

社群媒體分析 期末報告 第 20 組

組員：N124320019 吳邦齊(組長)

N124320018 邱 翔

N124320020 陳軍弦

N124320012 鄭義璋

報告影片連結：<https://youtu.be/Ec-iwoKeOrl> (Whisper 生成字幕版本)

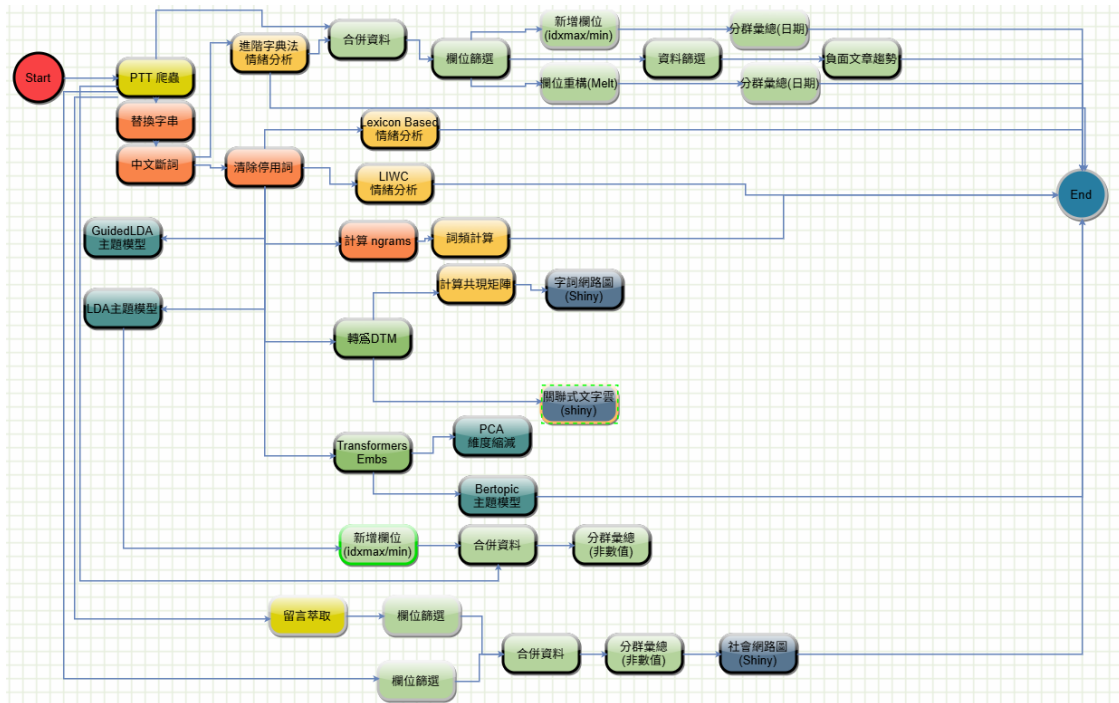
<https://youtu.be/cXKIBY5Vnso?si=JlIHlg2RIYIWLAVu>(備份連結)

1.動機和分析目的

隨著全球對電動車 (Electric Vehicle, EV) 產業的關注日益提升，相關品牌與技術在台灣社會中亦逐漸受到高度討論與關注。除了主流媒體的報導之外，社群論壇如 PTT 已成為民眾表達觀點、討論產品與政策、交流使用經驗的重要平台。特別是在股票、汽車、機車、生活與地方看板中，關於電動車的話題橫跨科技創新、消費選擇、政策補助、交通環境與生活品質等多重面向。

本研究希望透過蒐集 PTT 上與電動車相關的文本資料，搭配關鍵字設計與文本探勘技術，進行系統化的語料分析。分析內容包含詞頻與特徵提取、文字雲視覺化、LDA 主題模型、主題分布與情感分析等，並進一步結合共現矩陣與詞彙社會網絡圖，以掌握民眾對電動車的討論主軸與潛在情緒傾向。

2.Tarflow 流程圖



3.資料來源

PTT 爬蟲 (4)

選擇看板

marriage(婚姻)
medstudent(醫學生)
Militarylife(軍旅)
MobileComm(手機通訊)
MobilePay(行動支付)
mobilesales(手機買賣)

搜尋關鍵字

特斯拉
Tesla
Model 3
Model Y
Model X

排除關鍵字

以換行區隔, e.g.
壽山動物園
猴子
...

搜尋起始日期

2024/01/01

搜尋結束日期

2024/12/31

儲存更改

版面：car(汽車)、biker(機車)、movie(電影)

關鍵字：特斯拉、Tesla、Model 3、Model Y、Model X、Model S、比亞

迪、BYD、海豹電動車、小鵬汽車、蔚來汽車、裕隆鴻華、納智捷、n7、

Gogoro、宏佳騰、Ai-1、電動車、電動機車、充電樁、快充站、家用充電、里程


焦慮、電池壽命、固態電池、電池、自駕系統、自動駕駛、OTA 更新、OTA、碳權、電動車補助、電動車推薦、電動車維修、電動車好嗎

時間：2024/01/01~2024/12/31

我們使用 PTT 爬蟲，抓取與車子相關的汽車、機車和電影，搜尋相關關鍵字，針對 2024 年的資料作分析。

4.資料清洗

替換字串

≡ 替換字串  (27)

參數設定

Input - 4

任務結果

選擇處理欄位 *

artContent

選擇替換規則檔案 ❶

-----請選擇-----

替換字串設定 ❶

\n\n>> °

\n>> *

Sent from my JPTT on my \w+>>

Sent from BePTT on my \w+>>

(https?\:\/\/[^\/s]+)>>

儲存更改

使用符號將 PTT 的 BBS 句號、逗號、APP 簽名檔、網址等刪除。

斷詞及清除停用詞

≡ 中文斷詞  (15)

參數設定

Input - 27

任務結果

選擇處理欄位 *

result

定義詞彙 ❶

有沒有 500

還沒有 500

沒有問題 500

充電槽 500

億美元 500

選取字典 ❶

-----請選擇-----

儲存更改

依據後續 n-grams 後文字雲結果，增加斷詞及停用詞。

5.情緒分析

Lexicon Based 情緒分析

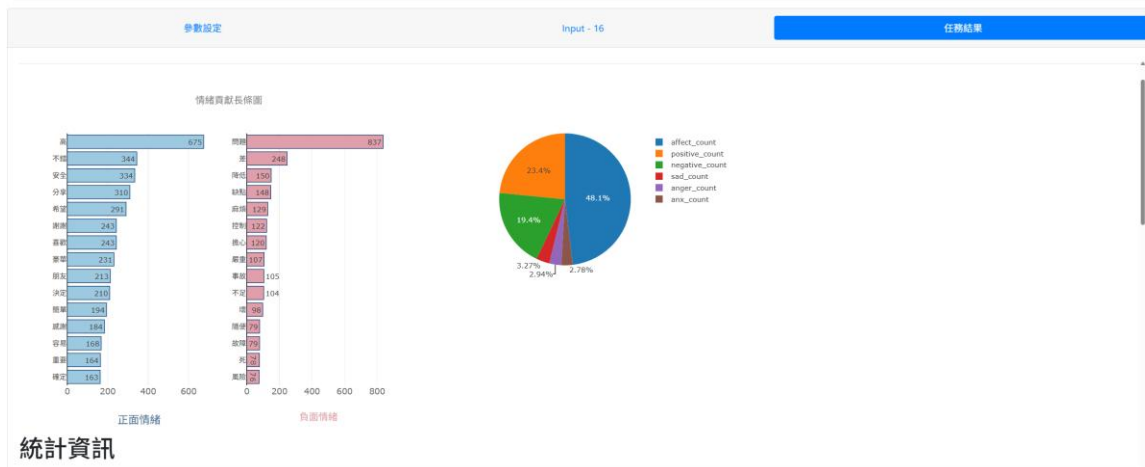
Lexicon Based 情緒分析 (34)



可看出正面詞彙主要為安全、不錯、分享、豪華、喜歡、朋友，表示電動車的
的安全為群眾最常討論的詞彙，而電動車的內裝也是網友討論的主題。

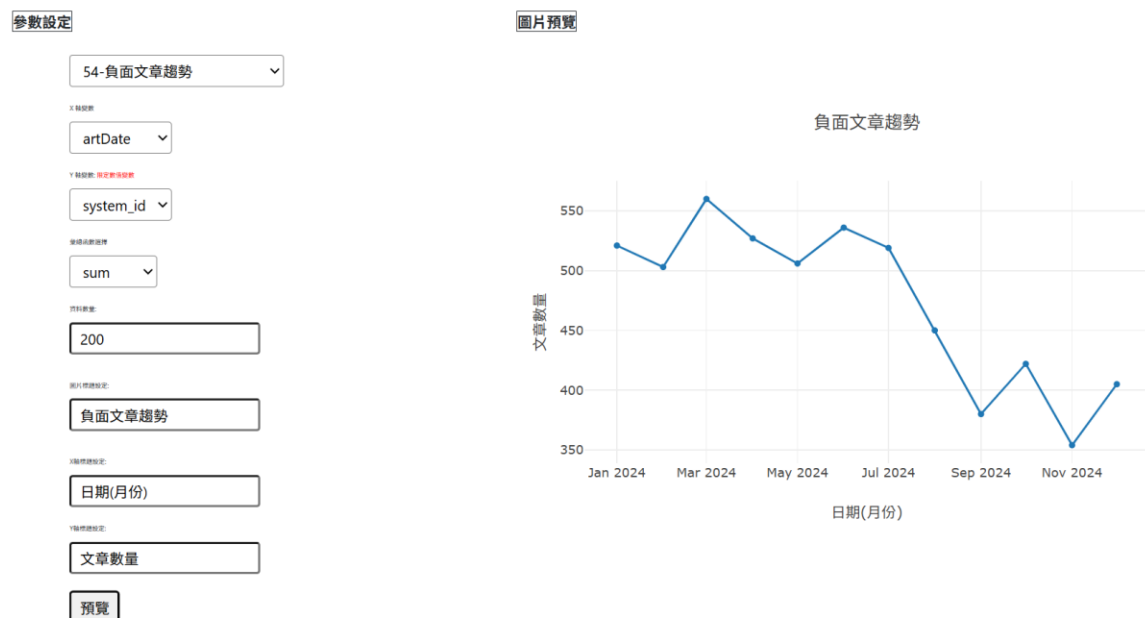
負面詞彙為問題、缺點、降低、事故、控制、風險、麻煩，表示最擔心的為
里程焦慮或是故障，以及控制上的問題。

LIWC 情緒分析



以 LIWC 來說，正面的詞彙仍佔大宗，其中問題並非一定是負面詞彙，但使用沒有問題之類的也無法斷詞，因此不管問題是否適中性字，則正面的評價仍是大於負面的評價。

負面文章趨勢



使用進階字典法後合併原資料 **sysid**，欄位篩選後依日期使用分群匯總，可繪製出負面文章趨勢，可看出整體負面文章是隨著日期慢慢地往下，因此是有逐漸減少的趨勢，表示資料中對於電動車的負面觀感是漸漸減少的。

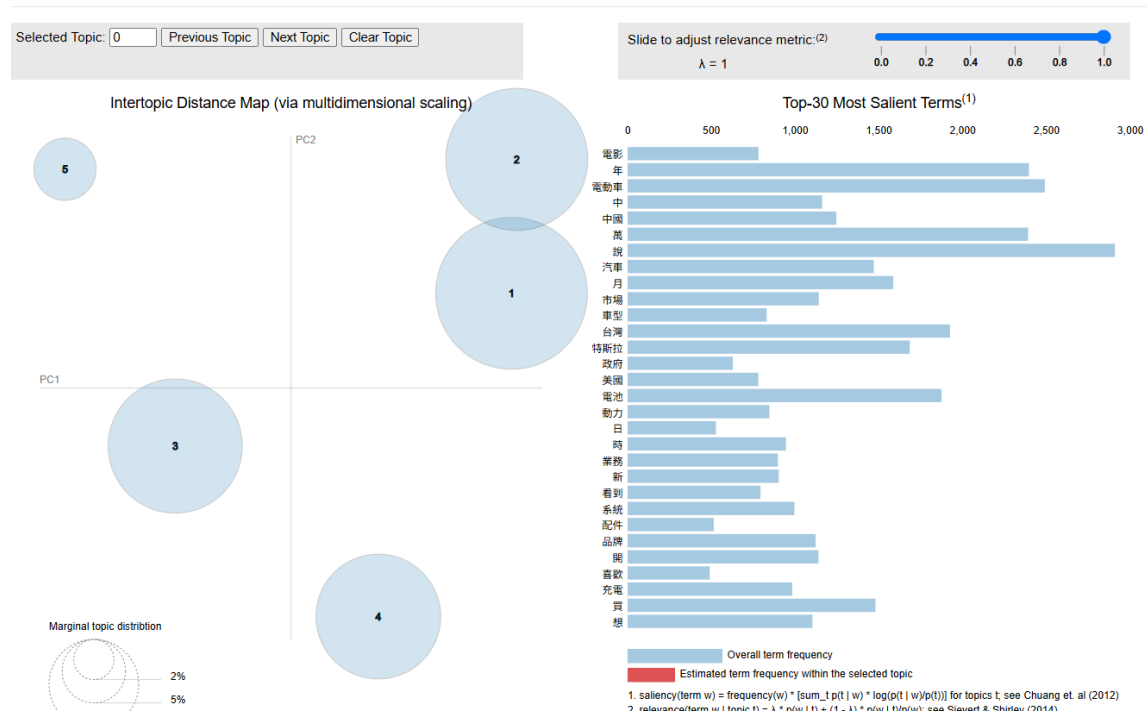
6.主題模型

LDA 主題模型

使用不同主題數，比較 **Cv** 及混淆度後，發現在 5 個主題是會有最高的 **Cv** 及最低的混淆度，因此採取 5 個主題。

LDA主題數	3	4	5	6	7
Cv	0.366	0.38	0.444	0.43	0.438
混淆度	1018.52	1020.6	916.03	948.15	901.24

LDA Vis



主題一關鍵字：空間、買、感覺、問題、喜歡、油耗，主要應是在購買電動車前，討論相對應的電動車相關問題及關注要點。

主題二關鍵字：電池、業務、充電、原廠、車主、配件，應與購買電動車時的一些資訊分享，比如說購車業務、購車配件、購車充電樁等。

主題三關鍵字：電動車、汽車、中國、台灣、特斯拉、市場、品牌、美國、政府，應是討論全球供應鏈的競爭下，各地的電動車銷售與比較，是較為宏觀的主題。

主題四關鍵字：車型、系統、動力、toyota、設計、電池、車款，應是討論電動車的一些規格、新資訊、新變化。

主題五關鍵字：電影、導演、作品：應是討論電動車在電影中的出現。

以這五個關鍵字來說，可看出 PTT 在討論電動車主題上，從宏觀的總體經濟，如中美日各國的電動車銷售與進展，到電動車的新功能及智慧駕駛的突破，到購車前與購車後的相關配件討論，另外因為有找電影版，所以也有電動車在電影中的出現與地位，可看出電動車的討論是由上而下由遠到大的。

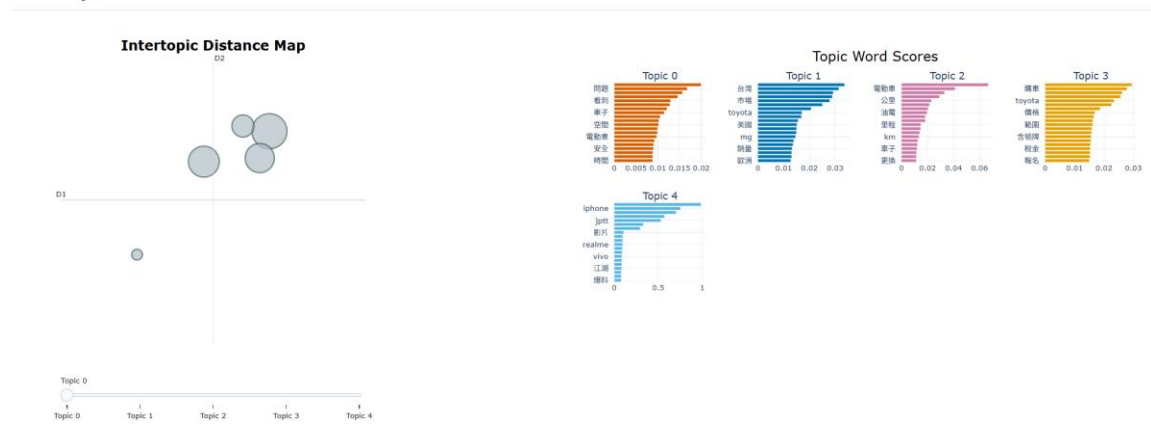
topic	system_id@count
0	1393
1	489
2	824
3	1225
4	156

由分群匯總可看出，主題 1 及主題 4 為最多人，因此對於電動車議題來說，PTT 相關論壇在 2024 年，主要是討論電動車的新科技，以及購買前的討論，為最大

宗，表示消費者傾向於購買前討論做足功課，而購買後的相關分享資訊就會相對較少；而電影中的電動車出現場景，則是較少討論，顯示出目前電動車在電影產業中仍未建立其汽車文化與情懷，電影拍攝會讓人印象深刻的車輛仍為那些經典的老車，比如說華爾街之狼中李奧納多撞壞的那台藍寶基尼 Contach。

Bertopic 主題模型

BERTopic Vis



由 Bertopic 的分析，可看出 0,1,2,3 是較為靠近的，而主題 4 是較遠的，主題 0 為電動車空間、安全等購買前的討論，主題 1 為台灣、美國、歐洲等市場銷量討論，主題 2 為電動車里程、更換等電動車科技的討論，主題 3 為購車時的稅金、價格、購買後相關問題的討論，至於較遠的主題 4 則為電影方面的討論。

7.字詞網絡圖

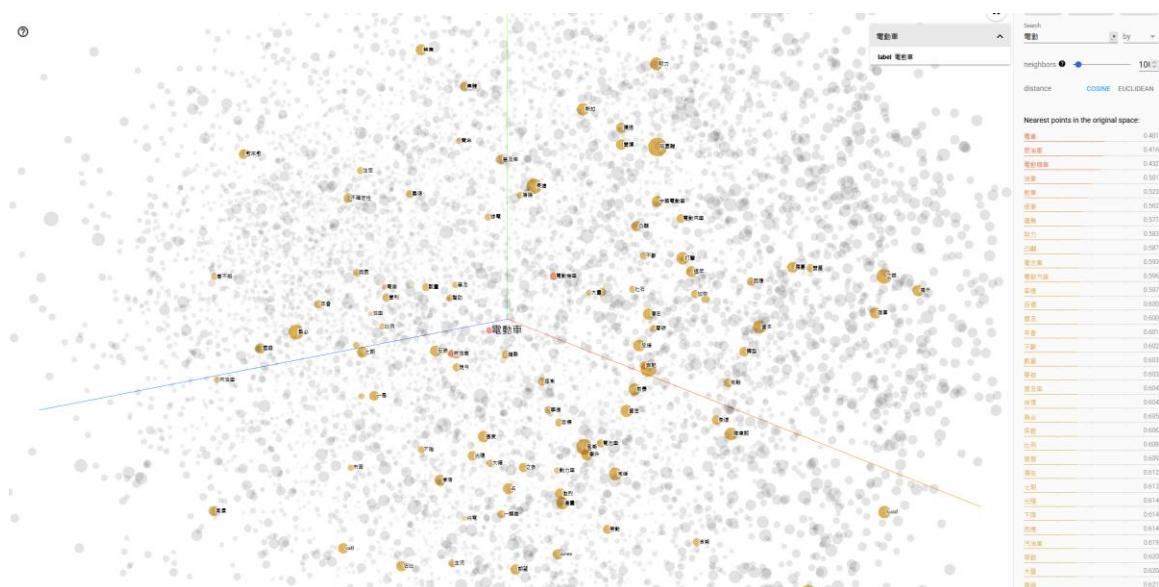
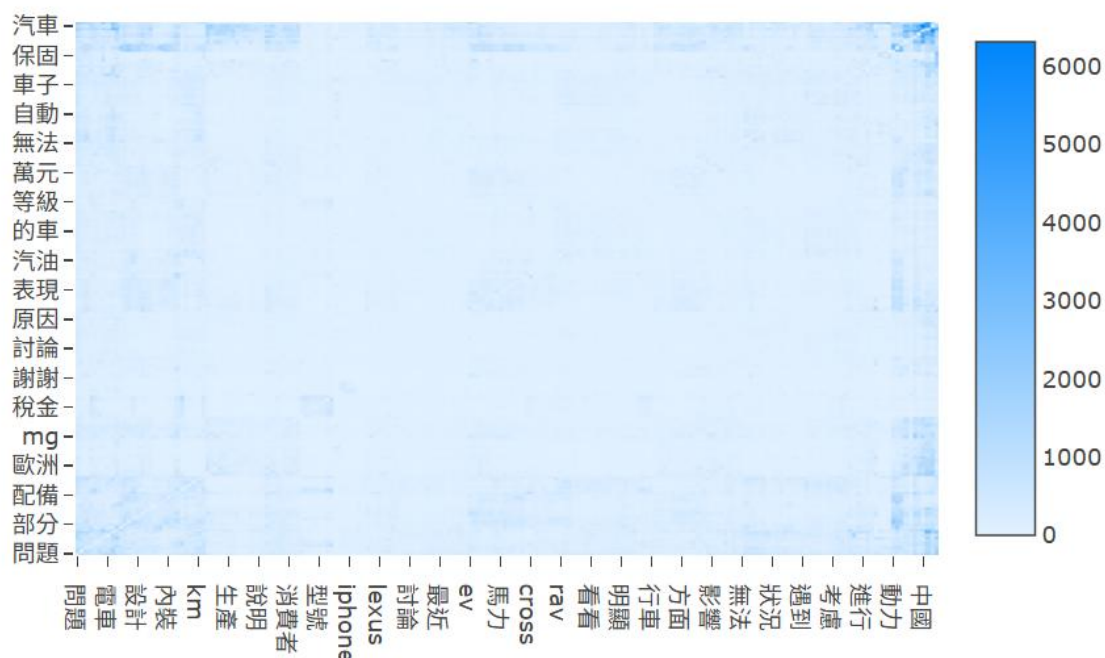
我們使用清除停用辭後轉為 DTM，繪製關聯式文字雲(shiny)，也在計算共現矩陣後，欲繪製字詞網路圖(Shiny)，但由於 Shiny 卡住無法顯示，因此我們轉為使用

Word2Vec，討論各字詞的向量輸出成 tsv 檔案，將 word_emb 和 word_metadata

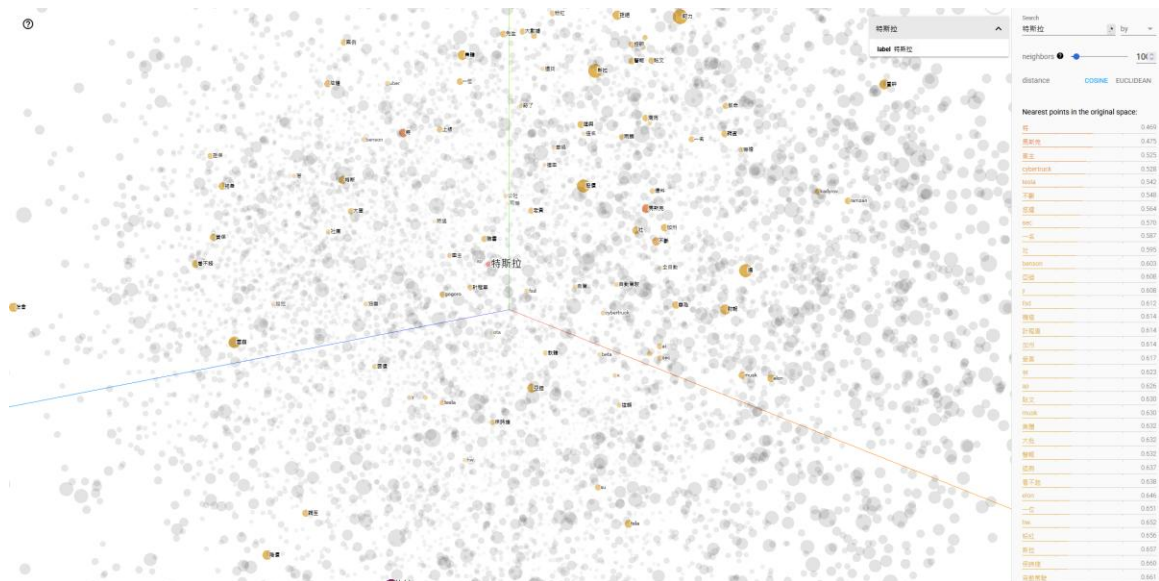
上傳至 tensorflow 網站，觀察各字詞的共現性。

下圖為共現矩陣，但轉為 shiny 時輸出失敗，可能是期末報告太多人使用資源

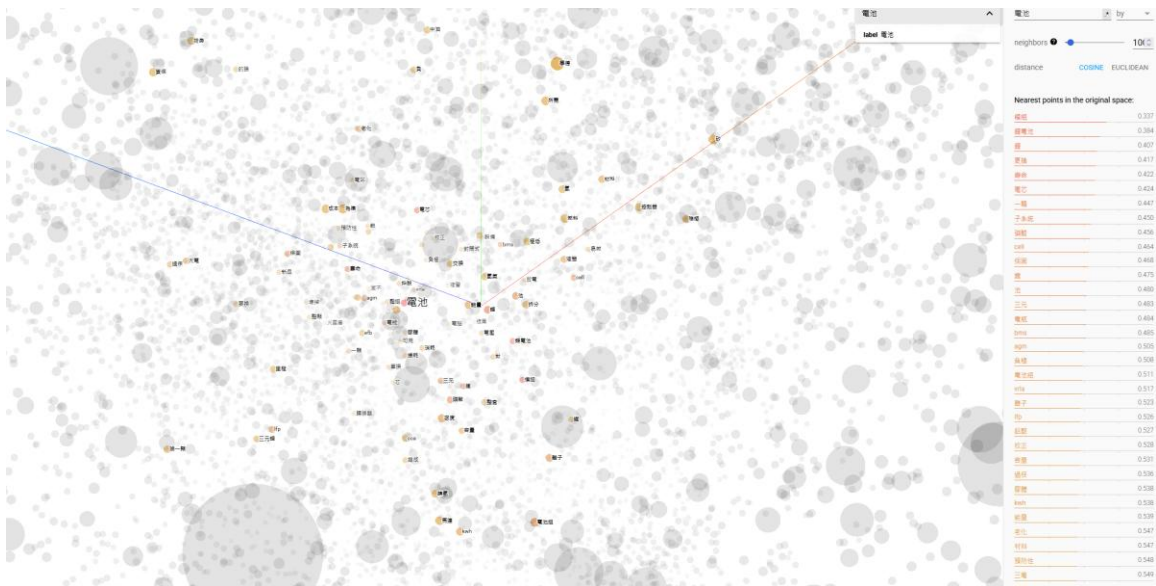
熱圖



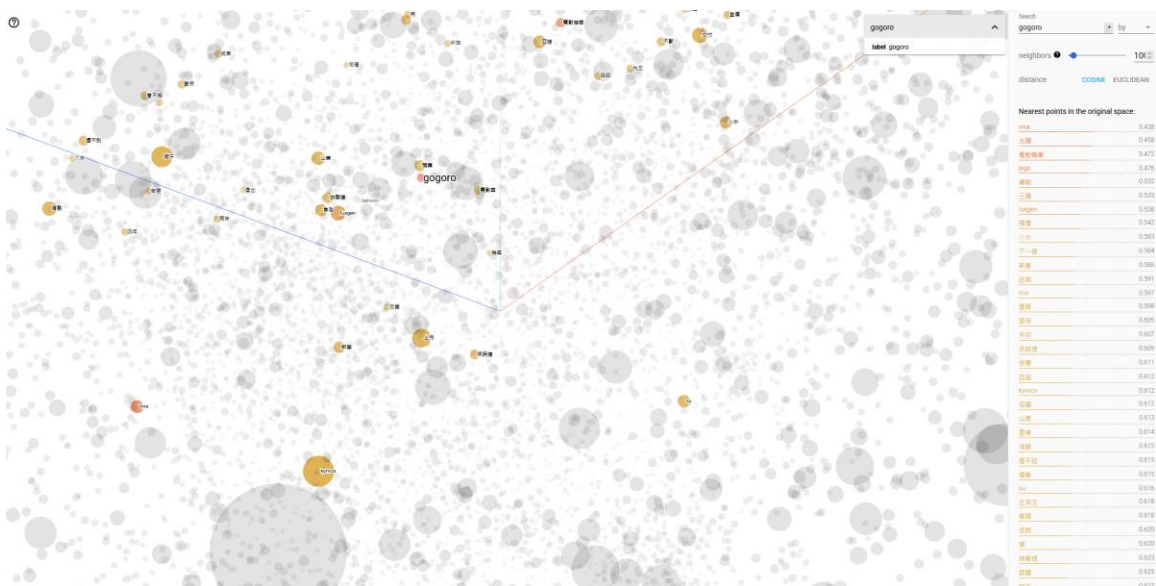
Tensorflow 如上圖繪製後，首先查詢關鍵字電動車，其貢獻較多的字詞由多至少為燃油車、電動機車、油車，表示很多電動車的討論都會將電動車與燃油車等做比較，而後續有趨勢，寧德(電池廠商)等等的字詞；點擊寧德後，由寧德電池為中心往外的是寧德固態、矽、氫燃料、氫能源、石磨，表示寧德電池在電動車領域擁有領先討論地位，許多人在討論中國的稀土元素以及電動車電池發展。



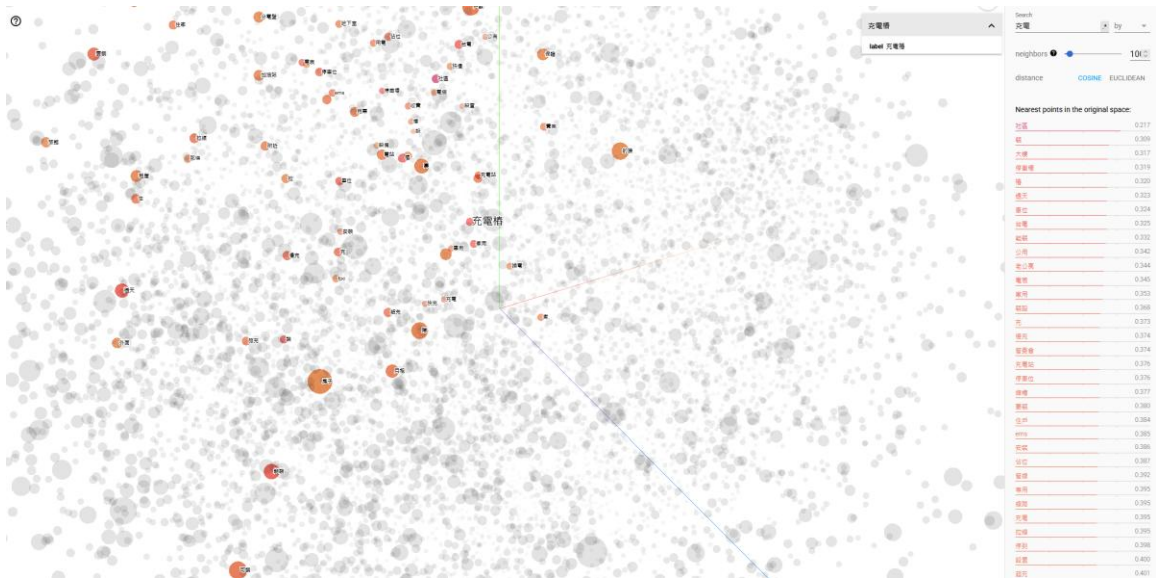
上圖為查詢特斯拉，前幾大共現字詞為馬斯克、車主、cybertruck、炫耀、比亞迪、計程車、加州，表示特斯拉討論的主題，多著墨於特斯拉新科技的 cybertruck 以及 robotaxi，以及特斯拉與加州的關係。



查詢電池後，最多的是模組、鋰電池、更換、壽命、磷酸、cell、三元，表示電池方面，大家會關注電池的使用壽命，種類不同，容量衰退，以及安全性等。



搜尋 gogoro，發現會有 viva(新產品名稱)、光陽、jego(新產品名稱)，然而往下會有看到雪崩，表示 gogoro 除了新產品及競爭對手受到矚目一起討論以外，也有針對銷量減少的現象有所討論。



針對充電樁的討論，台灣仍有社區、透天、台電、停車場、老公寓、管委會、住戶、ems 等討論，因詞由字詞可看到台灣的居住型態造成電動車使用者對於充電樁的安裝仍有諸多限制。

8.社會網絡資料

分群彙總 (非數值) (103)

參數設定

Input - 99

任務結果

使用...欄位進行分群(按住ctrl(Windows)或command(MAC)可以複選) *

system_id

comment_idx

cmtPoster

artPoster

匯總函數 *

count

nunique

min

max

first

last

sum

計算欄位(按住ctrl(Windows)或command(MAC)可以複選) *

system_id

comment_idx

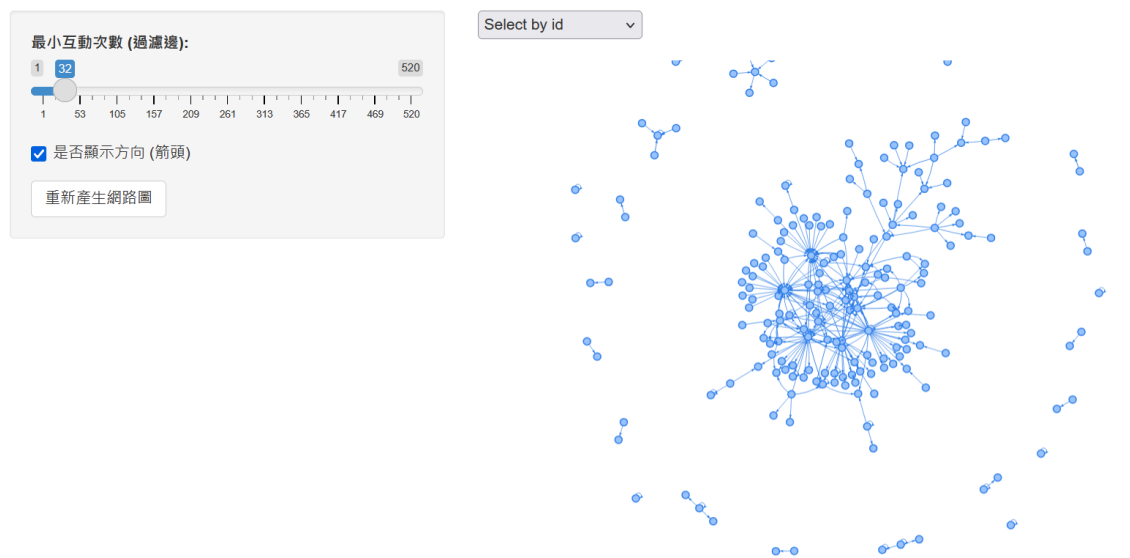
cmtPoster

artPoster

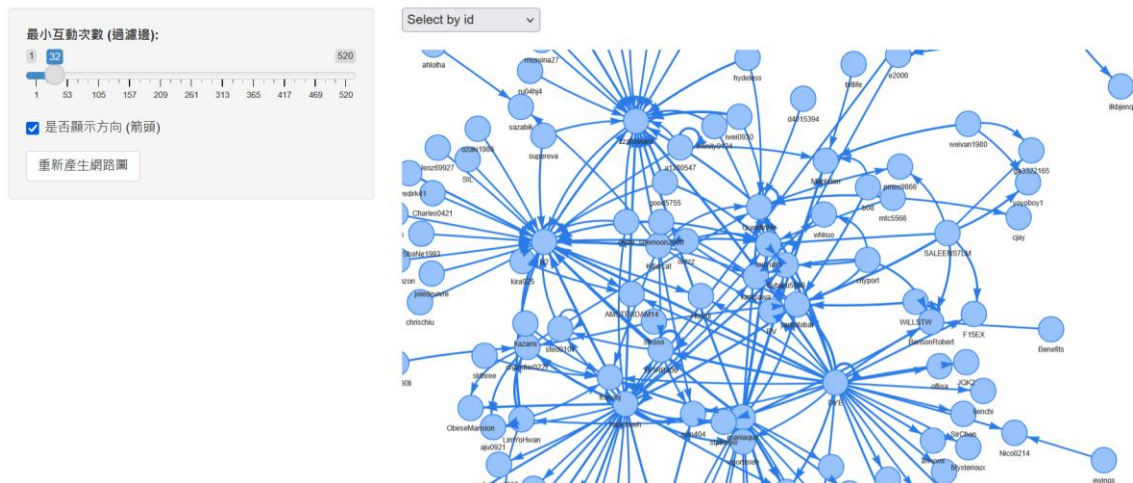
使用留言萃取功能，將討論文章中的留言萃取，利用欄位篩選分別篩選後，使用 `system_id` 做交集將留言者 `comment_idx`、`cmtPoster` 及發文者 `artPoster` 做資料合併，使用分群匯總後使用 Shiny 網路圖繪製。

因 Tarflow 的 Shiny 社會網路圖功能有時候會卡住，因此我們組員將分群匯總後的數值，使用 R 語言繪製。

PTT 互動網路圖

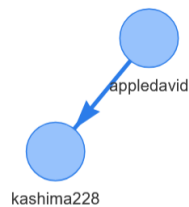


PTT 互動網路圖

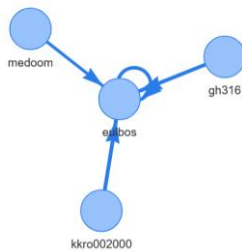


將最小互動次數增加到 32，可以看出中間有許多意見領袖，而邊緣有幾個兩兩成對的使用者，此網絡型態除了周圍的衛星型態以外，有一點點像老師上課所說的 **small world graph**，而造成周圍衛星型態，放大觀察各衛星，大多都是單向的，意思就是可能是某一個人 **PO** 文，另一個人他去該文章下熱烈推文討論(單箭頭)，或是兩人在文章下面討論該議題(會看到發文者有個迴旋箭頭)

單箭頭示意圖，表示 **appledavid** 在 **kashima** 的文章裡發表高見



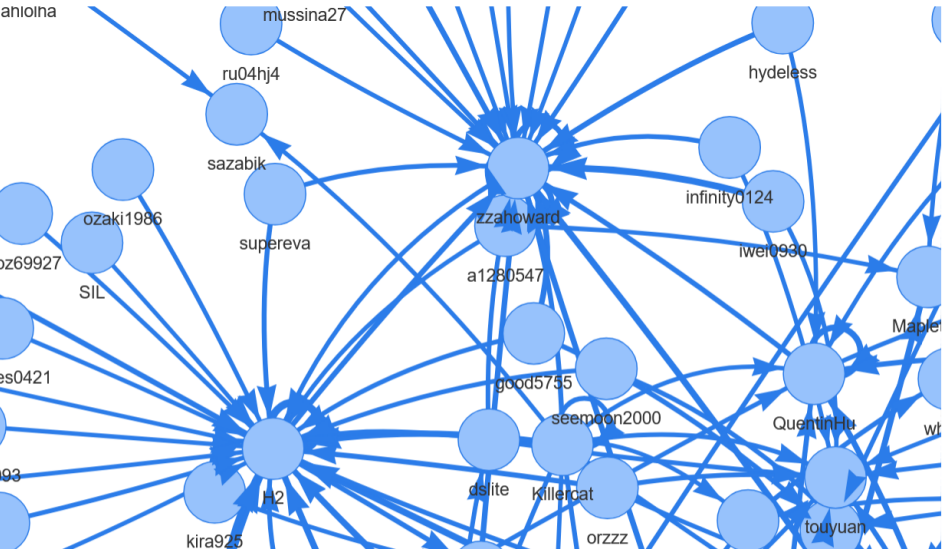
此三人包含發文者 **eulbos**，在 **eulbos** 的文章裡面熱烈討論



如將中心網絡處放大看，如有使用者是多箭頭指向他，表示他發表的文章是較有共鳴或是回響多的主題，也有可能是爭議性較大的主題，實際查詢較多人回應的 **zzahoward**、**H2**；我去汽車版查詢了這兩個作者的文章，**zzahoward** 在 2024 年在 **car** 汽車版就發表了 90 篇文章，大概 4 天就一篇，有發現他除了電動車新聞以外，也針對電動車的使用以及疑慮，有很多的回文，他將自身電動車的使用經驗，

分享至論壇上推廣，並引起很多人有共鳴討論；查詢 H2 在 2024 年 car 汽車版更是發表了 150 篇文章，他主要是針對一些電動車的 youtube 影片或是爭議新聞畫面，拿到討論版上分享，也引起了很多的討論。

兩位電動車議題活躍使用者 zzahoward 以及 H2



zzahoward 在汽車版的討論文章，對於電動車及汽車有很多討論

【板主:Pellaeon】				[汽車] 禁止政治相關推文		系列《car》
[<-]離開 [→]閱讀 [Ctrl]-P]發篇文章 [d]刪除 [z]精華區 [i]看板資訊/設定 [h]說明						
編號	日期	作者	文章	標題	人氣:348	
224	5310/30	zzahoward	[新聞]	Mazda 6 接班人售價出爐！純電、油電雙		
225	+1511/02	zzahoward	R: [問題]	Toyota被盜刷手排Supra導致崩盤？		
226	6111/08	zzahoward	R: [新聞]	台中電動車數破11萬 民愛大樓地下停車場		
227	2311/09	zzahoward	R: [閒聊]	電動車開起來像油車是好事嗎		
228	1611/15	zzahoward	R: [新聞]	南韓電動車今晨連續「起火自燃」 皆使去		
229	+2211/17	zzahoward	[閒聊]	Carwow 2024 top 10高妥善率排名		
230	+1911/18	zzahoward	[情報]	Polestar 成都、重慶關廠		
231	+1311/24	zzahoward	R: [閒聊]	Volvo EX90 超感人的汽車廣告		
232	+2611/29	zzahoward	R: [新聞]	日產高層直言「只剩14個月」挽救公司		
233	+512/02	zzahoward	R: [心得]	電動車大電池原廠保固的真相（文長）		
234	+2012/04	zzahoward	R: [新聞]	Peugeot與Citroen台灣代理商換人 尚騰		
235	+2212/04	zzahoward	R: [閒聊]	新莊Volvo濫告消費者		
236	+4912/04	zzahoward	R: [閒聊]	新莊Volvo濫告消費者		
237	+1112/09	zzahoward	[閒聊]	Mercedes新一代MMA (BEV+HEV)平台		
238	M3912/11	zzahoward	[心得]	車損、賠償以及保險		
239	+爆12/15	zzahoward	R: [閒聊]	汽車關稅降為0根本圖利進口車商		
240	+6312/16	zzahoward	[新聞]	電動車數量成長 交通部預估2030年開收汽		
241	+2012/17	zzahoward	R: [問題]	xc40電車避震升級建議		
242	+8712/26	zzahoward	R: [心得]	台灣賓士讓你體驗新車開12000換引擎		
243	+90 1/15	zzahoward	[閒聊]	有社區似乎開槍地下室禁停有"鋰電池"車輛		
文章總覽				(v)回應(X)推文(^X)轉錄 (= <>)相關主題(/?a)找標題/作者 (b)進板畫面		

至於 DYE 他有很多箭頭往外，表示他不常發文，但是常常在很多人的文章下面，留很多的留言，表示他雖然有一些專業知識和發表，但是他不同於有些人會採用的回文，反而是使用留言的方式表達，或許他是有些低調，但是又是十分熱心有分享慾的，如果他持續高品質且正確中立的發言，使更多的使用者受益或是有共鳴，那久而久之就容易成為電動車或是汽車版的意見領袖。下圖為 DYE 針對電動車討論

14f → DYE: BZ3 雖然三電來自BYD，但沒有證據指出整車是中國車款貼牌。	11/28 10:34
15f → DYE: 小心造謠水桶。	11/28 10:34
23f → DYE: 如同以前很多車款用王子引擎搭配變速箱，但沒有人說這是貼牌	11/28 10:55
24f → DYE: 車。	11/28 10:55
26f → DYE: 對於電動車的三電供應商，重要程度等同以前油車的引擎加變速	11/28 10:56
27f → DYE: 箱。	11/28 10:56
28f → DYE: 或是例如Pagani用MB的引擎，也沒人說這樣就是貼牌車。	11/28 10:59
29f → DYE: 而把一台豐田的車款，因為三電的供應商是中國品牌，而說成是	11/28 11:05
30f → DYE: 貼牌車，這個抹紅程度是比較高的。例如特斯拉也用BYD電池與	11/28 11:05
31f → DYE: 電控，加上也用中國中科三環的電機，這樣也沒人說他是貼牌車	11/28 11:05
32f → DYE: 。	11/28 11:06
33f → DYE: 以前說BZ3是貼牌BYD，主要是諷刺豐田做不好電動車而去找BYD	11/28 11:11
34f → DYE: 幫忙，但現在本板不允許諷刺，例如ZZ因為諷刺台奧CEO而水桶	11/28 11:11
35f → DYE: ，請發文注意。	11/28 11:11
43f → DYE: 油車改epower，因為電池小沒什麼重量增加問題。油車改增程就	11/28 12:01
44f → DYE: 跟油改電一樣有電池放置與車重的問題，只是增程的電池量還沒	11/28 12:02

結論

由此次的分析，我們藉由 tarflow 的 PTT 爬蟲，抓取了關於電動車的討論議題，經過情緒分析、主題模型、字詞共現、社群網絡等多個面向，主要發現台灣電動車討論，探討主題為購車前考量，以及購車相關資訊，也關心全球市場競爭以及技術規格討論，正處於認知建立與接受度提升期，在論壇中有經驗的消費者與潛在消費者的互動，融合情緒分析，表示市場由初期的懷疑漸漸有轉向積極了解，而其中消費者最關注的議題為實用性、充電便利性、與燃油車的優劣，以及全球電動車品牌的競爭。那對於車廠的行銷建議，可由社群討論尋找電動車痛點解決，並提升

電動車爽點，尋找社群意見領袖，邀請真實體驗分享而非單純產品宣傳，由論壇與 youtube 影片雙管齊下提升討論度。

然而從充電樁的討論，以及對於電動車缺點的續航中可看到，台灣在電動車推廣除了價錢因素以外，仍有集合式住宅安裝電動樁不易的高門檻，因此如果想要提升電動車的銷售，這方面痛點仍賴業者與政府溝通，取得人民信賴一同建立有效解決方案。