

社群媒體分析-讀書會 II

振興五倍券 使用情緒分析

第7組

林建玄 田又仁 林志勳

曾子恬 林明禾 陳雅詩 施善騫

500

振興五倍券

刺激消費 加碼加倍 —— 專區



Agenda

- 研究動機與分析目的
- 分析流程與資料前處理
- Corpus Based 情緒分析
- 情緒分類器訓練

任務要求

1. 包含一個自行訓練的分類器 (分類任務不限)
2. 需包含corpus-based 情緒分析

研究動機與分析目的

- PTT八卦版蒐集自 **2021/08/01~11/30** 和振興券相關議題之討論文章, 共計 **858筆** 資料
- 振興券從政策討論到正式實施至今, 已有一段時間, 想探討在這段期間 **內鄉民們的情緒變化**

搜尋關鍵字

五倍卷
5倍卷
振興券
藝fun券
好食券
農遊券
國旅券
動滋券
客庄券
i原券
地方創生券
台北熊好券
熊好券
新北加倍券
加倍券
桃園好市乘雙券
好市乘雙券
數位桃園1000+
新竹買爆券
買爆券
滿竹券
台南好康券
好康券
高雄券

5

欄位數



858

資料筆數



任務結果

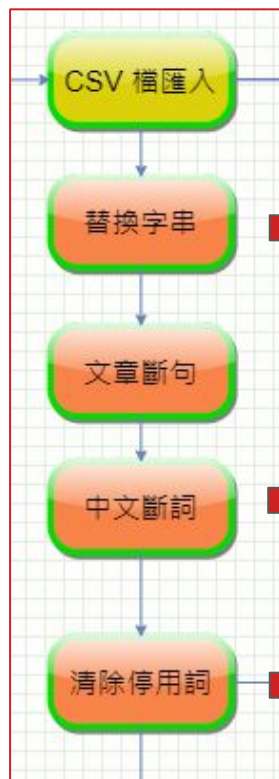
Show 10 entries

system_id	artTitle	artDate	artTime	artUrl	artContent
38	[問卦]覺得振興紓困發錢好還是券	2021/08/05	03:40:59	https://www.ptt.cc/bbs/Gossiping/M.1628134861.A.11E.html	餓死抬頭_振興經濟選擇有很多_抽獎 發錢 消費券振興券等_確實發現金有北存起來 (畢竟省本_是不是比較好的方_卦? _Sent
55	[新聞]陳其邁:中央若推「振興券」 高雄	2021/08/07	04:55:34	https://www.ptt.cc/bbs/Gossiping/M.1628312138.A.E5C.html	陳其邁:中央若推「振興券」 高雄配合加碼_https://www.bcc.com.tw/newsView.6 之後, 高雄市長陳其邁今天(6日)也表示, 如果中央決定發放「振興券」, 高雄振興券, 我們會再加碼, 譬如之前消費券時, 高雄推出過_夜市券、商圈券、時時抽 陳其邁說, 如果中央推出「振興券」, 高雄會加碼, 加強

分析流程圖



資料前處理



替換字串設定 ①

```
\n>>  
https?:\V(www\,)?[-a-zA-Z0-9@:%_\+~#={1,256}\.[a-zA-Z0-90]{1,6}b([-a-zA-Z0-90]@:%_\+~#?  
&/=]*)>>  
\W>>  
Sent from JPTT on my>>
```

Sent from JPTT on my>>
Sent from BePTT on my>>
5倍券>>五倍券
3倍券>>三倍券
藝FUN>>藝放

定義詞彙 ①

不如預期 100
不念書 100
不後悔 100
不要領 100
不能用 100
不能信用 100

自定義停止詞

完整
連結
內文
文章
違者
刪除

實體券 10000	數位轉型 100
實體五倍券 3000	數位券 100
蔡英文 1000	數位加碼 100
數位五倍券 1000	銀行 100
實體三倍券 1000	臺灣價值 100
電子支付 1000	熊好券 100
業者 1000	跨縣市 100
超商 1000	裝逼 100
健保卡 1000	綁定 100
紓困金 1000	新台幣 100
振興券 1000	損龜 100
夜市券 1000	搶頭香 100
沒有幫助 1000	搶米 100
台灣人 1000	想問 100
五倍券 1000	嫌麻煩 100
三倍券 1000	搭綠班 100

去除版規/新聞連結/注意事項等用語

每日文章數量

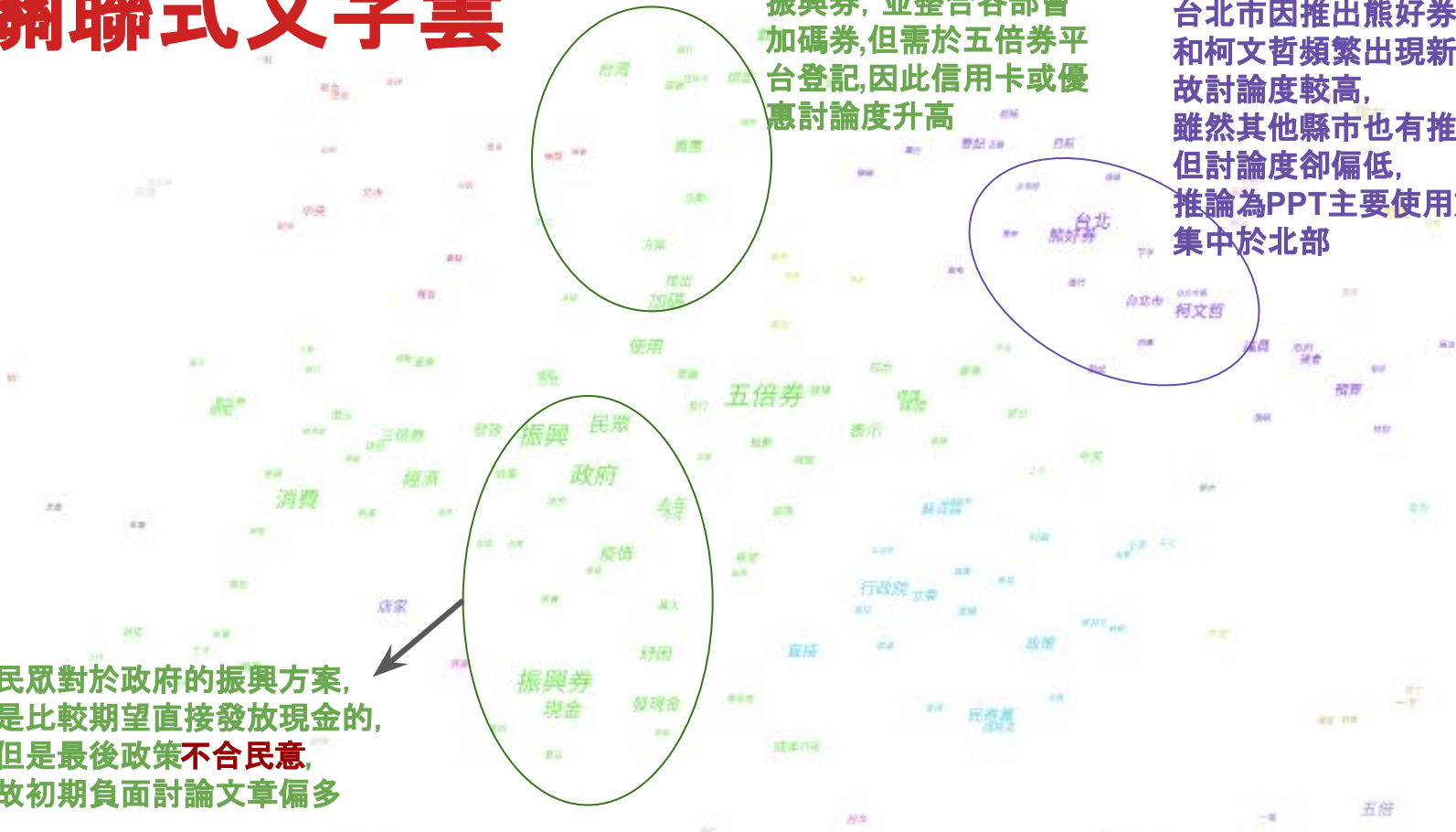


關聯式文字雲

政府鼓勵民眾綁訂數位振興券，並整合各部會加碼券，但需於五倍券平台登記，因此信用卡或優惠討論度升高

台北市因推出熊好券
和柯文哲頻繁出現新聞，
故討論度較高，
雖然其他縣市也有推出加碼券，
但討論度卻偏低，
推論為PPT主要使用族群
集中於北部

民眾對於政府的振興方案，
是比較期望直接發放現金的，
但是最後政策**不合民意**，
故初期負面討論文章偏多



字詞網絡關係圖

關鍵字網路分析工具

詳細設定

社會網路圖

others

選擇次層詞類

五倍券

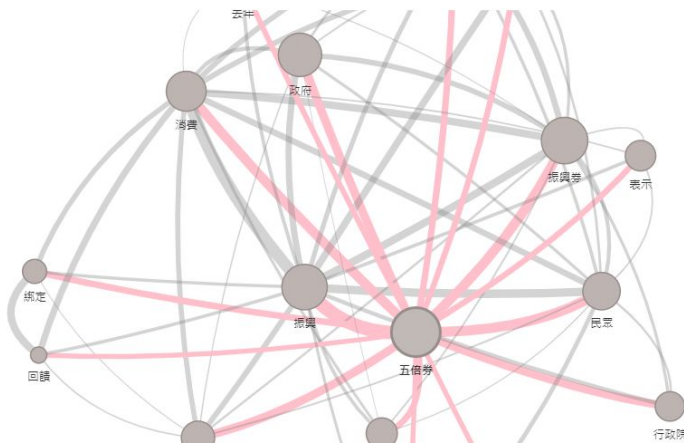
設定網路節點數量



依關聯強度篩選鏈結



others



Recent Co-Accurance (200): 五倍券

是這樣的最近版上幾乎都在噴**五倍券**的政策但我覺得明明就還稱讚多了一小筆錢可以去買自己喜歡的東西這樣有不好嗎 有沒有掛

1 媒體來源 水果日報2 記者署名 鄭柏揚3 完整新聞標題 **五倍券**異見多 綠委也有雜音 蘇貞昌堅持 振興券限期使用比發現金好4 完整新聞內文 政府去年發放振興三倍券刺激經濟 今年遇到本土疫情大爆發 在許多行業受到前所未有的重創下 行政院5日證實會參考去年經驗 再推**五倍券**救經濟 但此舉引起在野黨不滿 主張應發現金較為妥當 就連綠營黨內也有雜音 行政院原定祕

2-grams 文字雲

從圖上可觀察出除了振興券領取之方式討論外，
相關議題還有是否要改發放現金或者期待地方政府之加碼優惠

振興 五倍券

想要發放現金的討論

地方政府或業者 加碼優惠之討論

Corpus Based 情緒分析

完成資料前處理後，我們將整理好的資料集進行情緒分析，再藉由分類彙總，來對資料進行視覺化圖表與解讀

Mini Workflow Design Platform

Corpus Based 情緒分析 (126)

參數設定

Input - 15

任務結果

統計資訊

476

負向情緒數



0

正向情緒數



381

總數



任務結果

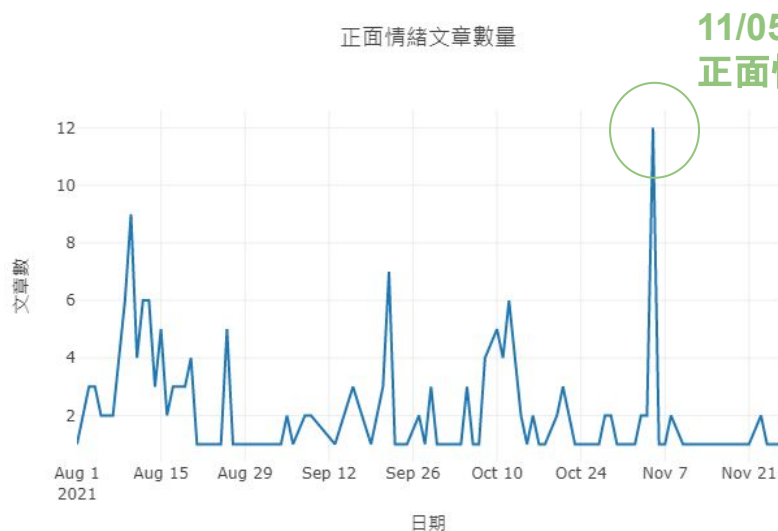
Show 10 entries

Search:

system_id	sentence_idx	sentiment_value
1	1	0.000
2	1	0.000
3	1	0.000
4	1	0.088

Corpus Based 情緒分析

由下面兩張圖可以看出振興券實施後，民眾對於振興券之情緒反應，從政策討論期(08/1~09/22)可以看出負面文章明顯較多，但開始實施後，開始有少部分正面情緒文章開始出現



情緒分類器訓練 取得分類的csv資料

從PTT爬蟲下載的原始資料集，沒有情緒分數可以進行分類標籤，故透過文字分析平台裡 內建的「文章+類別+情緒」輸出有情緒評分欄位的 csv檔，並在csv檔裡面 **建立資料的分類標籤 (情緒正向、負向)** (分類標籤設定條件 $\text{positive_emotion_grade} > 0.8$ 標籤為正向， $\text{negative_emotion_grade} > 0.8$ 為負向，如果上述兩條件皆不符合則不標籤也採用)，分類完成的資料集另存成 csv檔提供訓練分類器使用

中山管院文字分析平台

帳號管理

選擇的文件集: Project1

文件集

設定

抽樣處理

全部處理

資料探索

第二步:選擇輸出欄位

輸出欄位

文章 + 類別 + 情緒

請選擇類別辭典

重新整理

設定辭典

所有類別辭典

space_default

family_default

insight_default

case_100family

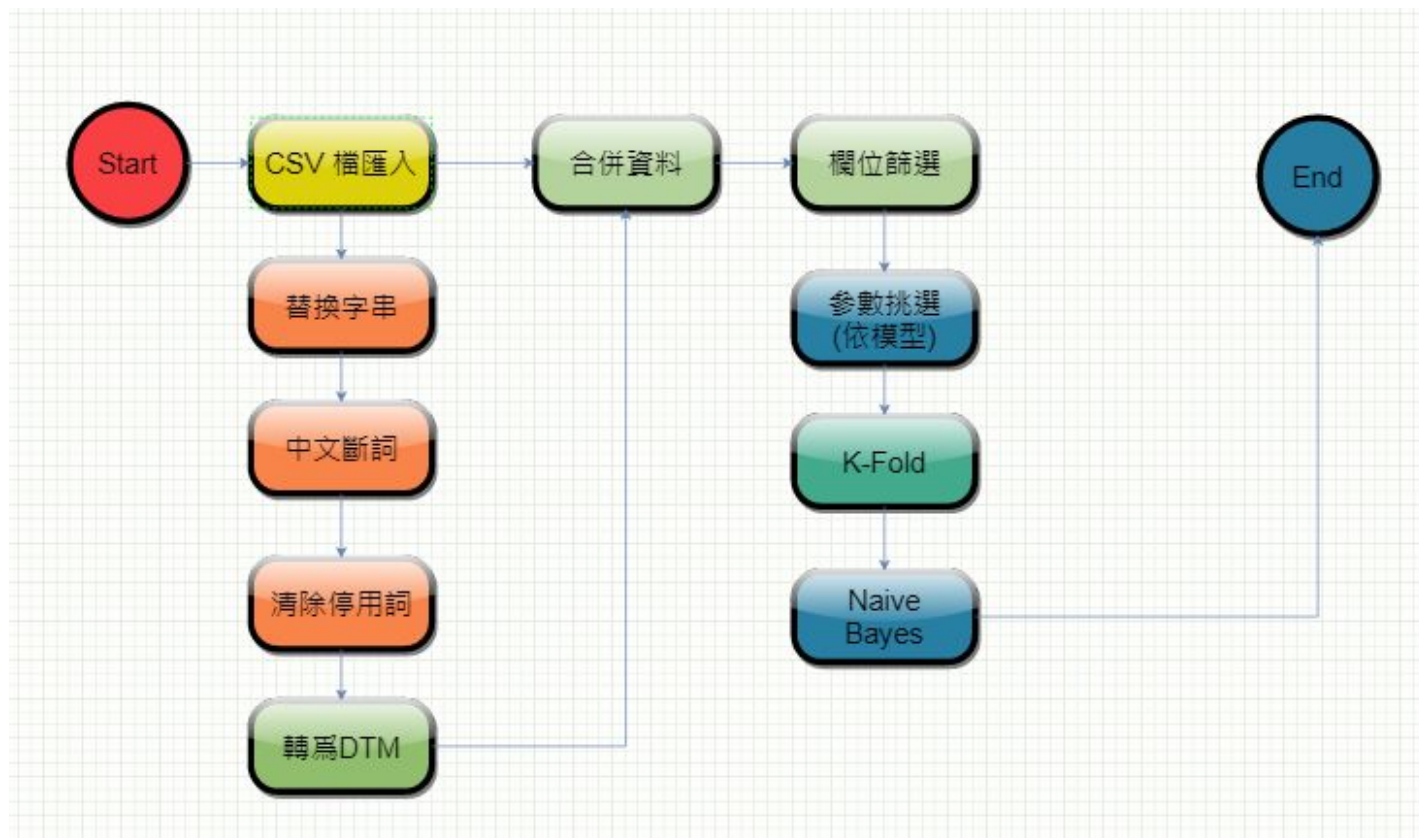
已選擇類別辭典

money_default

artTitle	artDate	artTime	health(種類1)	family(種類2)	positive_emotion	negative_emotion	neutral_emotion
文章標題1	2018/03/06	08:45:00	3	5	0.33	0.66	0
文章標題2	2018/03/07	08:45:31	4	0	0.25	0.25	0.5

d	artTitle	artDate	artContent	positive	negative	neutral	emotion	type
1	Re:[新聞]1個月當機4次	2021/11/29	我好奇的是誰要負責?_阿凱	0	1	0	負向	
2	[問卦]高雄卷到底是在	2021/11/26	最近高雄開了幾家飯店小弟	0	1	0	負向	
3	[問卦]對政府印鈔票無	2021/11/21	靠北房價一直漲已經是本板的	0	1	0	負向	
4	[問卦]欸大家台北熊好	2021/11/19	我新北人啦 但是中獎了_當然	0	1	0	負向	
5	[問卦]中獎機率35%還	2021/11/8	聽說新北加倍券開獎卷_我登	0	1	0	負向	
6	[問卦]淺談台灣結構性	2021/11/8	說真的現在台灣消費力大不如	0	1	0	負向	
7	[問卦]熊好券店家說沒	2021/11/6	拿著台北市民稅金的熊好券	0	1	0	負向	
8	[問卦]熊好券掃捲完店	2021/11/6	剛吃飯完,要用熊好券付款時	0	1	0	負向	
9	[問卦]熊好券的店家查	2021/11/5	是否有專板_本板並非萬能問	0	1	0	負向	
10	[問卦]熊好券適用店家	2021/11/5	是否有專板_本板並非萬能問	0	1	0	負向	
11	[爆卦]熊好券安卓不能	2021/11/5	因為我的豆腐頭遺落在暖男的	0	1	0	負向	
12	Re:[爆卦]柯文哲電梯前	2021/11/3	先不評論柯文哲和呱吉平常的	0	0.875	0.125	負向	
13	[問卦]恭喜身分證771.	2021/11/2	是這樣子的,五倍卷抽獎,	0	1	0	負向	
14	[問卦]大家新北加倍券	2021/11/1	上周新聞新北加倍券_5000變	0	1	0	負向	
15	[問卦]有沒有最黑的身	2021/10/29	安安 我肥宅啦_這個智障政府	0	1	0	負向	
16	[問卦]熊好券要被延	2021/10/27	如題_剛剛看到新聞_議會又	0	1	0	負向	
17	[問卦]農遊券app登不	2021/10/25	小弟上週爽中農遊券,今天還	0	1	0	負向	
18	[問卦]五倍券還沒花怎	2021/10/24	如題_第一天拿了_到現在放	0	1	0	負向	
19	[問卦]新陳之國府建國	2021/10/23	新陳之府建國建國時_就最	0	1	0	負向	

情緒分類器訓練一流程圖



情緒分類器訓練 匯入已分類的csv檔

≡ CSV 檔匯入  (30)

參數設定

任務結果

統計資訊

9

欄位數



303

資料筆數



任務結果

Show 10 entries

Search:

system_id ▲	artTitle	artDate	artTime	artUrl	artContent	positive_emotion_grade
-------------	----------	---------	---------	--------	------------	------------------------

					我好奇的是誰要負責?_阿凱出來處理一下_洗一洗風向_台北通當機都用_「零件在海上」_買不到帶過_應該宣一宣_系統是那家公司承包的_是不是都是這家做的?_背後水是不是很深_不過在八卦_這件事情會被_小凱洗推_來看看9/22的新聞_行政院推出的振興經濟五倍券,今(22)日上午開放民眾上網登記數位券登	
--	--	--	--	--	--	--

情緒分類器訓練 特徵值調整與訓練結果檢視

參數挑選 (依模型) (20)

參數設定

Input - 18

目標欄位 *

emotion_type

最大特徵數

200

Naive Bayes (22)

參數設定

Input - 21

任務結果

統計資訊

0.004

訓練時間



0.009

推論時間



0.672

測試資料準確度



0.672

測試資料micro-F1



0.665

測試資料macro-F1



0.679

測試資料加權F1



0.672

測試資料micro精確率



0.67

測試資料macro精確率



0.708

測試資料加權精確率



0.672

測試資料micro召回率



0.684

測試資料macro召回率



0.672

測試資料加權召回率



全屏瀏覽

點我下載完整CSV資料

點我下載完整Rdata

最大特徵值200

情緒分類器訓練 特徵值調整與訓練結果檢視

≡ 參數挑選 (依模型) (20)

參數設定

目標欄位 *
emotion_type

最大特徵數
30

Naive Bayes (22)

參數設定Input - 21任務結果

統計資訊


0.002 訓練時間	0.009 推論時間	0.59 測試資料準確度	0.59 測試資料micro-F1
0.586 測試資料macro-F1	0.576 測試資料加權F1	0.59 測試資料micro精確率	0.68 測試資料macro精確率
0.73 測試資料加權精確率	0.59 測試資料micro召回率	0.654 測試資料macro召回率	0.59 測試資料加權召回率

全螢幕瀏覽點我下載完整CSV資料點我下載完整Rdata

最大特徵值30

最大特徵值 30

情緒分類器訓練 訓練結果檢視與調整

≡ 參數挑選 (依模型)  (20)

參數設定

目標欄位 *
emotion_type

最大特徵數
40

Naive Bayes  (22)

參數設定Input - 21任務結果

統計資訊

0.003 訓練時間	0.011 推論時間	0.738 測試資料準確度	0.738 測試資料micro-F1
0.736 測試資料macro-F1	0.743 測試資料加權F1	0.738 測試資料micro精確率	0.766 測試資料macro精確率
0.827 測試資料加權精確率	0.738 測試資料micro召回率	0.789 測試資料macro召回率	0.738 測試資料加權召回率

全螢幕瀏覽點我下載完整CSV資料點我下載完整Rdata

最大特徵值調整到
40時，訓練器可以獲得最佳的準確度

最大特
徵值 40

情緒分類預測結果



將一開始未分類的 **858筆** PTT 爬蟲原始資料集丟入訓練好的分類器進行預測

分類預測 (16)

參數設定 Input - 15 任務結果

Show 10 entries Search:

system_id	result
1	正向
2	正向
3	正向
4	負向
5	正向
6	正向
7	負向
8	正向
9	負向
10	正向

Showing 1 to 10 of 100 entries Previous 1 2 3 4 5 ... 10 Next

