter sentiment classification. In *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics*, 1, 1555-1565. <https://doi.org/10.3115/v1/P14-1146>
- Tannen, D. (2013). The medium is the metamessage: Conversational style in new media interaction. In Tannen, D., & Trester, A.M. (eds.) *Discourse 2.0. Language and New Media* (pp. 99-118). Georgetown University Press.
- van Dijk, T.A. (2006). Ideology and discourse analysis. *Journal of Political Ideologies*, 11(2), 115-140. <https://doi.org/10.1080/13569310600687908>

- Wang, T., Lu, K., Chow, K. P., & Zhu, Q. (2020). COVID-19 sensing: Negative sentiment analysis on social media in China via BERT model. *IEEE Access*, 8, 138162-138169. <https://doi.org/10.1109/ACCESS.2020.3012595>
- Wilson, D. (2017). Irony, hyperbole, jokes and banter. In J. Blochowiak et al. (eds.) *Formal Models in the Study of Language* (pp. 201-219). Springer.
- Ye, Z. (2019). The emergence of expressible agency and irony in today's China: A semantic explanation of the new bèi-construction. *Australian Journal of Linguistics*, 39(1), 57-78. <https://doi.org/10.1080/07268602.2019.1542933>
- PTT 鄉 民 百 科 。 反 串 文 。  
<https://pttpedia.fandom.com/zh/wiki/%E5%8F%8D%E4%B8%B2%E6%96%87> 。 [PTT Netipedia. Fake post. <https://pttpedia.fandom.com/zh/wiki/%E5%8F%8D%E4%B8%B2%E6%96%87>.]
- 呂秋遠 (2018) 。 呂秋遠專欄：台灣人患了巨嬰症？蘋果新聞網 。  
<https://tw.appledaily.com/forum/20180914/JS3WXKYWXFOI5XOTWLLN343LGI/> 。  
 [Lu, C.-y. (2018). Chiou-yuan Lu's Column: Taiwanese suffer from Macrosomia (giant baby syndrome)? *Apple Online*. <https://tw.appledaily.com/forum/20180914/JS3WXKYWXFOI5XOTWLLN343LGI/>.]
- 施孟賢、段人鳳、鍾曉芳 (2021) 。 中文新聞文本之宣傳手法標記與分析。中文計算語言學期刊，26(1)，79-104。 [Shih, M.-H., Duann, R.-f., & Chung, S.-F. (2021). The analysis and annotation of propaganda techniques in Chinese news texts. *International Journal of Computational Linguistics & Chinese Language Processing*, 26(1), 79-104.]
- 陳韋帆、古倫維 (2018) 。 中文情感語意分析套件 CSentiPackage 簡介。圖書館學與資訊科學，44(1)，24-41。 [https://doi.org/10.6245/JLIS.201804\\_44\(1\).0002](https://doi.org/10.6245/JLIS.201804_44(1).0002) [Chen, W.-F., & Ku, L.-W. (2018). Introduction to CSentiPackage: Tools for Chinese sentiment analysis. *Journal of Library and Information Science*, 44(1), 24-41. [https://doi.org/10.6245/JLIS.201804\\_44\(1\).0002](https://doi.org/10.6245/JLIS.201804_44(1).0002) ]
- 鍾文翔 (2018) 。 新聞導言之智能生成。中央大學資訊管理學系碩士論文。 [Zhong, W.-X. (2018). Intelligent Generation of News Lead (Mater's Thesis). National Central University.]

