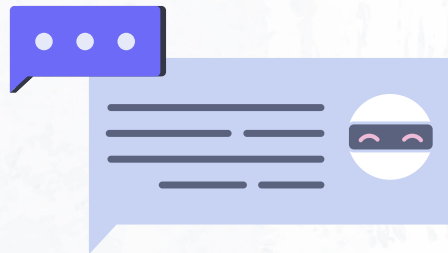
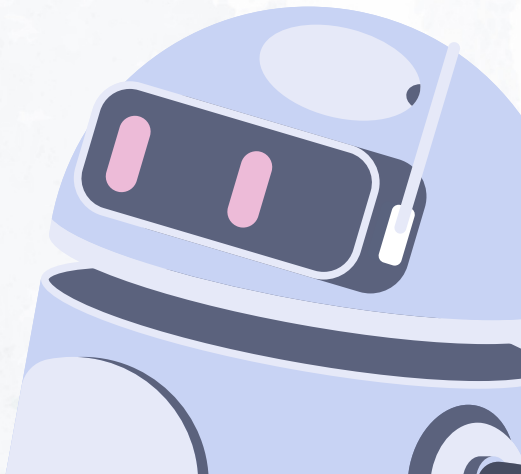


歌詞情緒分析結合Line Notify



資管碩一 611630020 尹嫻雯



目錄

- 01 → 動機及目的
- 02 → 使用的工具
- 03 → 方法及流程說明
- 04 → 程式碼
- 05 → 結果呈現

01 動機及目的



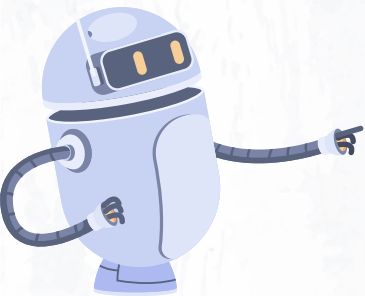
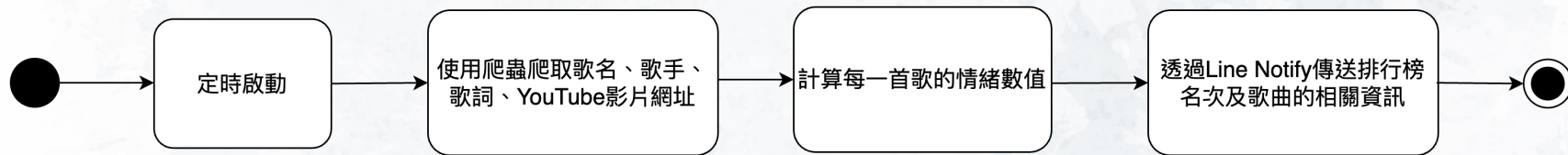
- 我們希望提供一個便捷的方式讓使用者在早上通勤時間聆聽音樂時，能夠了解目前大眾喜愛的音樂排行榜資訊。許多人在通勤時都喜歡聽音樂，但常常只會聽自己熟悉的歌曲。對於沒有使用音樂串流平台的人來說，他們可能需要透過YouTube去搜尋自己喜歡的音樂，這樣就難以了解目前大眾所喜愛的音樂趨勢。
- 因此，我們將提供使用者每天早上即時接收到當天的KKBOX風雲榜中華語新歌日榜前5名的排行榜資訊。使用者可以透過Line Notify在早上起床後，直接收到包括名次、歌名、歌手和情緒分數的通知訊息。這樣的服務能夠讓使用者了解目前大眾喜愛的音樂，並且擁有更廣泛的音樂選擇。
- 另外，透過情緒分析的方式，計算每首歌曲的情緒數值。這有助於使用者根據自己當天的情緒狀態，選擇聽哪一首歌曲。例如，如果使用者感到快樂或充滿活力，他們可以選擇聆聽情緒正向的歌曲；相反地，如果使用者感到憂鬱或沮喪，他們可以選擇聆聽情緒舒緩或安慰的歌曲。這樣的選擇能夠更好地符合使用者的心情需求，提升他們的音樂體驗。
- 此外，也會提供每首歌曲的YouTube影片網址，讓使用者能夠直接透過連結觀看歌曲的音樂影片。這樣一來，使用者可以更深入地感受和體驗歌曲所傳達的情感和意境。

02 使用的工具

- 爬蟲：selenium
- 文字處理：jieba、OpenCC
- 情緒分析：SnowNLP
- 訊息傳送：requests



03 方法及流程說明

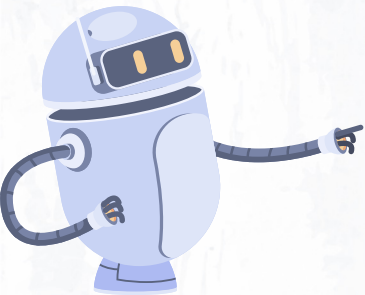


03 方法及流程說明

Step1 :

爬取KKBOX風雲榜中華語新歌

日榜前5名的歌名、歌手



KKBOX 風雲榜

即時榜 日榜 週榜 年度累積榜 音樂趨勢 台灣

華語新歌日榜

類型: 2023-06-14 全部試聽 +

當日 前日 發行日期

排名	日期	歌名	歌手	發行日期
1	1	摯友	周興哲	2023-05
2	2	撲火	A-Lin	2023-05
3	3	沒有開口 (Don't Leave Me)	九澤CP	2023-05
4	8	缺愛·盡 - 现场	盛宇D-SHINE, Faye詹雯婷	2023-06
5	4	大人	梁靜茹	2023-05

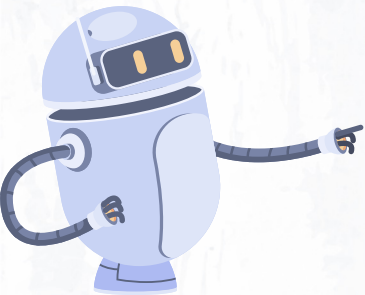
曲風: 華語 西洋 日語 韓語 台語 粵語

排行榜計算說明:
依據 KKBOX 全體會員在特定期間完整
聆聽歌曲的有效點播總次數排序而成,
並由系統自動定期更新。


03 方法及流程說明

Step2 :

透過Webdriver去點擊每一首歌，
進入到歌曲頁面，爬取每首歌的歌詞



KKBOX特色功能訂閱方案娛樂情報會員服務關於我們搜尋下載



歌曲

摯友

Eric 周興哲

試聽

聽全曲

出自專輯

摯友
2023/05/17
專輯曲目

1 摯友

想當星辰 卻像路燈
若愛一個人 切忌愛得太深
酒後傳的訊息 你別當真
我總感情用事 忘了不可能

也委屈你 長期容忍
但我們之間 該用什麼相稱
難道非要我愛 其他的人
你才心安理得 卸下了責任

我們不討論的關係 很接近卻不是愛情
擁有無數交焦 要丟棄太可惜
我演的恨 真不誠懇
你最清楚 我是怎樣的人

03 方法及流程說明

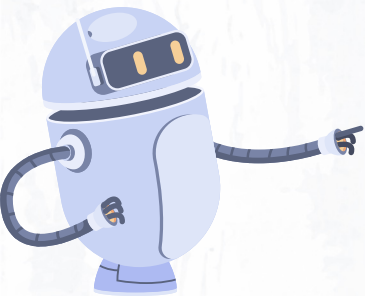
Step3 :

透過

[https://www.youtube.com/results?search_query=](https://www.youtube.com/results?search_query=歌名+歌手)

[歌名+歌手](https://www.youtube.com/results?search_query=歌名+歌手)

前往每一首歌的YouTube搜尋結果頁面，
並抓取第一個結果的影片網址



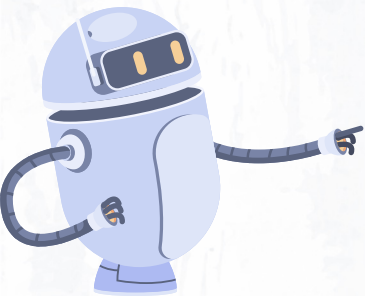
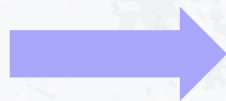
03 方法及流程說明

Step4 :

將爬取的歌詞以jieba進行斷詞(搭配自定義字典及停用詞字典)，再用OpenCC將繁體中文轉為簡體中文，最後使用SnowNLP計算每個詞語的情緒數值，最後將歌曲中全部詞語的情緒值取平均，得到每一首歌的情緒分數



jieba
OpenCC
SnowNLP



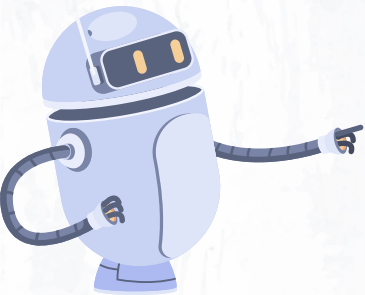
03 方法及流程說明

Step5 :

將爬取的歌名、歌手、YouTube網址及計算
的歌曲情緒數值透過Line Notify傳給使用者。



LINE Notify



04 程式碼

爬蟲

```
1 import requests
2 from selenium import webdriver
3 import time
4 from snowlp import SnowNLP
5 import jieba.analyse
6 from opencv2 import OpenCC
7
8 def web_crawler():
9     #KKBOX華語新歌日榜網址
10    url="https://kma.kkbox.com/charts/daily/newrelease?terr=tw&lang=tc#"
11    count=[2,3,4,5,6]
12    options = webdriver.ChromeOptions()
13    driver=webdriver.Chrome(chrome_options=options,executable_path="chromedriver")
14    driver.get(url)
15    time.sleep(3)
16    content=[]
17    lyrics=[]
18    #抓取前5名的歌曲資訊
19    for i in range(len(count)):
20        time.sleep(2)
21        #抓取歌曲名稱及歌手名稱
22        song=driver.find_element("xpath", '/html/body/div[3]/div/div[2]/ul/li['+str(count[i])+']/a/div/div[1]/span[1]/span[1]').text
23        singer=driver.find_element("xpath", '/html/body/div[3]/div/div[2]/ul/li['+str(count[i])+']/a/div/div[1]/span[1]/span[2]').text
24        #點擊歌曲名稱進入歌曲資訊頁面並抓取歌詞
25        driver.find_element("xpath", '/html/body/div[3]/div/div[2]/ul/li['+str(count[i])+']/a/div/div[1]/span[1]/span[1]').click()
26        lyric=driver.find_element("xpath", '/html/body/main/div[2]/div[1]').text.replace('\n', '')
27        time.sleep(2)
28        #回到排行榜頁面
29        driver.get(url)
30        #到Youtube搜尋該歌曲的YT網址
31        yt_url="https://www.youtube.com/results?search_query="+song+" "+singer
32        driver1=webdriver.Chrome(chrome_options=options,executable_path="chromedriver")
33        driver1.get(yt_url)
34        time.sleep(2)
35        yt=driver1.find_element("xpath", '//*[id="video-title"]').get_attribute("href")
36        #將歌曲名稱、歌手名稱、YT網址及歌詞存入content中
37        content.append([song,singer,yt])
38        lyrics.append(lyric)
39    print(content)
40    driver.close()
41    driver1.close()
42    return content,lyrics
```

04 程式碼

情緒分析

```
44 def sentiment_analysis(lyrics):
45     cc = OpenCC('t2s')
46     jieba.set_dictionary('dict.txt.big')
47     jieba.load_userdict('userdict.txt')
48     with open('stops.txt', 'r', encoding='utf8') as f:
49         stops = f.read().split('\n')
50     avg=[]
51     for i in range(len(lyrics)):
52         text=lyrics[i]
53         total=0
54         count=0
55         for sent in text.split("\n"):
56             sent_words=[t for t in jieba.cut(sent) if t not in stops and t.split() and len(t)>1] #斷詞
57             for i in sent_words:
58                 new_i=cc.convert(i) #繁體轉簡體
59                 s=SnowNLP(new_i)
60                 total+=s.sentiments #取得每一個詞語的情緒數值
61                 count+=1
62             average=total/count #計算每一首歌的平均情緒數值
63             avg.append(average)
64     return avg
```


04 程式碼

Line Notify傳遞訊息

(未來會設定每日固定時間爬取資訊及進行分析，提供使用者當日的排行榜資訊)

```
66 # LINE Notify 權杖
67 token = 'nVrffNorEJJVJgkXHiP8UGQn3YRf1BFTLlyEhjrS6l6'
68
69 content, lyrics=web_crawler()
70 avg=sentiment_analysis(lyrics)
71 for i in range(0, len(content)):
72     content[i].append(avg[i])
73
74 # 要發送的訊息
75 message = '\n今日華語新歌日榜前5名:\n\n'
76 for i in range(len(content)):
77     message1='第'+str(i+1)+'名:'+str(content[i][0])+'\n歌手:'+str(content[i][1])+'\nYT網址: '+
78     str(content[i][2])+'\n情緒數值:'+str(content[i][3])+'\n\n'
79     message+=message1
80 message+="註>情緒數值愈接近1，歌曲愈正面；數值愈接近0，歌曲愈負面"
81 print(message)
82
83 # HTTP 標頭參數與資料
84 headers = { "Authorization": "Bearer " + token }
85 data = { 'message': message }
86
87 # 以 requests 發送 POST 請求
88 requests.post("https://notify-api.line.me/api/notify",
89               headers = headers, data = data)
```


05 結果呈現



LINE Notify





【歌詞情緒分析小幫手】
今日華語新歌日榜前5名:

第1名:摯友
歌手:周興哲
YT網址:<https://www.youtube.com/watch?v=kYOUbDO6hXg&pp=ygUQ5pGv5Y-LIOWRqOiliOWTsg%3D%3D>
情緒數值:0.5897173080975732

第2名:撲火
歌手:A-Lin
YT網址:<https://www.youtube.com/watch?v=IhGbYGT6QJl&pp=ygUM5pKy54GrIEEtTGlu>
情緒數值:0.5944473607525866

第3名:沒有開口 (Don't Leave Me)
歌手:九澤CP
YT網址:<https://www.youtube.com/watch?v=jLrpW6SumdM&pp=ygUm5rKS5pYj6ZaL5Y-jlChEb24ndCBMZWF2ZSBNZSk5Lmd5r6kQ1A%3D>
情緒數值:0.5670246964033705

第4名:訣愛·盡 - 现场
歌手:盛宇D-SHINE, Faye詹雯婷
YT網址:<https://www.youtube.com/watch?v=y-eBNrCciTA&pp=ygUx6Kij5oSbwrfnm6EgLSDnrDlnLog55ub5a6HRC1TSEIORSwgRmF5Zeipuembr-Wptw%3D%3D>
情緒數值:0.621810096348143

第5名:大人
歌手:梁靜茹
YT網址:https://www.youtube.com/watch?v=Lh_pbcAxxhU&pp=ygUQ5aSn5Lq6iOaigemdnOiMuQ%3D%3D
情緒數值:0.6102842337678454

<註>情緒數值愈接近1，歌曲愈正面；
數值愈接近0，歌曲愈負面

下午 5:07

輸入訊息



The End