

期中專案報告

gogoro

指導教授:黃三益

第15組_組員

N104020016 楊世華

N104020008 游淑媛

N104020029 陳勇任

B096060020 黃湘安

影片連結



大綱

01 動機與目的

02 資料清洗

03 分析流程視覺化結果

04 結論





01

選擇主題動機與目的

動機與目的

氣候變遷

電動運具成為趨勢

台灣運輸部門的二氧化碳
排放量占總量約14%

台灣電車龍頭

Gogoro

智慧電動機車
電池交換系統
GoShare共享機車
業務範圍擴及全球

台灣主要論壇

PPT

用戶年齡分佈多為
有購車需求與能力的成年人

02

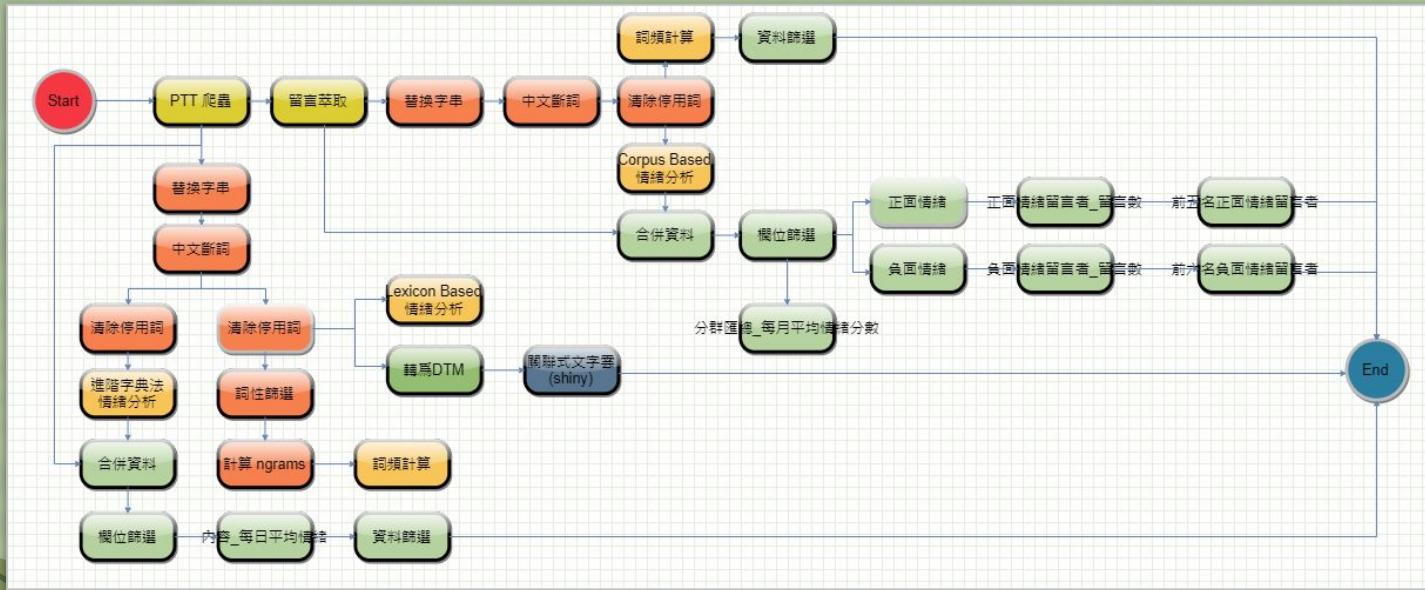
資料清洗



分析工具

使用文字探勘工作流程設計平台(Tarflow)進行文字探勘

工作流程名稱為PROJECT_1, 由左至右為文章內容與留言



爬蟲

1. 抓取資料區間: 2022年4月10日~2023年4月10日，共373筆。
2. 目的: 獲取分析標的。
3. 資料抓取標的: PTT機車板和八卦板。
4. 搜尋關鍵字: gogoro、狗肉、狗狗肉、GOGORO、Gogoro。
5. 排除關鍵字: 殺肉、戰爭、林志穎 ...等，包含後續反覆檢查抓取結果後，回饋修正添加而成。



替換文字

1. 目的: 將特殊符號或特定字串替換, 以利後續分析。
2. 替換字串設定: 去除標點符號、特殊字元及超連結, 共 6種符號
變化需替換為空白, 並將產品名稱格式統一



中文斷詞

1. 目的：將句子進一步切割成適當詞語，以利後續分析。
2. 定義詞彙：將具相關性詞語進行權重調整，共 101 個詞彙。



清除停用字

1. 目的：過濾掉無實際含意或非分析主體等詞彙，以利後續分析。
2. 定義詞彙：使用預設停止詞且轉換小寫英文，並清除換行符號、html tag、數字、特殊標點符號。除了系統設定外，亦加入自定義停止詞，內容包括較無意義的語助詞、數字與轉貼新聞報導的詞彙（例：問卦、轉載），共停用120個詞彙，清除 34,619筆文章詞彙、97,925筆留言詞彙。

統計資訊

34619

清除總數



統計資訊

97925

清除總數

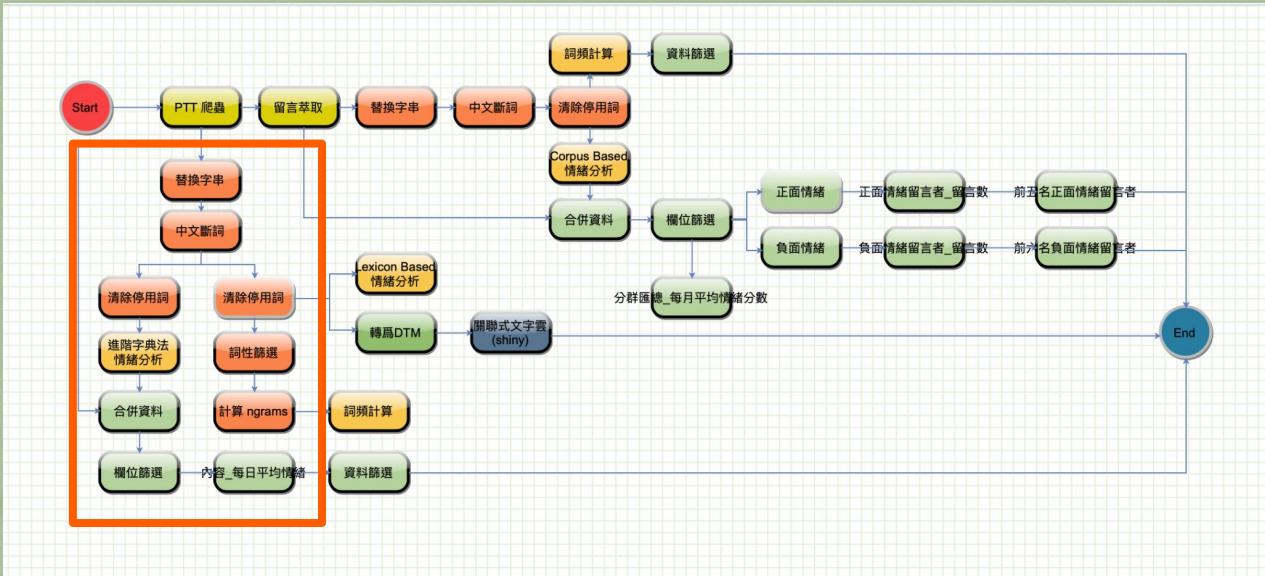


03

分析流程與視覺化結果



分析內容1



分析內容-1

1. 詞性篩選：

(2) 目的：根據停用詞結果，分別篩選標題中「專有名詞」與「名詞」，並進行詞頻計算，以檢視熱門關鍵字。

(3) 處理方式：保留詞性→專有名詞、名詞。

1. 計算 ngrams

(4) 目的：計算詞頻關聯度，找出連續出現相關性大的兩個詞。

(5) 處理方式：設定 $n=2$

1. 詞頻計算

(6) 目的：接續1-2進行計算，以檢視熱門關鍵字。。

(7) 處理方式：選擇前100組詞彙組合

分析內容-1

1. Lexicon Based情緒分析

- (2) 目的：接續清除停用詞處理，進行內容詞彙之情緒分析。
- (3) 定義正面詞彙。
- (4) 定義負面詞彙。

1. 進階字典法情緒分析

- (5) 目的：接續清除停用詞，針對內容詞彙進行進階字典法之情緒分析。
- (6) 設定部分因留言可能出現否定詞，於是在清除停用詞部分選擇不清除單字詞，反覆將輸出結果透過NTUSD字典法進行正負面情緒分析。

內容分析-1

1. 合併資料+欄位篩選

(2) 目的:將進階字典法情緒分析與原爬蟲資料合併，並篩選所需欄位進行分析。

(3) JOIN KEY:system_id。

(4) 篩選欄位: system_id、artDate、artPoster及sentiment_value。

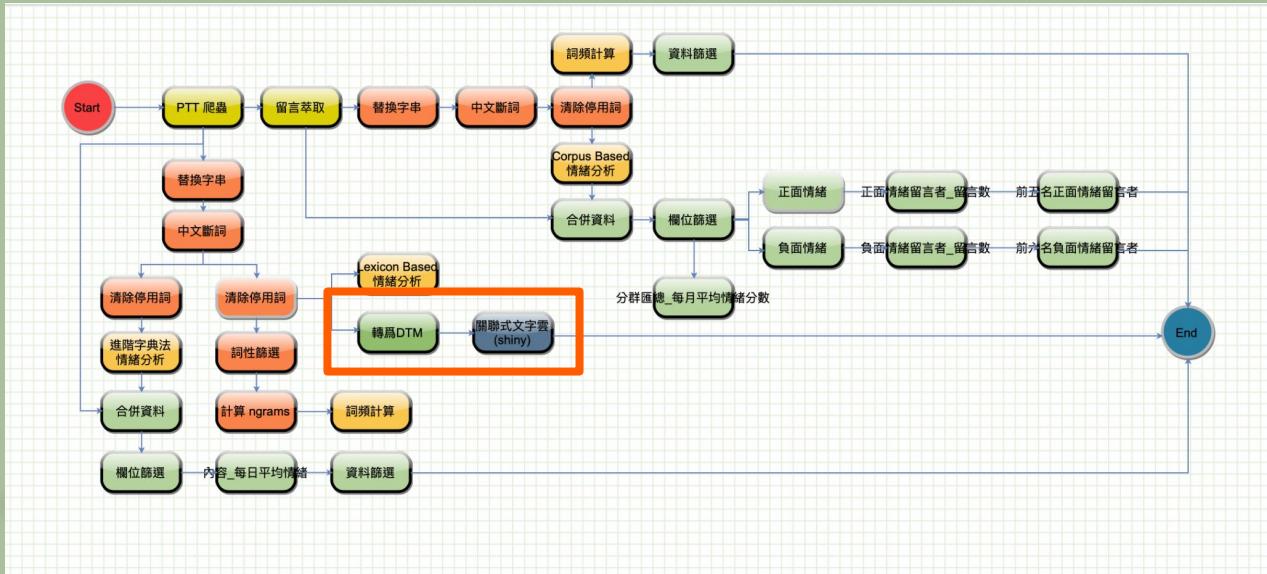
1. 分群匯總_每日平均情緒分數

(5) 目的:計算每日平均情緒分數。

(6) 處理方式:使用「artDate」作為日期欄位，並依日進行分析；計算「sentiment_value」欄位平均數，並保留「artDate」欄位。

(7) 執行結果:以日期分群，計算每日平均情緒分數。

內容 分析2



內容分析－文字雲

1. 轉為DTM

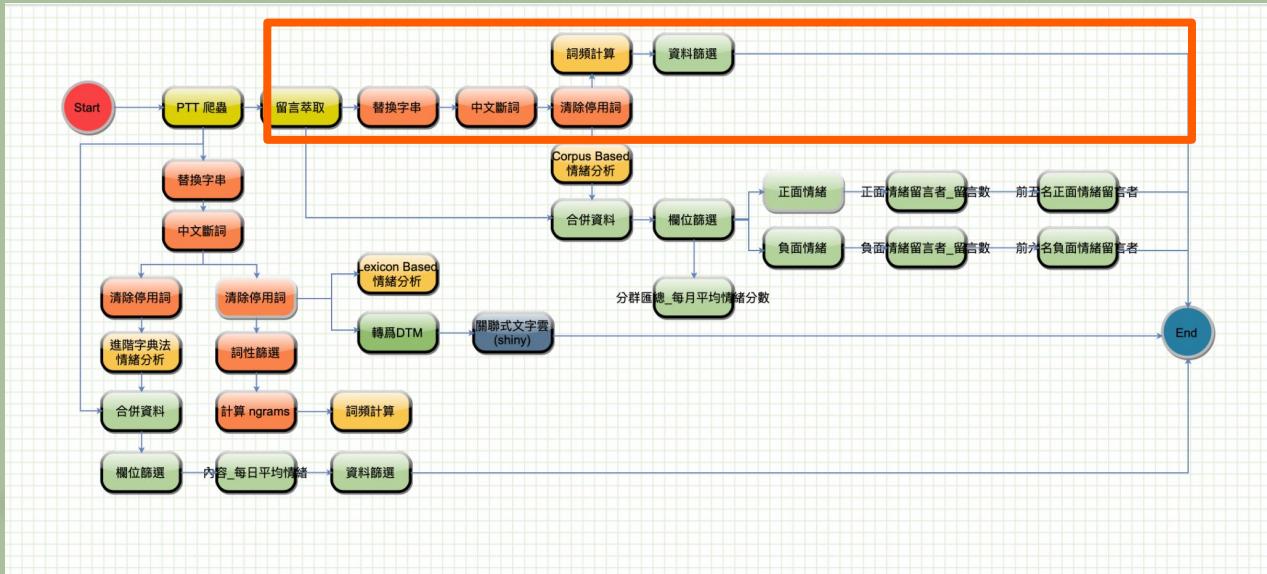
(2) 目的: 接續清除停用詞處理, 將斷詞後結果轉製成 200維度的矩陣。

1. 關聯式文字雲

(3) 目的: 接續3-1處理, 進行詞彙分群計算共線關聯性。

(4) 處理方式:以關聯性最高的前 50 個字詞, 分成 20 群顯示。

留言 分析1



留言分析－品牌名稱

1. 詞頻計算：

(2) 目的：接續清除停用詞處理，進行留言出現詞彙之頻率計算。

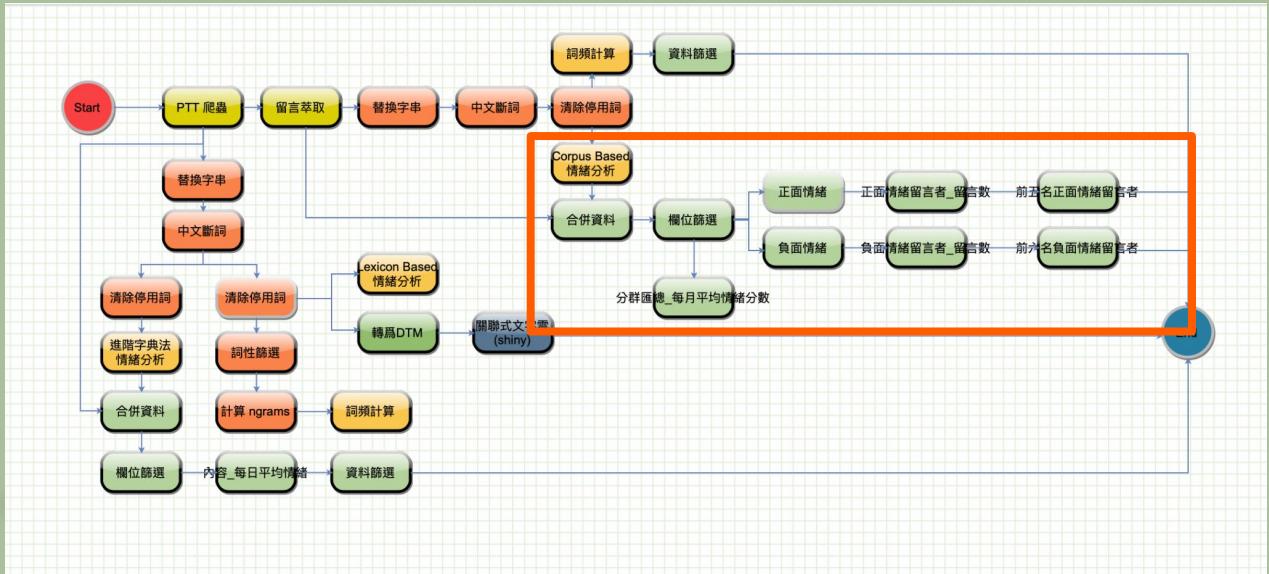
(3) 篩選詞彙數量：50。

1. 資料篩選：

(4) 目的：篩選出指定詞彙之詞頻計算結果。

(5) 條件式：'\$Term=="光陽"'或'\$Term=="gogoro"'...等7個電動車品牌名稱的條件。

留言 分析2



留言分析－情緒分析

1. Corpus Base分析：

- (2) 目的：接續清除停用詞處理，針對留言詞彙進行Corpus Base分析取得情緒分數。
- (3) 輸出結果：得到數據以0.5為基準，大於等於為正面情緒，反之為負面情緒。

1. 合併資料+欄位篩選：

- (4) 目的：將Corpus Base分析與原留言萃取資料合併，並篩選所需欄位進行分析。
- (5) JOIN KEY：comment_idx及system_id。

1. 分群匯總_每日平均情緒分數

- (6) 目的：計算每日平均情緒分數。
- (7) 處理方式：使用「cmtDate」作為日期欄位，並依日進行分析；計算「sentiment_value」欄位平均數，並保留「cmtDate」欄位。
- (8) 執行結果：以日期分群，計算每日平均情緒分數。

正面情緒分析

1. 資料篩選_正面情緒

(2) 目的:篩選出情緒分數大於及等於 0.5, 即呈現出偏正向之留言。

(3) 條件式:'\$sentiment_value>=0.5'。

1. 分群匯總_留言者之正面留言數

(4) 目的:以留言者與同篇文章為計算標的, 計算留言數。

(5) 處理方式:使用「cmtPoster、system_id」作為分群依據, 並計算「sentiment_value」欄位之數量。

1. 資料篩選_正面情緒留言者排名前五名

(6) 目的:篩選出發表正向留言大於 90 之留言者。

(7) 條件式:'\$sentiment_value@count>90

負面情緒分析

1. 資料篩選_負面情緒

- (2) 目的:篩選出情緒分數小於0.5, 但不為0, 即呈現出偏負向之留言。
- (3) 條件式:{'AND':[{'\$sentiment_value<0.5'}, {'\$sentiment_value>0'}]}

4-2 分群匯總_留言者之負向留言數

- (1) 目的:以留言者與同篇文章為計算標的, 計算留言數。
- (2) 處理方式:使用「cmtPoster、system_id」作為分群依據, 並計算「sentiment_value」欄位之數量。

4-3 資料篩選_負面情緒留言者排名前六名

- (1) 目的:篩選出發表負向留言大於90之留言者。
- (2) 條件式:'\$sentiment_value@count>90'。

分析結果一內容:詞頻分析

依據篩選名詞及專有名詞的結果進行詞頻計算，以「gogoro」最多，佔645次；「機車」佔230次，及「電池」佔181次，檢視分析結果，以專有名詞篩選詞性的結果算是相對精準，沒有出現明顯分類錯誤的詞語。



分析結果一留言: 詞頻分析

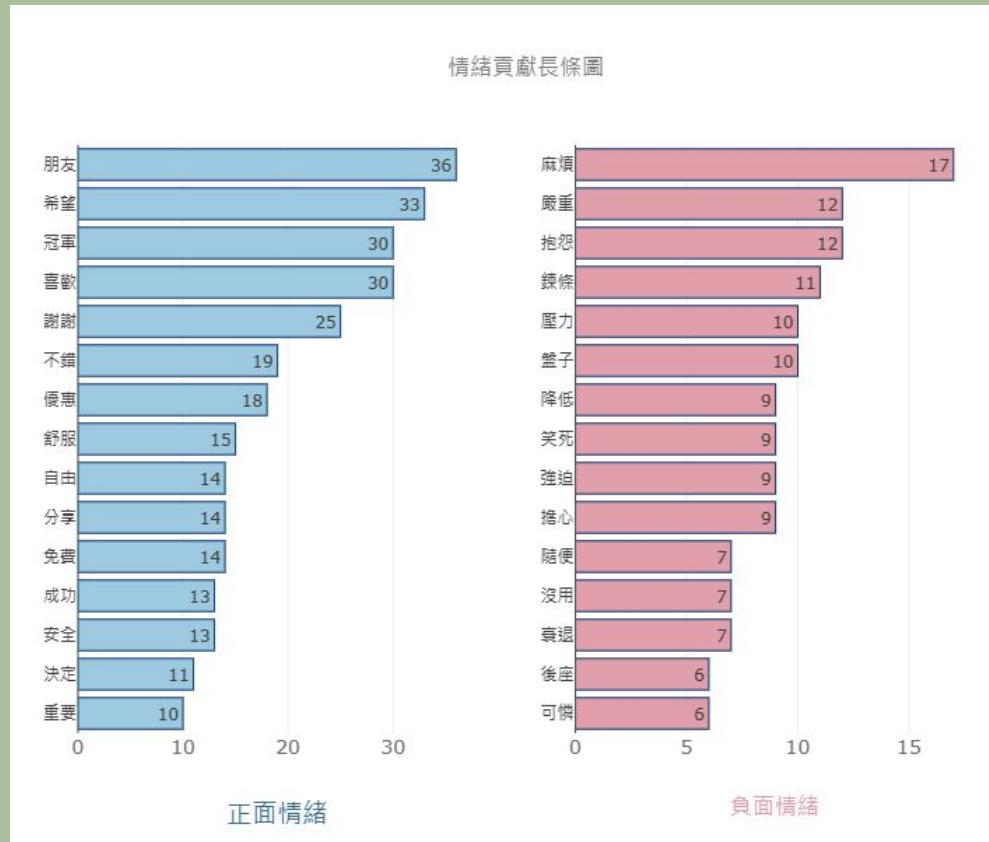
依據清除停用字後的結果進行留言內容詞頻計算，以「gogoro」最多，佔1320次；「電池」佔791次，及「油車」佔684次，比較特別是出現了「盤子」、「笑死」等詞彙。



分析結果一內容:情緒分析

Lexicon Based 情緒分析:

文章內容正面情緒較多，如朋友、希望、冠軍、喜歡等。負面情緒詞彙有麻煩、嚴重、抱怨、鍊條、盤子...等。



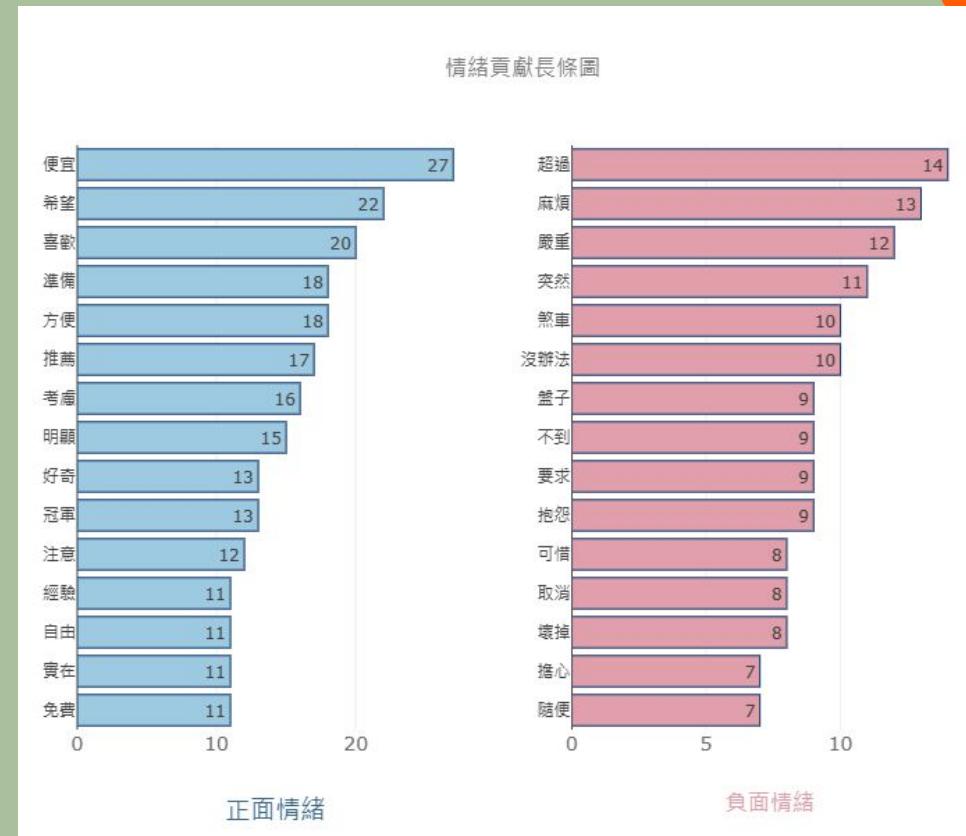
分析結果一內容:情緒分析

NTUSD進階字典法情緒分析:

正面情緒以「便宜」出現最高；

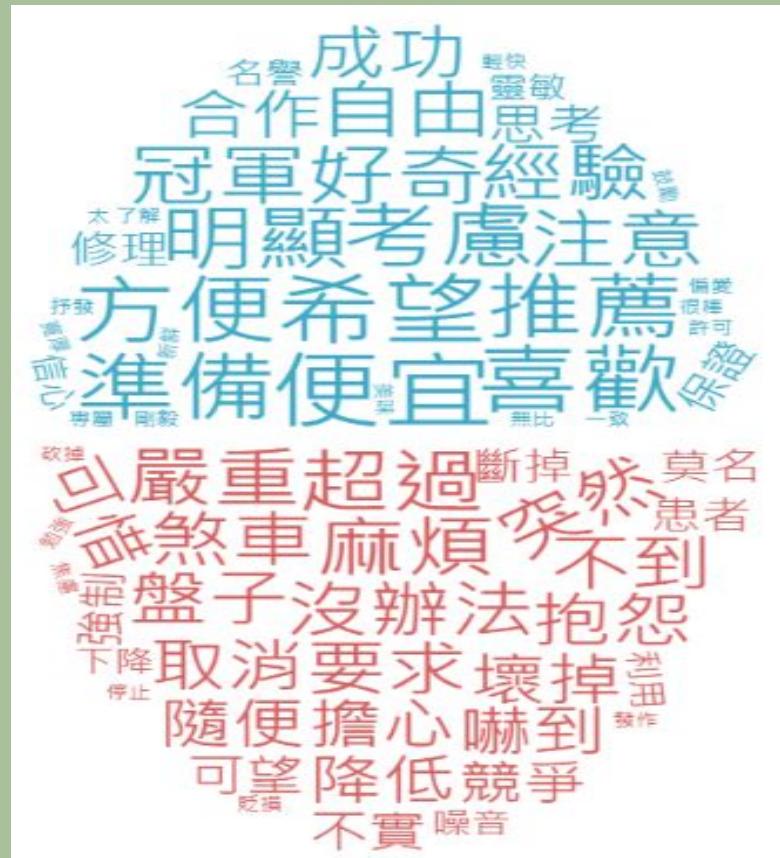
而負面情緒留言詞彙出現最多的是「超過」，其次為麻煩、嚴重、突然、煞車...等等，其中從正、負面情緒字中也有大量出現機車、感謝、光榮、謝謝、盤子...情緒詞。

情緒貢獻長條圖



分析結果一內容:情緒分析

透過字典法分析詞彙，雖然gogoro相關文章內容之正面情緒較多，但僅能代表發文內容之情緒多為正面，內容不一定專指gogoro，因此也不一定為對gogoro之正面評價。



分析結果一留言：情緒分析

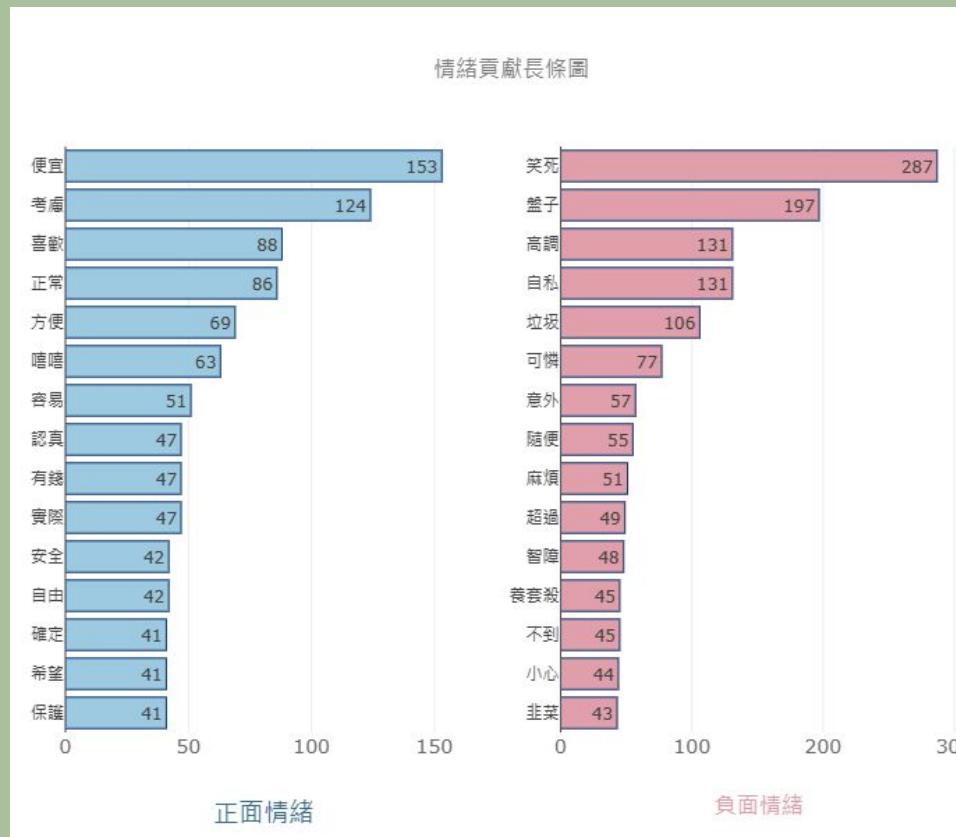
Lexicon Based 情緒分析：
留言內容則與文章內容呈現相反結果，留言內容負面情緒較多，如笑死、盤子、自私、可憐、垃圾、養套殺、割韭菜...等；正面則有不錯、喜歡等。



分析結果一留言:情緒分析

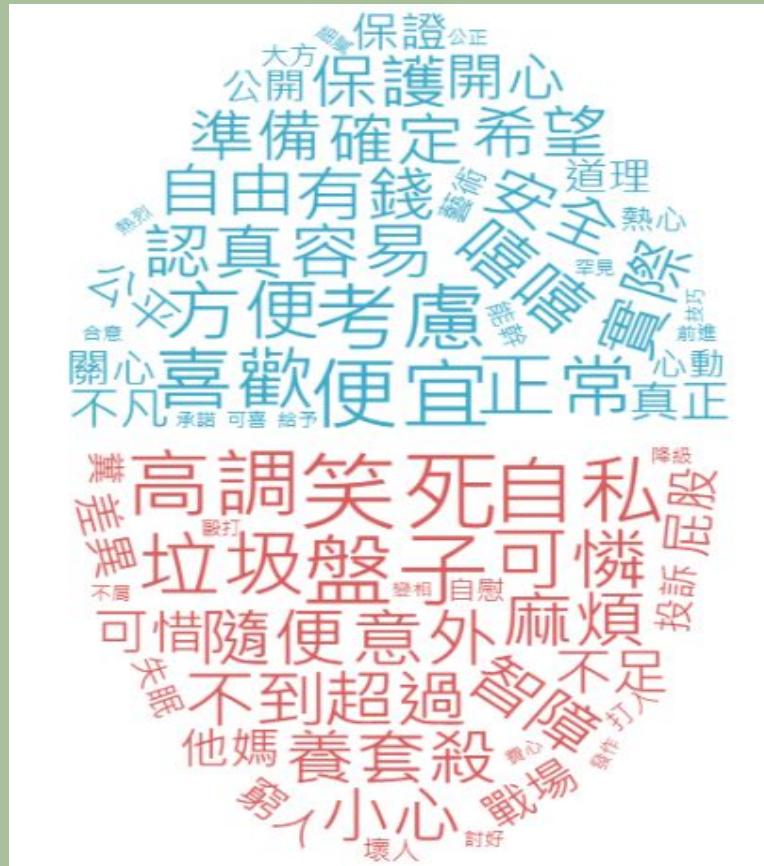
NTUSD進階字典法情緒分析：

留言內容負面情緒較多，如笑死、盤子、自私、可憐、垃圾、養套殺、韭菜...等；正面則有便宜等。



分析結果一留言：情緒分析

透過字典法分析留言詞彙可見gogoro相關文章留言之負面情緒較多；由內容得知，大多為網友爭論油車與電車孰優孰劣，而gogoro作為電車代表，因此網友之負面情緒多為針對gogoro。



分析結果一內容:情緒分析

合併進階字典法：

觀察指定資料區間發文內容情緒分數的變化，

因發文的來源大多為新聞轉發或是網友心得，由情緒分數可以看出沒有太明顯的變化，篩選出情緒分數最高及10/17最低分別有11/21及3/4。

文章主題依序分別為：

[新聞] IONEX 全面升級, 突破 8kW 成為市售最大
動力白牌電動機車

[心得] 第一次後悔買gogoro(鑰匙一定要放身上

[心得] EC-05與Gogoro選購試乘分享



分析結果－留言：情緒分析

合併進階字典法：

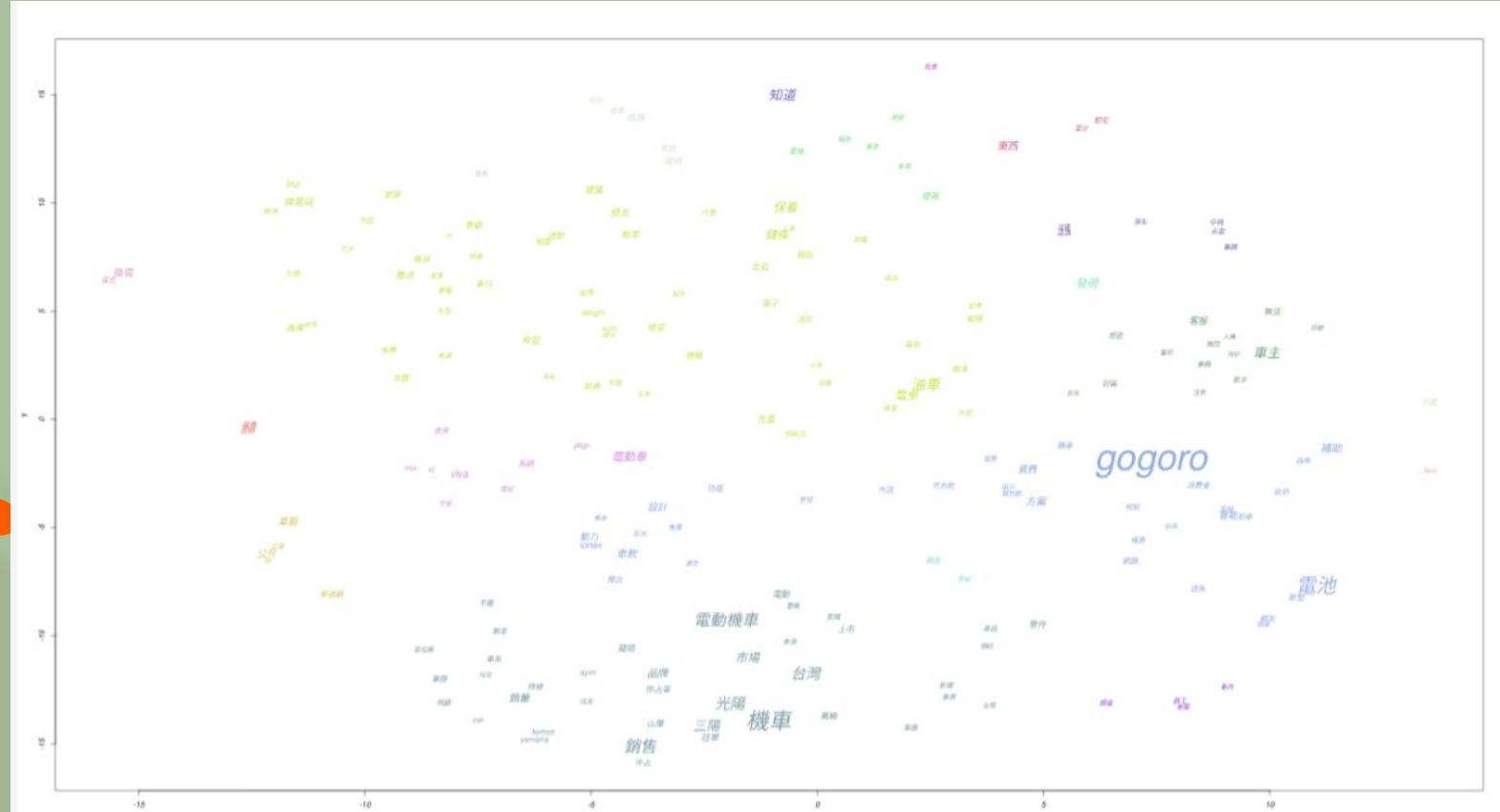
觀察指定資料區間留言區情緒分數的變化，可見7/5
當天平均情緒分數最低，8/11次低，10/8則為第三
低。

● 討論主題依序分別為：

- [新聞] 沒立中柱歪進隔壁車位夫妻被貼「自私」
- [問卦] 內湖GOGORO區經理有多厲害啊？
- [新聞] Gogoro「吃到飽宣佈退場」時間點曝光全場
驚：油車要流行了



分析結果一內容:共線相關圖



分析結果一留言: 詞頻分析

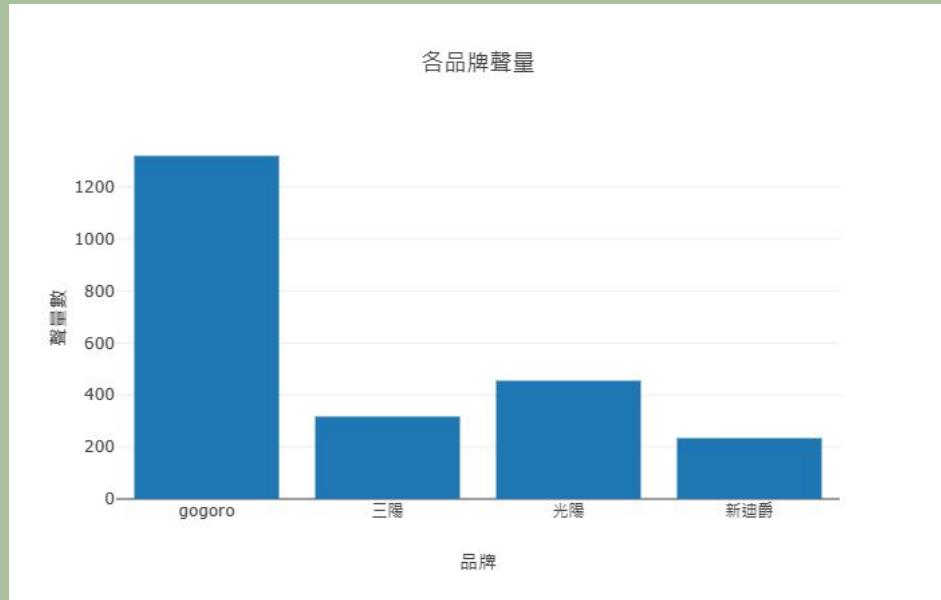
The word cloud highlights several key themes from the comments:

- Gogoro** is the most prominent word, indicating its central role in the discussions.
- 電動車 (Electric Vehicle)** and **電池 (Battery)** are also highly frequent.
- 光陽 (Gogoro)** and **三陽 (SanYang)** are mentioned, representing two major manufacturers.
- 的車 (The car)** and **車主 (Car owner)** appear, referring to the users of these vehicles.
- 充電 (Charging)** is a common topic, along with **充電站 (Charging station)**.
- 價錢 (Price)** and **便宜 (Inexpensive)** are frequently used when discussing the cost of these vehicles.
- 環保 (Environmental protection)** is highlighted, emphasizing the eco-friendly aspect.
- 換電 (Battery swapping)** is another notable feature mentioned.
- Other terms like **新迪爵 (XinDiJue)**, **東西 (Things)**, **法院 (Court)**, **月租 (Monthly rent)**, **垃圾 (Garbage)**, **側柱 (Side pillar)**, **樓上 (Upstairs)**, **外送 (Delivery)**, **根本 (根本)**, **皮帶 (Belts)**, **鏈條 (Chain)**, **吃到飽 (All you can eat)**, **高調 (High-profile)**, **自私 (Selfish)**, **里程 (Mileage)**, **盤子 (Plate)**, **補助 (Subsidy)**, **保 (Insurance)**, **環保 (Environmental protection)**, **建議 (Suggestion)**, **考慮 (Consideration)**, **不錯 (Not bad)**, **還好 (Pretty good)**, **根本 (Root cause)**, **廣告 (Advertisement)**, **車廂 (Vanity)**, **維修 (Maintenance)**, **認證 (Certification)**, **設計 (Design)**, **台灣 (Taiwan)**, **中柱 (Middle column)**, and **騎車 (Cycling)** are scattered throughout the cloud.

分析結果一留言：詞頻分析

上頁詞頻分析中：以「電池」最多、「電車」「油車」等次之，可見網友將 GOGORO 作為電車的代表，演變成 電車、油車之爭。

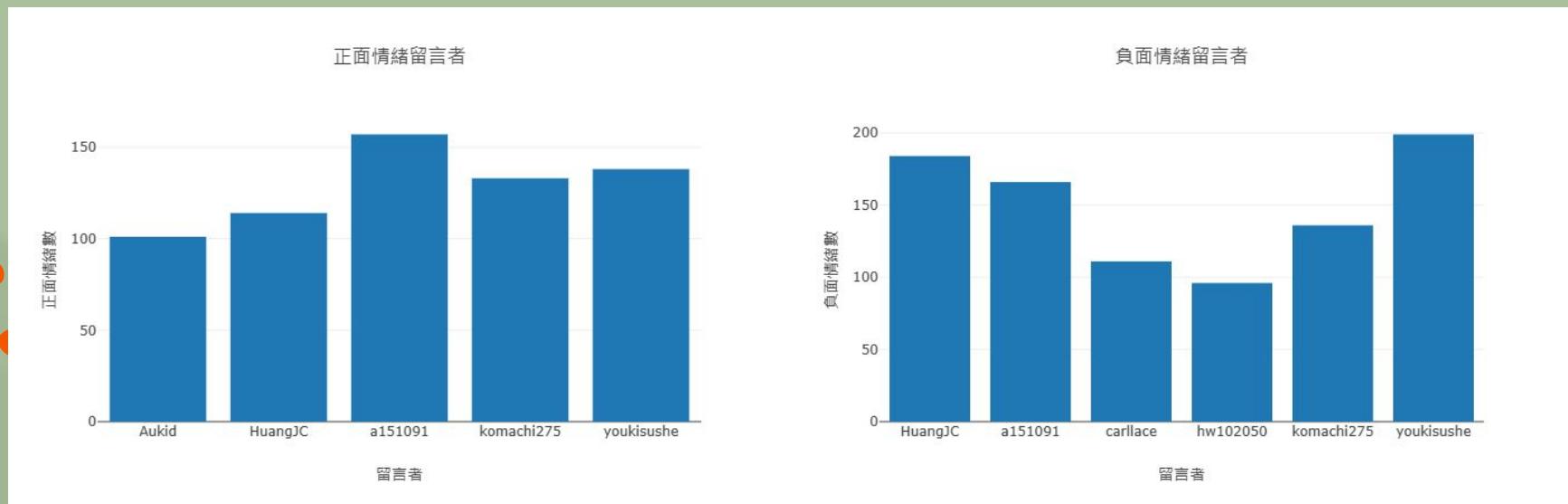
因同時出現其他機車品牌，因此針對其他品牌進行分析，可發現除了 gogoro 本身，最常出現的品牌為三陽、光陽及新迪爵，其中僅有新迪爵沒有參與任何電動車事業。



分析結果一留言：情緒分析

依據情緒分數將正面及負面分開討論，篩選出各自留言者是偏正面(左圖)或負面(右圖)。

其中較特殊的為youkisushe、komachi275、a151091、HuangJC四位留言者，留言中出現正向及負向詞彙比例都相當高，如之前所呈現，相關討論無法得知討論之對象為gogoro或是其他機車品牌，僅能推測這四位版友高度關注電動車相關議題，因留言相對多，導致其佔正、負向詞彙比例都很高。



04

結論



結論

一、正向情緒分析：

由文章內容、留言之情緒分析結果，不約而同有「便宜」、「方便」等字詞；但因相關主題同時提及多家廠牌，因此僅能推論「便宜」、「方便」等正向情緒字詞，是網友較為在意的面向。

二、負向情緒分析：

負面情緒則差異較大，文章內容主要的負面情緒字詞為「麻煩」、「煞車」等，屬於電車保養及操作的相關問題；而留言內容之負面情緒則為「笑死」、「盤子」、「養套殺」、「高調」等，主要著重在近一年 gogoro的爭議事件。

三、總結：

文章內容之來源多為新聞或網友心得，因此能較為客觀的評論，而留言內容為不特定網友，大多數網友仍是使用油車，gogoro雖為電車龍頭，但僅佔機車總體市佔約 10%，在網路上的聲量仍是油車>電車情況下，呈現在資料分析結果，正面情緒字詞不一定是指 gogoro，但負面情緒字詞則多描述 gogoro。

Thanks

CREDITS: This presentation template was created by [Slidesgo](#), including icons by [Flaticon](#), and infographics & images by [Freepik](#)