



# 關鍵字提取及廣告文稿評分系統

~ 以化妝品行業應用為例 ~





# 目錄

01

動機、目的、功能介紹及架構

02

熱搜關鍵字萃取

03

廣告文稿評分系統

04

發現問題

# 爆品文案

  
Integrated Planning:  
Always-on +  
Campaigns

  
Online  
Activations

  
SEO:  
Organic  
Search

  
Paid search:  
Adwords

  
Social media  
marketing

  
Email  
marketing -

  
Multi-channel  
Analytics

# 專案動機

## 撰寫廣告文稿

您是否也有以下情況？



- 使用哪些關鍵字？毫無頭緒、關鍵字下得不夠精準...
- 撰文流程繁瑣(小編腦力激盪、上級主管再審核...)
- 文稿品質好壞未知、響應度未知

# 核心功能



## 熱搜關鍵字 萃取排序

自文稿篩選產品熱搜關鍵字，節省撰稿時間與人力



## 文稿評分 深度比對

新建文案與資料庫優質文稿進行相似度比對&深度評分

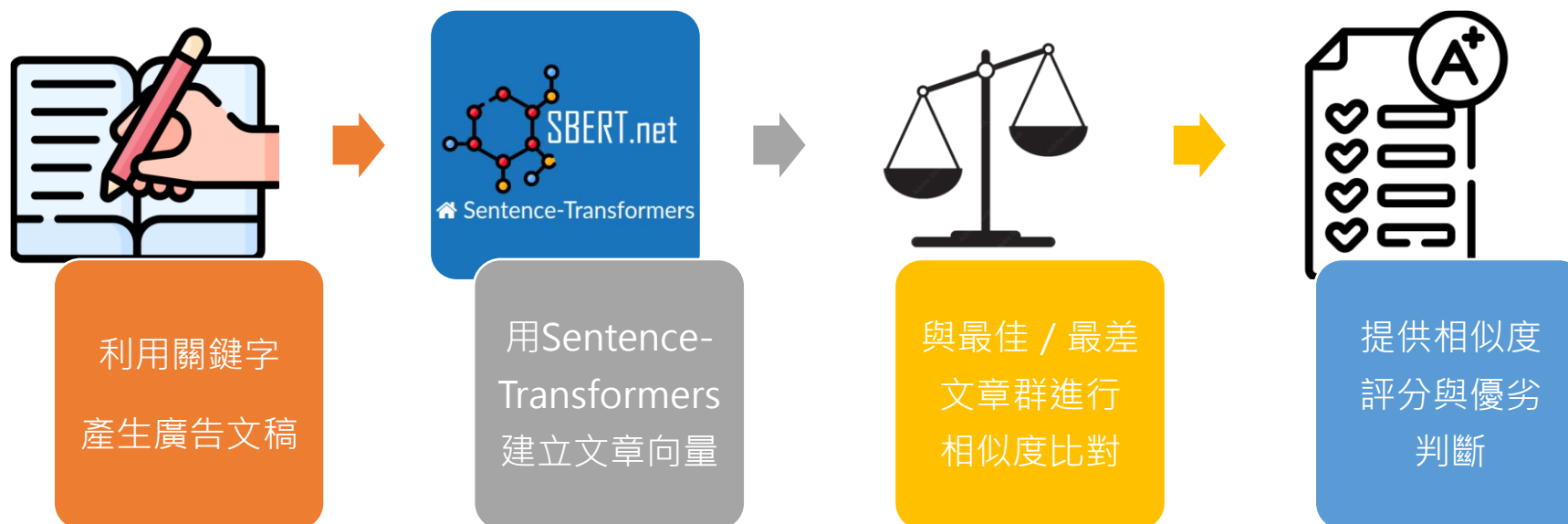
# 流程架構

## 熱搜關鍵字萃取



# 流程架構

## 廣告文稿評分系統





# 熱搜關鍵字萃取

最新最夯關鍵字輕鬆獲取





網路爬蟲  
YouTube熱門  
產品影片

將影片音檔  
轉為文字檔案

文字資料輸入  
斷詞工具演算

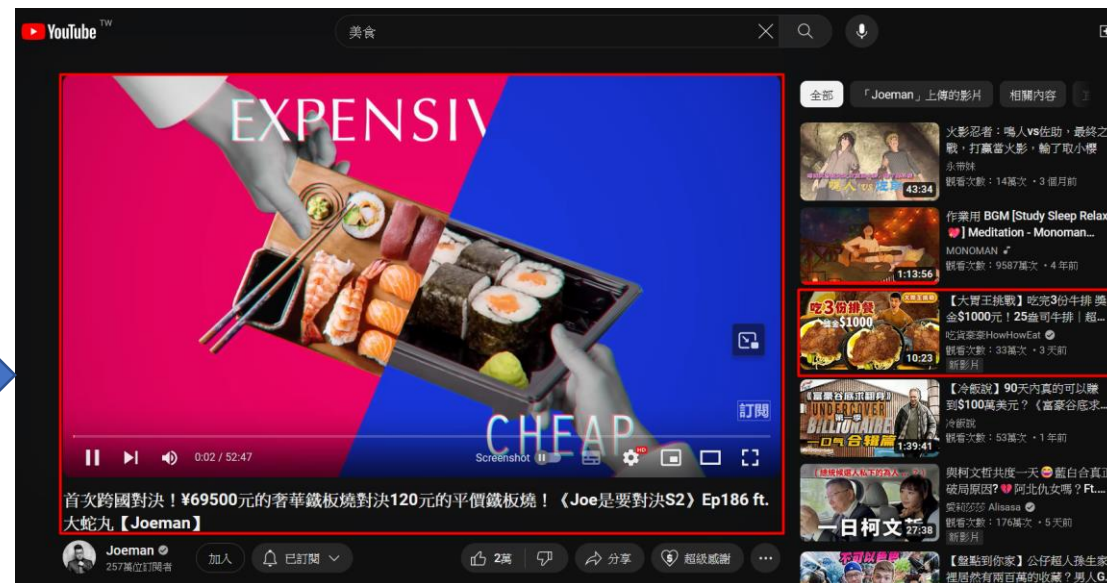
TF篩選詞彙  
剔除停用詞

提供TOP20  
熱搜關鍵字



## Selenium抓取YouTube影片

Selenium設計爬蟲程式，使用firefox抓取YouTube搜尋結果影片的觀看數、讚數、倒讚數以及這些影片的相關影片，以及相關影片的相關影片...以此類推。



網路爬蟲  
YouTube熱門  
產品影片

將影片音檔  
轉為文字檔案

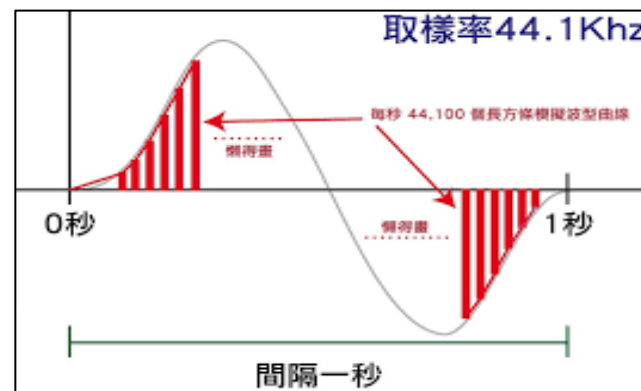
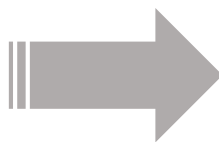
文字資料輸入  
斷詞工具演算

TF篩選詞彙  
剔除停用詞

提供TOP20  
熱搜關鍵字



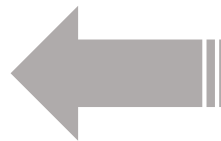
PyTube抓取影片中聲音編碼



AudioSegment  
處理聲道及取樣率



儲存到SQLite



Whisper將聲音轉成文字



網路爬蟲  
YouTube熱門  
產品影片

將影片音檔  
轉為文字檔案

文字資料輸入  
斷詞工具演算

TF篩選詞彙  
剔除停用詞

提供TOP20  
熱搜關鍵字



YouTube轉換文案原稿  
(330份)



使用中研院  
CKIP Tagger斷詞模組

- 辨識繁體中文文稿  
成功率更高

Tool	(WS) prec	(WS) rec	(WS) f1	(POS) acc
結巴系統	90.51%	89.10%	89.80%	--
CkipTagger	97.49%	97.17%	97.33%	94.59%

## 文稿範例

非常簡單的兩個步驟，就可以完成  
像我這樣子立體嘟嘟的唇妝。這種  
畫法真的超級適合天生跟我一樣，  
就是上嘴唇很單薄的朋友。



### 斷詞系統

☒ Show POS tagging

非常(Dfa) 簡單(VH) 的(DE) 兩(Neu) 個(Nf) 步驟(Na)  
就(D) 可以(D) 完成(Vc) 像(P) 我(Nh) 這樣子(VH) 立體  
(VH) 嘟嘟(Nb) 的(DE) 唇妝(Na) 這(Nep) 種(Nf) 畫法  
(Na) 真的(D) 超級(Dfa) 適合(VH) 天生(VH) 跟(P)  
我(Nh) 一樣(VH) 就(D) 是(SH) 上(Ncd) 嘴唇(Na) 很  
(Dfa) 單薄(VH) 的(DE) 朋友(Na)

網路爬蟲  
YouTube熱門  
產品影片

將影片音檔  
轉為文字檔案

文字資料輸入  
斷詞工具演算

手刻TF-IDF找出詞彙的權  
重

提供TOP20  
熱搜關鍵字

### Term Frequency

詞頻 ( term frequency · TF ) 指的是某一個給定的詞語  
在該檔案中出現的頻率。這個數字是對詞數 ( term  
count ) 的標準化

$$tf_{i,j} = \frac{n_{i,j}}{\sum_k n_{k,j}}$$

$$idf_i = \log \frac{|D|}{|\{j : t_i \in d_j\}|}$$



相關影片330份



不相關影片330份

TF-IDF



每一份

將相同詞彙的權重  
加總平均而得

唇膏	唇部	眼影	顏色	嘴唇	底妝	皮膚
0.0078	0.0068	0.0066	0.0065	0.0052	0.0047	0.0045

護唇膏	色號	睫毛	眼妝	遮瑕	唇彩	腮紅
0.0043	0.0042	0.0041	0.0039	0.0037	0.00368	0.00364

蜜粉	餐具	膚色	質地	唇釉	口紅	妝效
0.0035	0.0034	0.0033	0.0032	0.00316	0.00313	0.003

網路爬蟲  
YouTube熱門  
產品影片

將影片音檔  
轉為文字檔案

文字資料輸入  
斷詞工具演算

TF篩選詞彙  
剔除停用詞

提供TOP20  
熱搜關鍵字

## { 關鍵字 }

排序	1	2	3	4	5	6	7	8	9	10
詞彙	唇膏	唇部	眼影	顏色	嘴唇	底妝	皮膚	護唇膏	色號	睫毛
權重	0.0078	0.0068	0.0066	0.0065	0.0052	0.0047	0.0045	0.0043	0.0042	0.0041

■  
■  
■

## { 一般字 }

排序	31	32	33	34	35	36	37	38	39	40
詞彙	現在	示範	乾淨	嘟嘟	比較	一下	畫出來	老氣	滑過	蓋
權重	0.0025	0.0024	0.0023	0.0021	0.0020	0.0019	0.0018	0.0017	0.0015	0.0014



# 廣告文稿評分系統

高下立判你跟爆品文案的距離



利用TOP20  
關鍵字產生  
廣告文稿

用Sentence-  
Transformers  
建立文章向量

與最佳 / 最差文章群  
進行相似度比對

提供相似度  
評分與優劣判斷

關鍵字生成廣告文案

以**粉底液**為例

**Keywords** : {唇膏、唇部、眼影、顏色、嘴唇、底妝}

ChatGPT prompt : XXXXXXXX **{粉底液}** XXXXXXXXXXXXXXX, 包含以下關鍵詞 **{Keywords}**



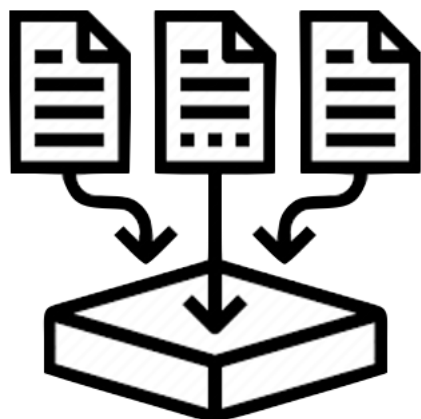
探索妝容的無限可能，從**底妝**開始！我們的粉底液融合獨特配方，打造輕盈自然的**底妝**感。質地細膩，妝容持久，猶如第二層肌膚，展現無瑕妝感。更令人振奮的是，這款粉底液能夠完美搭配**唇膏**，突顯**嘴唇**的立體感，同時與**唇部**、**眼影**等元素融為一體，打造出充滿層次感的彩妝風格。挑戰更多**顏色**，展現屬於你的美麗光彩！

利用TOP20  
關鍵字產生  
廣告文稿

用Sentence-  
Transformers  
建立文章向量

與最佳 / 最差文章群  
進行相似度比對

提供相似度  
評分與優劣判斷



YouTube轉換文案原稿  
(330份)



使用Sentence-Transformers  
換算整篇文章向量



1,548,200,774 views

👍 844,177

👎 7,314,092

依照互動率進行文章優劣排

( 互動率 = (讚數-倒讚) / 觀看數 )



	ID	title	description	link
	過濾	過濾	過濾	過濾
1	1	2023春季開架口...	#2023 #開架口...	https://www.yo...
2	2	黃黑皮👩vs冷白...	《這不是合作影...	https://www.yo...
3	3	簡單兩步驟變立...	《這不是合作影...	https://www.yo...
4	4	🔥淘寶爆紅彩妝...	🌸花西子官方進...	https://www.yo...
5	5	精挑細選！近期...	近期買的口紅試...	https://www.yo...
6	6	超美刷具13支只...	《本片感謝 台糖 ...	https://www.yo...
7	7	令人失望的熱門...	《本影片感謝帕...	https://www.yo...
8	8	雙11推薦&避雷...	《本片感謝NAFU ...	https://www.yo...
9	9	雷到吐的一集口史...	《本片感謝Ana L...	https://www.yo...
10	10	彩妝新手別再亂...	《這不是合作影...	https://www.yo...

產出文章互動率排序列表

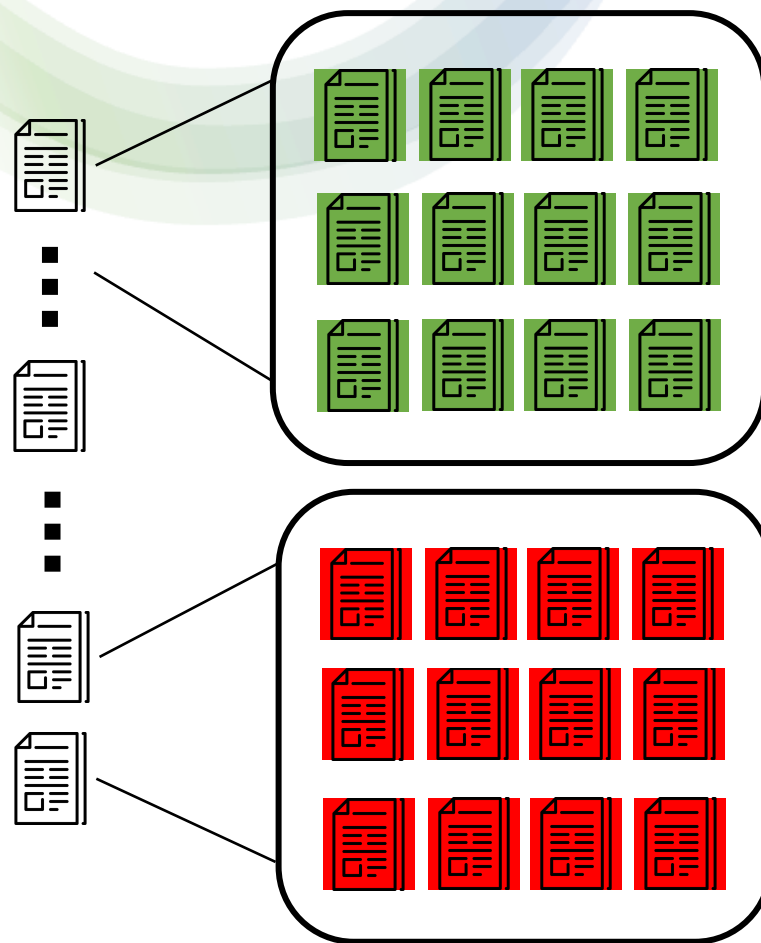


利用TOP20  
關鍵字產生  
廣告文稿

用Sentence-  
Transformers  
建立文章向量

與最佳 / 最差文章群  
進行相似度比對

提供相似度  
評分與優劣判斷



個人文章

與**最佳(前10%)**文章群中  
文章進行相似度一一比對

**最相似10篇**文章

相似度平均分數

**X**

個人文章

與**最差(倒數50%)**文章群中  
文章進行相似度一一比對

相似度平均分數

**Y**

假使  $X > Y$ ，個人文稿分類為**優**；假使  $X < Y$ ，個人文稿分類為**劣**

利用TOP20  
關鍵字產生  
廣告文稿

用Sentence-  
Transformers  
建立文章向量

與最佳 / 最差文章群  
進行相似度比對

提供相似度  
評分與優劣判斷

文章比對

美妝

先上一層的乳吧看起來反光就是這樣然後拿這個畫框這樣子有沒有看出來左邊是after右邊是before這樣有看到嗎下一個遮瑕膏把髒髒的地方都蓋起來R2是射擊鈕看到不好就把它蓋下去右邊也滿滿滿滿滿滿下一個這個粉底露跟精華粉底精華Kai你有沒有看到這邊也補一層舒服舒服拿這個奶油刀這樣子抹在臉上然後再來粉撲撲額頭也給它蓋一層人中不要忘記上髮很棒你看得得到嗎看不到的話我們再更近一點看一下你看有看到吧這個是什麼這個誰彩筆是高中的時候留到現在終於可以派上用場了你有沒有看到我再講一看第三次看到這邊的水美眉你們都有福了你們就照著我這樣子美美美啪啪啪就可以像我一樣美麗了你看有沒有看到然後這個是不一樣的粉紅色是不是比較好看等一下問你男朋友一定不知道這個粉紅色差別在哪裡你們看是不是很漂亮I'm a Barbie girlIn the Barbie world再來是裸光蜜粉欸左邊撲撲右邊也撲撲漂亮再來這個三色修容餅到底要修幾次呢你看是不是before after又不一樣了讚中間的修容到底是差別在哪裡呢右邊刷刷上面也刷刷然後眉毛不要浪費也刷刷再來顏彩盤這個就是拿來畫眼睛的有沒有看到這四個顏色我現在用的是第一個Pink Pinky畫我的漂亮眼睛你是不是看到眼睛裡面有光第二個眼皮也要有光我說眼皮有光就有光下面的眼皮也不要忘記也要有光喔漂亮不知道看哪裡第四種的粉紅色粉紅色Pinky這個是斐濟貝一樣的妝容感讓大家知道我們非常的新鮮Fresh這個臥蠶棒這個不是拿來按摩的這個是拿來畫我們的臥蠶的你看這樣笑起來才會有肌肉這樣子稍微縮起來笑得很開心的感覺我沒有在笑你就感覺我在笑你看我的網美手這個是拿來畫眉毛的你看這樣刷上去是不是好方便這個是橡皮擦嗎這個是Beauty什麼不知道在幹嘛的刷眉毛的你看我的眉毛這樣是不是很有神魔法來囉Before跟After都一樣了不要看我的手看這個是眼線液筆溫柔中我們看這個有多溫柔你要這樣輕輕的畫你看是不是越來越溫柔了呢暴力之氣都不見了下面也要用溫柔中畫一下鼻孔太大了各位觀眾我們的工程現在已經快要接近到尾聲了我們把這個睫毛再給它收尾夾翹一下不要夾錯不要夾到眼皮再用睫毛膏幫我們的睫毛定型你有看過那個蟑螂腳吧就是把它畫成像蟑螂腳那樣子長長的然後我也不知道為什麼要把我的睫毛這樣一束一束的把它夾起來可能真的有比較漂亮吧你說呢Baby你看我的眼睛是不是很漂亮怎麼了你終於看不清楚過來化妝一輩子不小心沾到了然後再拿我這個國小流盛的水彩筆在我的眼睛上方還有我的顴骨上方都刷亮刷亮看起來更有精神還有嘴唇這好滑順喔哪裡都可以擦擦的然後這個嘴巴這樣blend blend的再加強剛剛忘記畫了這邊也再畫一點為什麼大家上幾種唇色這看起來不小心咬到自己嘴巴一樣怎麼突然變成櫻桃色了呢Magic我們這次真的要進入到最後尾聲了你看最後的上光像果凍一樣你以為這樣就完了嗎沒有脖子也不能忘記了你以為這樣就結束了還有帶上我們奇怪的墨鏡漂亮蘋果肌愛心愛心金妮亞帶上墨鏡就完成了喜歡的話請訂閱支持啊還有分享出去我們下次再見啦

比對

產品  
品牌  
唇膏  
色號  
質地  
光澤  
漂亮  
推薦  
嘴唇  
霧面  
包裝  
不錯  
保濕  
彩妝  
訂閱  
膚色  
好用  
皮膚  
好看  
特別  
自然  
系列  
搭配  
上妝  
底妝  
腮紅

相似度: 54.89 % good

# 參考資料

- 影片轉換音檔

<https://stackoverflow.com/questions/67233609/extract-audio-from-youtube-video>

- CKIP Lab

<https://github.com/ckiplab>

- TF-IDF與余弦相似性的應用

<https://www.ruanyifeng.com/blog/2013/03/tf-idf.html>

[https://www.ruanyifeng.com/blog/2013/03/cosine\\_similarity.html](https://www.ruanyifeng.com/blog/2013/03/cosine_similarity.html)

- 中文常見停用詞詞庫

<https://blog.csdn.net/shijiebei2009/article/details/39696571>

- Sentence transformer使用介紹

<https://blog.csdn.net/zwqjoy/article/details/130200952>