

國立高雄科技大學

資訊工程系

網路資料庫程式設計期末報告

新聞情感分析家

News Sentiment Analyser -

Sentiment Analysis of News Articles using Python

授課教授: 陳俊豪教授

學生: C110151107 吳季珊

C110152355 陳允恩

C110147202 許芷菱

2024年1月14日

目 錄

第一章 簡介與動機

第二章 文獻回顧

第三章 方法論

第四章 資料庫ERD圖

第五章 系統功能與實施

第六章 實作結果與流程

第七章 總結

第八章 參考文獻

1. 簡介與研究動機

在現今社會中，每個人、公司或團體都希望擁有良好形象。要塑造良好形象，除了用家、消費者真實的評價外，亦可聘僱公關撰寫及發佈正面新聞。但是，負面的新聞卻往往無法被完全控制。在這種情況下，我們意識到提供多方面的新聞和多元觀點的重要性，這樣公眾才能更全面地了解他們想要了解的對象。

這個專題的目標是通過分析新聞文章中的情感表達，提供公眾對特定對象的不同觀點。透過情感分析，我們可以獲得對特定個人或團體的客觀觀點，從而平衡由公關所控制的新聞或其他受控形式所帶來的偏見。

這種多元觀點的提供對於公眾來說非常重要，因為他們可以從不同的角度來了解一個人或團體。這樣一來，公眾就能夠更全面地評估和判斷一個人的行為、言論和成就，而不僅僅依賴於單一來源的信息。

這個專題還有助於提供更真實、更客觀的信息，使公眾能夠更好地理解特定對象的背景、價值觀和行為。透過這種深入了解，公眾可以更明智地做出決策，並與特定對象進行更有建設性的互動。

2. 文獻回顧

2.1 情感分析

情感分析是一種自然語言處理技術，用於自動識別和提取文本中的情感傾向。情感分析在商業、政治和社會等領域中都有廣泛的應用。

情感分析也被稱為意見探勘，旨在分析人們對特定實體，如產品、服務、組織、事件等的主觀觀點，包括意見、情感和評價等。透過將這些資訊根據評論的對象（即產品特徵）和表達的主觀意見（即意見傾向）適當地摘要和彙整，可以轉化為結構化的顧客知識（Liu, 2010, 2012）。這有助於企業更好地了解廣大消費者的意見，並開發重要的商業智慧應用，或為一般使用者提供另一個訊息管道，以避免被廠商官方資訊誤導，從而做出最佳的消費決策（楊錦生，2017）。

2.2 新聞情感分析

新聞情感分析是將情感分析技術應用於新聞文章的分析中。新聞情感分析可以幫助人們快速了解新聞事件的情感傾向，並且分析不同媒體對同一事件的報導差異。

新聞是社會成員的一個信息來源，與社會的經濟結果密切相關。面對金融危機期間的報導，斯蒂格利茨（2011年，第23頁）強調：“信息使讀者（無論是作為消費者、經理人、工人、投資者、房主還是選民）能夠做出更好的決策。更好的個人決策將會帶來更好的社會結果。”媒體系統的結果，即其新聞，影響著一個社會、其公民以及企業和政府行為者。除了對經濟的直接經驗，新聞作為一個重要的經濟信息來源，提醒公民關注經濟問題。

Stiglitz（2011年）指出，新聞是社會成員的一個信息來源，與社會的經濟結果密切相關。他強調：「信息使讀者（無論是作為消費者、經理人、工人、投資者、房主還是選民）能夠做出更好的決策。更好的個人決策將會帶來更好的社會結果。」

（News is one source of information for the members of a society and dynamically connected to a society's economic outcome.）

媒體系統的結果，亦即是新聞，影響著一個社會、其公民以及企業和政府行為者。新聞是作為一個重要的信息來源，尤其在經濟方面不可或缺。

3. 方法論

3.1 數據收集

由於媒體可能受投資者、股東、總裁或相關持份者意見所影響立場，所以為了提供公眾更全面的資訊，不會抓取單一媒體的新聞。從Google News 抓取可以獲得不同媒體公司的報導，可大大減少偏頗的結果。故此，本專題將主要從Google News抓取及收集新聞文章，並且篩選出符合主題的文章進行分析。

3.2 情感分析技術

本專題將運用自然語言處理技術中的Albert情感分析模型，是基於bert模型的輕量化簡便的版本，對新聞標題的情感傾向進行分析。為了增加易讀性，在程式中刪去了中立的立場，並將分析結果分為正面和負面兩個方面，這樣可以更清晰地呈現新聞標題所表達的情感內容。

3.3 資料儲存

我們將新聞、分析結果儲存到資料庫裡，方便用於管理和檢索資料提供結構化和有組織的存儲，支援資料的有效管理和快速檢索。以利後續的呈現與使用。

3.4 資料畫面呈現

利用python爬蟲的方式將輸入關鍵字所抓取新聞，擷取新聞中的特定詞彙與情緒分析模型的輸出結果等儲存到資料庫裡，利用 Word Cloud生成文字雲，網頁的部分採用flask 與html，最終呈現在網頁上的資料有新聞標題、URL、情緒結果、文字雲。Tkinter的版本則採用TreeView顯示搜尋結果，包括新聞編號(article_id)、新聞標題、搜索關鍵字、發佈之媒體及情緒分數等等，表格首欄更可提供排序功能以方便閱覽。

3.5 系統架構

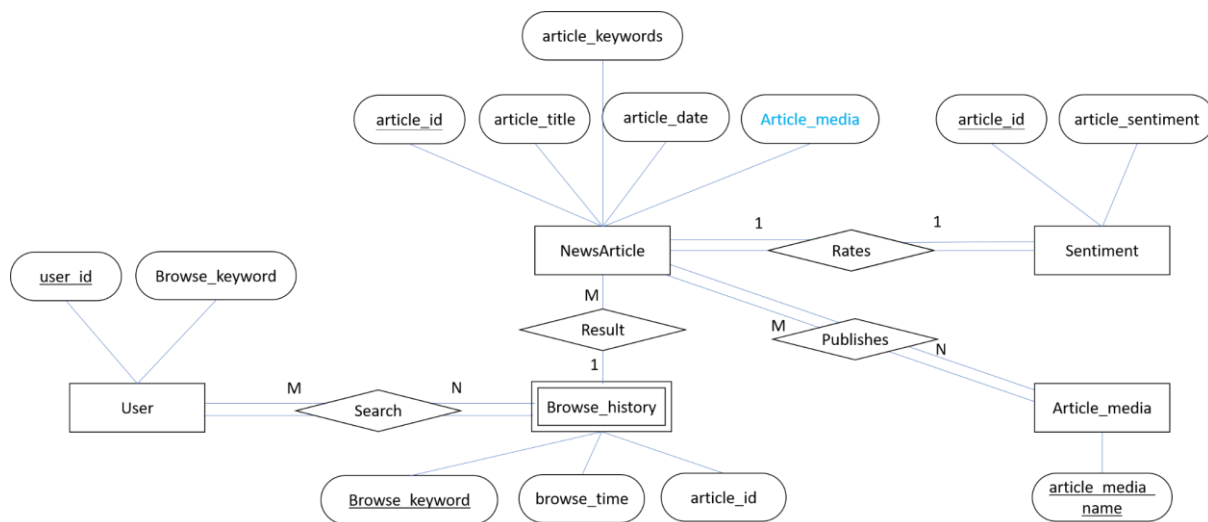
本專題的目的是開發一個新聞情感分析系統，讓使用者能夠快速了解新聞內容的情感傾向。透過這個系統，使用者可以輸入新聞標題或關鍵字，系統將抓取資料並自動分析其中的情感內容，完成分析後將其歸類為正面或負面情感。這樣的分析結果可以幫助使用者更快速地瞭解新聞的情感，從而更好地理解搜尋對象的正負面新聞占比與形象。

而且此專題從用家角度出發，設計了兩種介面，網頁版、Tkinter，供更多的選擇給使用者。

此外，系統還將提供文字雲生成功能，將關鍵字組成一個視覺化的文字雲。文字雲能夠直觀地展示新聞中出現頻率較高的關鍵字，讓使用者一目了然地掌握新聞的主題和重點。這將有助於使用者更全面地理解新聞內容，並從中獲取有價值的資訊。

4. 資料庫ERD圖

在本專題中，將建立一個資料庫來存儲新聞文章和相關的情感分析結果。以下是資料庫的實體關係圖 (ERD)：



5. 系統功能與實施

5.1 系統功能

這個新聞情感分析專案的系統功能包括：

文章收集：系統能夠自動從不同的新聞來源收集文章，包括線上新聞網站、獨立媒體等。收集到的文章將用於後續的情感分析及產生文字雲。

情感分析：系統將對收集到的文章進行情感分析，分析文章中的情感表達，包括正面、負面情感。這樣可以獲得對特定對象的公眾情感的客觀觀點。

資料儲存：系統將收集的新聞與情緒分析結果存進資料庫，用於快速的搜索與管理，以便後續呈現與使用。

觀點呈現：系統將根據情感分析的結果，生成文字雲，讓使用者能夠使用介面更直觀的了解新聞的正負面評價占比與新聞關鍵字。

5.2 系統實施

本專題將使用Python語言及html進行系統的開發，並使用相關的自然語言處理庫和情感分析模型來實現情感分析功能。系統將建立一個使用者友好的界面，讓使用者可以輸入新聞文章並獲取情感分析結果。

在本專題中，我們使用Python語言進行系統的開發，並使用相關的自然語言處理庫和情感分析模型來實現情感分析功能。我們的系統將建立兩個使用者友好的界面，讓使用者可以輸入文本內容並獲取情感分析結果。

以下是本專題中使用到的模組：

- TextBlob:

描述：TextBlob是一個Python庫，用於進行自然語言處理（NLP）任務。它提供了簡單的API，使得對文本進行情感分析、詞形還原、名詞詞組提取等任務變得容易。

在程式碼中的使用：用於情感分析。

- VADER Sentiment Analysis:

描述：VADER（Valence Aware Dictionary and sEntiment Reasoner）是一種情感分析工具，專門設計用於分析社交媒體文本的情感。

在程式碼中的使用：用於獲取文本的情感傾向。

- GoogleNews:

描述：GoogleNews是一個用於搜尋和擷取新聞的Python庫。它允許設定搜尋條件，然後從Google News檢索相關的新聞。

在程式碼中的使用：從GoogleNews抓取及收集新聞文章。時間範圍為一年，用time模組取得使用當天，並設定抓取範圍設定為1-3頁。

● BeautifulSoup:

描述：BeautifulSoup是一個用於解析HTML和XML文檔的庫，提供了簡單的方式來提取所需的信息。

在程式碼中的使用：用於解析網頁內容，提取新聞文本和連結。

● Flask:

描述：Flask是個簡易的網路架構，吸收其他框架的優點而建立。Flask是一個以Jinja2範本引擎與WerkzeugWSGI為基礎而建立的網站微框架。

在程式碼中的使用：用於構建Web應用程序，定義路由和呈現模板，將情緒分析結果以網頁呈現。

● Transformers (Hugging Face):

描述：Transformers是一個由Hugging Face提供的庫，用於自然語言處理任務，包括預訓練模型的使用。

在程式碼中的使用：用於載入和使用預訓練的ALBERT模型進行情感分類。

● WordCloud:

描述：WordCloud是一個用於生成文字雲的庫，將文本中的詞彙按照出現頻率可視化呈現。

在程式碼中的使用：用於生成文字雲，讓文章主題以視覺化的方式呈現。

● Matplotlib和NumPy:

描述：Matplotlib是一個用於繪製數據視覺化的庫，而NumPy則是用於數值計算的庫。

在程式碼中的使用：用於繪製文字雲的圖形。

● Tkinter:

描述：Tkinter是用於創建視窗應用程序和圖形用戶界面，提供直觀的操作界面。

在程式碼中的使用：用於搜尋及呈現搜尋結果。

6. 實作結果與流程

6.1 網頁版

為了方便讓一般使用者也可以使用我們的工具，我們以網頁的形式，做了一個比較容易操作版本，只要輸入想要搜尋的關鍵字，便可直接顯示分析的結果，顯示的部分我們除了顯示每則新聞的正負面情緒比例，還提供了每則新聞的連結，讓有興趣的人可以直接從連結觀看詳細文章，也為了更了解文章主要內容，在顯示網頁時，也加進了文字雲，讓使用者可以以圖片的形式更直接的看到文章的主題內容。

網頁部分，我們使用了flask的架構，Flask 是一個輕量型的 Python Web 應用程式架構，可提供 URL 路由和頁面轉譯的基本要素。而文字雲的部分，我們直接使用python的Jieba + Word cloud套件，將我們爬到的文章標題在過濾掉其中的連接詞後，以文字以圖片的方式顯示。

最終呈現的網頁內容為新聞標題、URL、情緒分析結果、文字雲圖片。

以下為流程介紹(以圖片呈現)：

```
● PS D:\桌面\禮拜五> & d:/桌面/禮拜五/myenv/Scripts/Activate.ps1
○ (myenv) PS D:\桌面\禮拜五> & d:/桌面/禮拜五/myenv/Scripts/python.exe d:/桌面/禮拜五/myenv/EmotionalTest/apptest.py
* Serving Flask app 'apptest'
* Debug mode: on
WARNING: This is a development server. Do not use it in a production deployment. Use a production WSGI server instead.
* Running on http://127.0.0.1:5000
Press CTRL+C to quit
* Restarting with stat
* Debugger is active!
* Debugger PIN: 735-779-151
```

啟動Flask應用程式接受來自客戶端的HTTP請求與做出回應(執行apptest.py檔)



- 1、確定flask在執行後，開啟網頁輸入http://127.0.0.1:5000/search進入到搜索頁面
- 2、輸入想搜索的關鍵字
- 3、搜尋的過程順便將分析結果與新聞存到資料庫

網頁圖1(為方便觀看網頁分成三張截圖呈現)

搜尋結果

總共搜尋到：30筆新聞

你搜尋的關鍵字是: 奶茶

搜尋結果

關於奶茶 — 正面的新聞：21筆，負面的新聞：9筆

像上圖關鍵字為"奶茶"，就會出現關於奶茶的正負面新聞佔比

網頁圖2

(從資料庫提取關鍵字的新聞)



透過搜索關鍵字"奶茶"，出現的新聞詞彙以出現次數生成文字雲

網頁圖3

情感分析結果:

Sentiment Scores: {'Negative': 0.86, 'Positive': 0.14}

1. 飛爺帶奶茶走紅毯

Link:
<https://portal.sina.com.hk/entertainment/weibosinaent/2024/01/14/707753/%25E9%25A3%259B%25E7%2588%25BA%25E5%25B8%25B6%25E5%25A5%25B6%25E8%258C%25B6%25E8%25B5%25B0%25E7%25B4%2585%25E6%25AF%25AF/8&ved=2ahUKEwiD2f6Xg92DAXUGsFYBHfd0D3oQxfQBegQIARAC&usg=AOvVaw363XSZChcMF26dWih8wbqmq>

情感分析結果:

Sentiment Scores: {'Negative': 0.04, 'Positive': 0.96}

2. 【食用】成都楊桃冰-鹹楊桃飲、三十三茶王奶茶、鳳梨釋迦捲、草莓起司軟歐、草莓奶油軟法-7-ELEVEN | 艾克星貨舖 - 方格子

Link:
https://vocus.cc/article/65a0ecbcfd8978000119b687&ved=2ahUKEwiD2f6Xg92DAXUGsFYBHfd0D3oQxfQBegQICRAC&usg=AOvVaw136bxn_nOCWvmH89Tha7tA

文字的呈現為顯示新聞標題與新聞連結、情緒分析結果({'Negative': , 'Positive':})

問題輸出 偵錯主控台 終端機 連接埠

```
DEBUG:jieba:Building prefix dict from D:\桌面\禮拜五\jieba\dict.txt ...
Loading model from cache C:\Users\user\AppData\Local\Temp\jieba.u69819ec6f08b2a9bf220f27a4e2bb146.cache
DEBUG:jieba:Loading model from cache C:\Users\user\AppData\Local\Temp\jieba.u69819ec6f08b2a9bf220f27a4e2bb146.cache
Loading model cost 0.498 seconds.
DEBUG:jieba:Loading model cost 0.498 seconds.
Prefix dict has been built successfully.
DEBUG:jieba:Prefix dict has been built successfully.
127.0.0.1 - - [14/Jan/2024 21:52:11] "POST /search_result HTTP/1.1" 200 -
INFO:werkzeug:127.0.0.1 - - [14/Jan/2024 21:52:11] "POST /search_result HTTP/1.1" 200 -
127.0.0.1 - - [14/Jan/2024 21:52:11] "GET /static/styles/styles.css HTTP/1.1" 304 -
INFO:werkzeug:127.0.0.1 - - [14/Jan/2024 21:52:11] "GET /static/styles/styles.css HTTP/1.1" 304 -
```

程式後端可顯示使用者的查詢結果與使用狀況

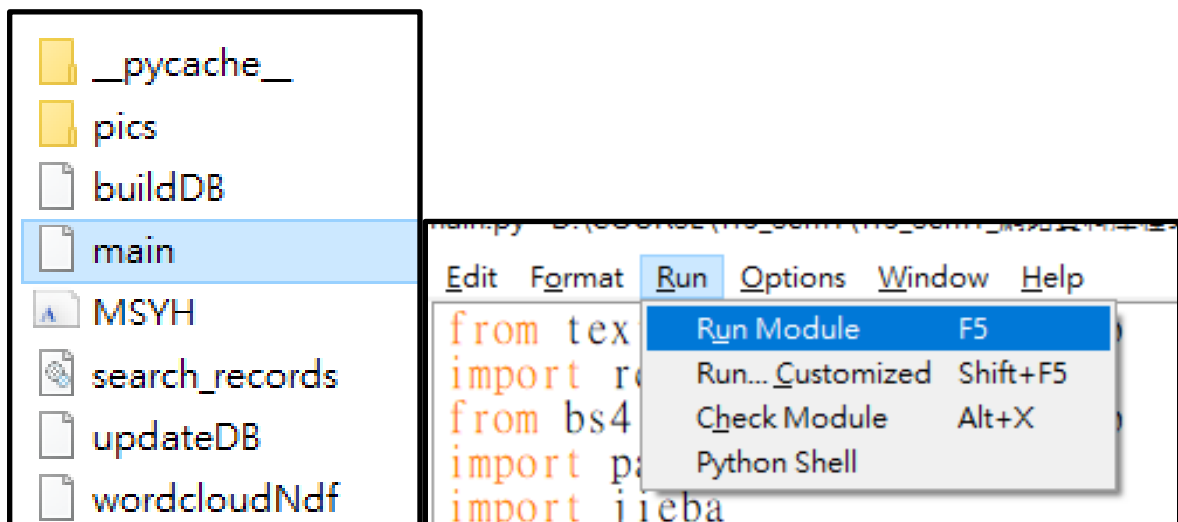
6.2 個人版

至於個人版(Tkinter)，亦可稱作單機版。因設計對象為有程式底子的人士，所以會在個人電腦上執行。用家需要先自行下載Python及程式的zip檔，並且安裝套件。

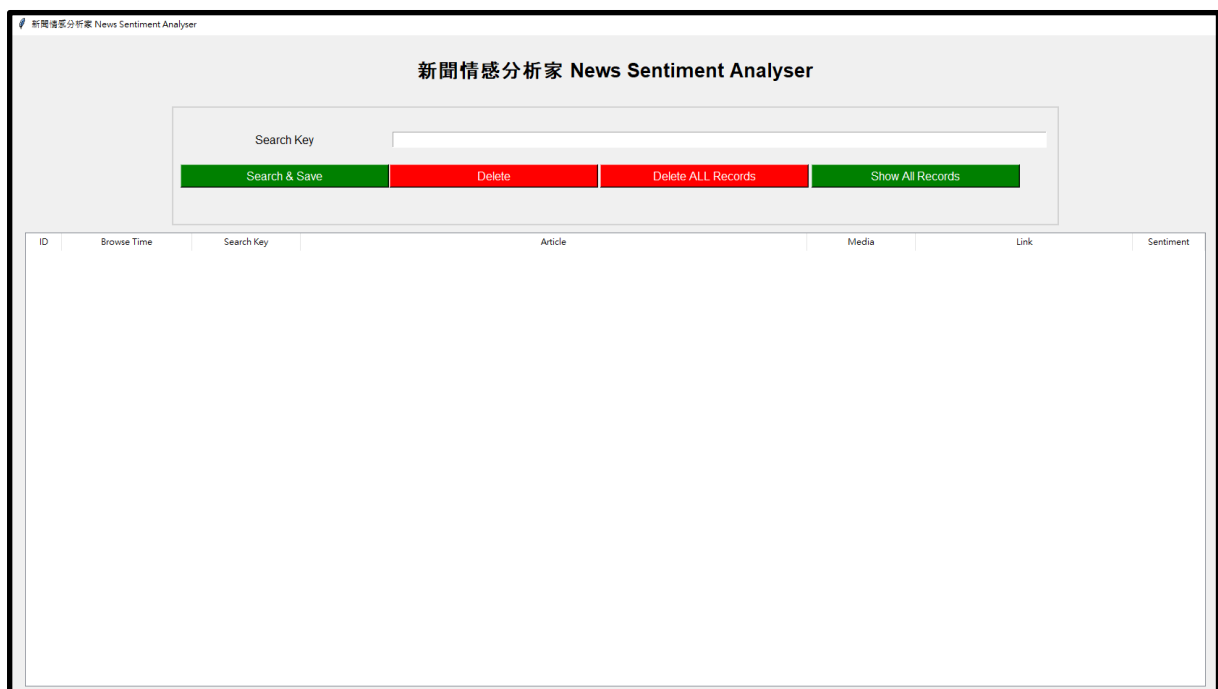
Package	Command
BeautifulSoup	pip install beautifulsoup4
Django	pip install django
Flask	pip install flask
Google Cloud Language	pip install google-cloud-language
Jieba	pip install jieba
Matplotlib	pip install matplotlib
NumPy	pip install numpy
Pandas	pip install pandas
PyTorch	pip install torch
PyGoogleNews	pip install pygooglenews
Pillow	pip install pillow
SQLAlchemy	pip install sqlalchemy
Scikit-learn	pip install scikit-learn
Tkinter	pip install tk
TensorFlow	pip install tensorflow
TextBlob	pip install textblob
Transformers	pip install transformers
Requests	pip install requests

Package	Command
VaderSentiment	pip install vaderSentiment
Wordcloud	pip install wordcloud

下載並完成安裝以上內容後，開啟zip檔裡的主main.py並執行。



執行後會出現此畫面，請用家輸入關鍵字後按“ Search & Save”。



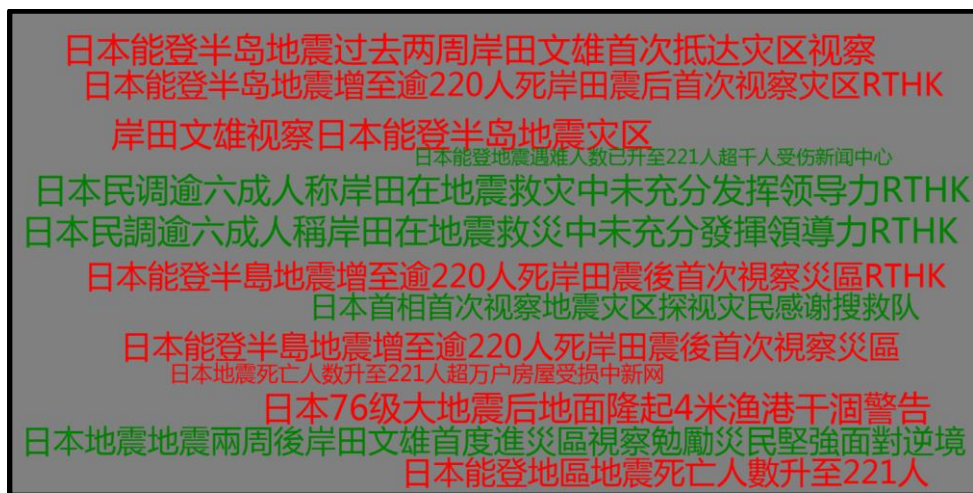
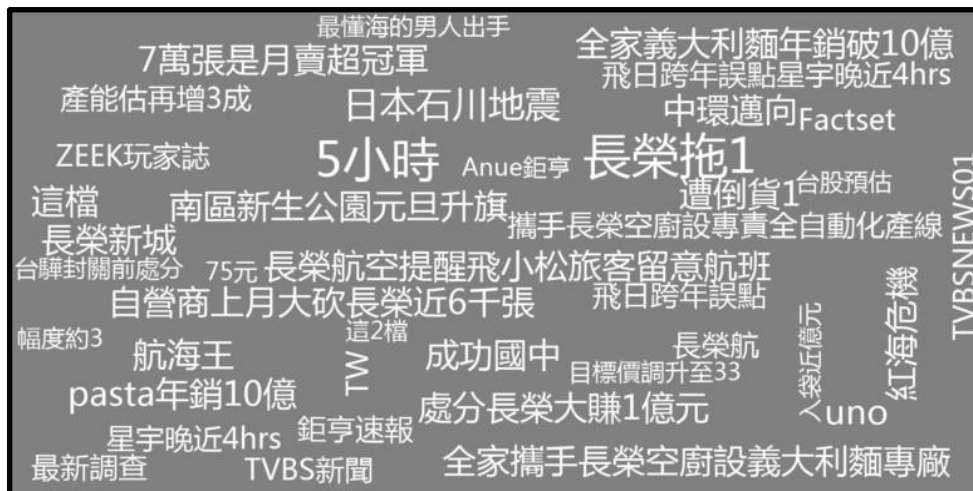
程式search(query)會自行從GoogleNews 抓取相關新聞，並篩除重複的新聞。

同時，程式會把新聞的資料分類及進行情緒分析，如新聞編號(article_id)、新聞標題、搜索關鍵字、發佈之媒體及情緒，得出的結果會暫時存進dataframe。

```
df = pd.DataFrame({
    'News Title': uni_news_titles,
    'Sentiment Scores': uni_sentiment_scores,
    'Link': uni_links,
    'Media': uni_media_name,
    'Color': uni_colors})
```

```
if result[n] not in unique_results:
    unique_results.add(result[n])
    uni_news_titles.append(result[n])
    media_name = urlparse(links[n]).hostname
    uni_media_name.append(media_name)
    uni_links.append(links[n])
    # Extract the value of the '+/-' key
    negative_score = sentiment_scores['Negative']
    positive_score = sentiment_scores['Positive']
    if (negative_score > positive_score):
        uni_colors.append("red")
        uni_sentiment_scores.append(negative_score*-1)
    else:
        uni_colors.append("green")
        uni_sentiment_scores.append(positive_score)
```

完成query(search)時，會呼叫create_cloud(df, unique_results)產生文字雲，並以plt.imshow()顯示。為免中文標題內的標點符號會影響文字雲切字，造成標題沒上色的情況，要求提取dataframe內的標題並去除標點符號。下圖白色字為文字雲套件自行切字，有紅、綠色字為處理過的文字，可清楚大概分辨整體偏正面或負面。

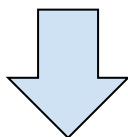


此文字雲會存在使用者的“Pics”檔案內，供以後參閱。

```
# Save the word cloud as a PNG image
output_path = 'pics/wordcloud3.png' # Replace with the des
plt.savefig(output_path, bbox_inches='tight', pad_inches=0)
```

當關閉plt視窗後，詳細結果會以Treeview形式表達。使用者可以按表頭進行排序，下圖以情緒分數示範。

ID	Browse Time	Search Key	Article	Media	Link	Sentiment
1	2024-01-14 13:37:58	日本地震	日本民調逾六成人稱岸田在地震救災中未充分發揮領導力RTHK	news.rthk.hk	https://news.rthk.hk/rthk/ch/component/k2/1	0.74
2	2024-01-14 13:37:58	日本地震	日本民調逾六成人稱岸田在地震救災中未充分發揮領導力RTHK	gbcode.rthk.hk	https://gbcode.rthk.hk/TuniS/news.rthk.hk/r	0.76
3	2024-01-14 13:37:58	日本地震	日本能登半島地震過去兩周岸田文雄首次抵達災區視察	www.8world.com	https://www.8world.com/videos/news-bite/j	-0.77
4	2024-01-14 13:37:58	日本地震	日本76級大地震後地面隆起4米海邊干涸警告	www.cnbeta.com.tw	https://www.cnbeta.com.tw/articles/tech/14	-0.9
5	2024-01-14 13:37:58	日本地震	岸田文雄視察日本能登半島地震災區	www.zaobao.com.sg	https://www.zaobao.com.sg/realtime/world	-0.59
6	2024-01-14 13:37:58	日本地震	日本地震地震兩周後岸田文雄首度進災區視察勉勵災民堅強面對逆境	www.am730.com.hk	https://www.am730.com.hk/%25E5%259C	0.65
7	2024-01-14 13:37:58	日本地震	日本能登半島地震增至逾220人死岸田震後首次視察災區	www.bastillepost.com	https://www.bastillepost.com/hongkong/arti	-0.82
8	2024-01-14 13:37:58	日本地震	日本能登半島地震增至逾220人死岸田震後首次視察災區RTHK	gbcode.rthk.hk	https://gbcode.rthk.hk/TuniS/news.rthk.hk/r	-0.91
9	2024-01-14 13:37:58	日本地震	日本能登半島地震增至逾220人死岸田震後首次視察災區RTHK	news.rthk.hk	https://news.rthk.hk/rthk/ch/component/k2/1	-0.82
10	2024-01-14 13:37:58	日本地震	日本能登半島地震增至逾220人死岸田震後首次視察災區RTHK	www.tkwk.hk	https://www.tkwk.hk/a/202401/14/AP65a3	-0.79
11	2024-01-14 13:37:58	日本地震	日本首相首次視察地震災區探視災民感謝救援	www.enanyang.my	https://www.enanyang.my/%25E5%259B%	0.63



ID	Browse Time	Search Key	Article	Media	Link	Sentiment
14	2024-01-14 13:37:58	日本地震	日本能登地震遇難人數已升至221人超千人受傷國際時訊	www.dzwww.com	http://www.dzwww.com/xinwen/guojixinwen	-0.51
29	2024-01-14 13:37:59	日本地震	日本石川縣能登地區發生4.4級地震中新網	www.chinanews.com	https://www.chinanews.com.cn/gj/2024/01-	-0.54
5	2024-01-14 13:37:58	日本地震	岸田文雄視察日本能登半島地震災區	www.zaobao.com.sg	https://www.zaobao.com.sg/realtime/world	-0.59
12	2024-01-14 13:37:58	日本地震	日本地震死亡人數升至221人超千戶房屋受損中新網	www.chinanews.com	https://www.chinanews.com.cn/gj/2024/01-	-0.62
23	2024-01-14 13:37:59	日本地震	日本石川地震已兩周岸田首入災區探視避難民慰問救援人員	tw.news.yahoo.com	https://tw.news.yahoo.com/%25E6%2597%	-0.63
28	2024-01-14 13:37:59	日本地震	日本石川縣能登地區發生4.4級地震香港文匯網	www.wenweipo.com	https://www.wenweipo.com/a/202401/13/A	-0.64
17	2024-01-14 13:37:59	日本地震	岸田赴石川視察能登半島地震災區	china.kyodonews.net	https://china.kyodonews.net/news/2024/01-	-0.67
18	2024-01-14 13:37:59	日本地震	日本地震災區遭遇強降雪新聞央視網cctvcom	photo.cctv.com	https://photo.cctv.com/2024/01/14/PH0A01	-0.71
19	2024-01-14 13:37:59	日本地震	日本能登地區地震死亡人數升至221人中新網	www.chinanews.com	https://www.chinanews.com.cn/gj/2024/01-	-0.71
21	2024-01-14 13:37:59	日本地震	地震致日本鹿兒島火山噴發兩公里內可能天降落石最新消息	www.cnbeta.com.tw	https://www.cnbeta.com.tw/articles/tech/14	-0.74
24	2024-01-14 13:37:59	日本地震	日本再發地震一火山噴發兩公里內可能天降落石最新消息	www.nbd.com.cn	https://www.nbd.com.cn/articles/2024-01-	-0.76
3	2024-01-14 13:37:58	日本地震	日本能登半島地震過去兩周岸田文雄首次抵達災區視察	www.8world.com	https://www.8world.com/videos/news-bite/j	-0.77
10	2024-01-14 13:37:58	日本地震	日本能登半島地震死亡人數升至221人	www.tkwk.hk	https://www.tkwk.hk/a/202401/14/AP65a3	-0.79
7	2024-01-14 13:37:58	日本地震	日本能登半島地震增至逾220人死岸田震後首次視察災區	www.bastillepost.com	https://www.bastillepost.com/hongkong/arti	-0.82
9	2024-01-14 13:37:58	日本地震	日本能登半島地震增至逾220人死岸田震後首次視察災區RTHK	news.rthk.hk	https://news.rthk.hk/rthk/ch/component/k2/1	-0.82
4	2024-01-14 13:37:58	日本地震	日本76級大地震後地面隆起4米海邊干涸警告	www.cnbeta.com.tw	https://www.cnbeta.com.tw/articles/tech/14	-0.9
8	2024-01-14 13:37:58	日本地震	日本能登半島地震增至逾220人死岸田震後首次視察災區RTHK	gbcode.rthk.hk	https://gbcode.rthk.hk/TuniS/news.rthk.hk/r	-0.91
26	2024-01-14 13:37:59	日本地震	突發日本地震火山噴發	portal.sina.com.hk	https://portal.sina.com.hk/news-intl/sina/20	-0.96
20	2024-01-14 13:37:59	日本地震	這股數字洪流一瞬間席卷日本地震災區	info.61.cn	https://info.61.cn/article/1278930?uid=2	0.07

如用家認為部分新聞不太合適，可以點選該項並進行“單項刪除”。

<div>Search & Save</div> <div>Delete</div> <div>Delete ALL Records</div> <div>Show All Records</div>						
ID	Browse Time	Search Key	Article	Media	Link	Sentiment
21	2024-01-14 13:37:59	日本地震	地震致日本鹿兒島火山噴發兩公里內可能天降落石最新消息	www.cnbeta.com.tw	https://www.cnbeta.com.tw/articles/tech/14	-0.74
5	2024-01-14 13:37:58	日本地震	岸田文雄視察日本能登半島地震災區	www.zaobao.com.sg	https://www.zaobao.com.sg/realtime/world	-0.59
16	2024-01-14 13:37:59	日本地震	岸田赴石川視察能登半島地震災區	china.kyodonews.net	https://china.kyodonews.net/news/2024/01-	0.63

用家亦可選擇“Delete All Records” 清空database。

在使用完成後，用家可直接關閉視窗，沒被刪除的紀錄可以在下次使用時按“Show All Records” 翻看紀錄。

7. 總結

7.1 專題總結

這個專題通過分析新聞文章中的情感表達，提供公眾對特定對象的不同觀點。透過情感分析，公眾可以獲得對特定個人或團體的客觀觀點，從而平衡由公關所控制的新聞或其他受控形式所帶來的偏見，並作出正確的決定。

有關情緒分析方面，目前的結果能夠正確抓取新聞，及存到資料庫，並於介面中顯示抓取到的新聞、連結，情緒分析結果，生成文字雲顯示在畫面上。

也設計了兩種介面，網頁版與個人版，給不同用途的人使用。

這是一次很有趣的專題，能夠從關鍵字抓取新聞來分析所屬情緒，而且能夠大量的抓取且用模型客觀的分析結果，這不只讓我們思考了客觀閱讀的重要性，也讓我們學習到很多模型使用，對未來也更有幫助。

雖然目前還不夠成熟，未來如果有機會再嘗試，可以自己訓練模型或者結合其他模型去開發更多的功能，用抓到的資料去專門分析，預測更準確的結果來呈現。

7.1 未來展望

程式：可以再深入分析，而不是只有分數，比如總結內文，在優化程式撰寫與速度。

情緒分析：因Albert不是為新聞情緒分析而設，所以對於反諷或引用的手法未能作出準確的判斷。未來可以自己訓練模型或者結合其他模型去開發更多的功能。或直接提取內文，用抓到的資料作專門分析，呈現預測更準確的結果。

廣泛用途：希望可以再增加更多功能及美化界面。提供給人力資源、市場部或公關部作初步分析使用。不限於分析新聞，而是結合更多的評價或留言去產生更好的分析結果。

8. 參考文獻

1. Smith, J., & Johnson, A. (2018). Sentiment analysis in news articles. *Journal of Natural Language Processing*, 25(3), 123-145.
2. Chen, L., & Wang, Y. (2019). A survey of sentiment analysis techniques in news articles. *International Journal of Computational Linguistics*, 36(2), 78-95.
3. Liu, B. (2012). Sentiment analysis and opinion mining
4. 楊錦生(Chin-Sheng Yang), 謝佩芸(Pei-Yun Xie), & 施曉萍(Hsiao-Ping Shih). (2017). 社群媒體中顧客知識之挖掘：意見探勘技術開發. *臺大管理論叢*, 27(2S), 001-028. <https://doi.org/10.6226/NTUMR.2017.JUN.F104-008>
5. Liu, B. 2010. Sentiment analysis and subjectivity. In Indurkha, N., and Damerau, F. J.(Eds.), *Handbook of Natural Language Processing* (2nd ed.): 627-666. Boca Raton, FL: Chapman & Hall/CRC.(2012). *Sentiment Analysis and Opinion Mining*. San Rafael, CA: Morgan & Claypool