東吳大學

巨量資料管理學院學士專題成果報告書

專題名稱：YouTube 留言蜚語

06170101 高子淇

06170114 劉　萱

06170139 許靖玟

06170145 周欣德

指導老師：吳政隆老師

中華民國 110 年 01 月

摘要

由於現代大多數人由電視轉向網路平台，在過去大家回到家後會優先打開電視，而如今逐漸轉變成人手一機，使用手機或電腦觀看影片，有鑒於傳統媒體的衰落，新媒體的迅速興起和快速發展，近 93%的民眾曾造訪過 YouTube。YouTube 逐漸成為現代人生活娛樂的重心，藉此本專題想深入研究民眾對於 YouTuber 的觀感、影響力及形象。YouTuber 目前僅能從觀看瀏覽紀錄或使用者行為等數據來追蹤其影片成效，而缺乏最直接表現使用者情緒的留言分析，因此想幫助 YouTuber 更加了解其觀眾對於他影片的感受，以影片底下的留言為依據，進行輿情分析。

本專題使用的資料擷取知名 YouTuber，分別為開箱、談話、生活短劇類型，平均留言數過千，分析 YouTube 發佈的影片底下留言，以一位使用者的留言為一個單位進行資料標記。將資料標記分為五大指標，針對影片的喜愛、YouTuber 的喜愛、激動、諷刺、腥羶色進行程度標記，針對五個標準的評分使用 FastText 模型進行訓練與預測。每部影片的分析結果以網站和 LINE Bot 呈現，使用者只要輸入影片網址，就能即時爬蟲追蹤留言，進行 YouTube 影片的輿情分析，提供視覺化呈現，將非結構化訊息整理為有系統的資料。

本專題的實現目的，近程為提供使用者服務，可根據所欲分析的資料，動態蒐集並回饋結果，對於 youtube 的頻道經營者或是個別粉絲都大有幫助。遠程可作為 youtuber 經營頻道的績效指標，並同時讓社會大眾審視目前社會的概況及趨勢。

目錄

圖目錄

表目錄

緒論

介紹背景

由於新興媒體的開放且可以無所不在和不受約束地表達，改變了傳媒業的秩序，也改變了全球人類的生活方式。相較於傳統媒體，新興媒體的發展更私人化、普及化、自主化，自媒體也應運而生。尤其以透過網路媒體作為發布管道者為甚，加上人手一機和攝影器材的普及，人人都可以拍攝影音資訊上傳到網路平台。根據東方線上 E-ICP 行銷資料庫自 2014 年至 2018 年的資料顯示（圖 1-1），台灣消費者看電視比例逐漸下降。特別是 20-34 歲消費者最為明顯，在 2014 年時仍有高達 8 成 9 有看電視，然而至 2018 年時其收看比例已降至 7 成 9，跌幅高達 1 成。另外在網路影視 APP 上最主要使用族群仍是 20-34 歲（圖 1-2），但 50-64 歲年長者才是使用網路影視 APP 成長最為快速的族群，其使用率五年間成長將近 2.7 倍。由此可見，消費者看電視之比例逐漸下降，網路影視 APP 使用率大幅上升，代表其收視習性已逐漸轉型。

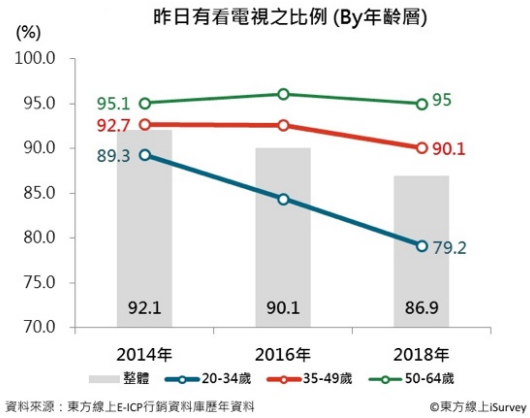

圖 1-1 昨日有看電視之比例

資料來源：許愷洋（2019）。跨代收視習慣大不同，長者網路收視率大漲。iSURVEY 東方線上。取自 https://www.smartm.com.tw/article/36303632cea3

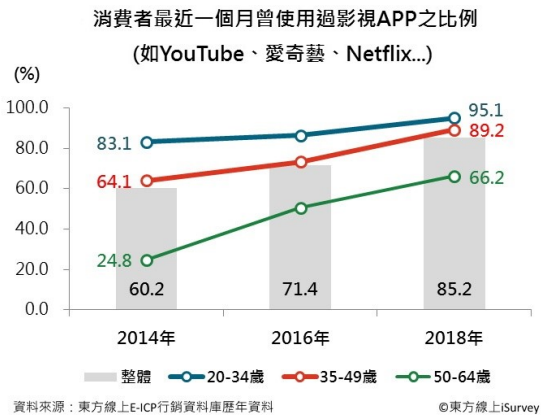

圖 1-2 消費者最近一個月曾使用過影視 APP 之比例

資料來源：許愷洋（2019）。跨代收視習慣大不同，長者網路收視率大漲。iSURVEY 東方線上。取自 https://www.smartm.com.tw/article/36303632cea3

動機

根據 Ipsos 益普索市場研究，台灣每日使用 YouTube 人數高達 7 成，有超過一半的使用者每天平均使用 YouTube 超過 1.5 小時，因此累積龐大的瀏覽量。其中有三分之一的民眾會在購物旅程中使用 YouTube，當中更有超過 8 成民眾表示 YouTube 會影響他們最後購物的決定。由此可見，影音社群平台成為推動品牌策略的得力助手。在「利用機器學習分析中文評論面向之應用 - 以 YouTube 影片評論為例」論文探討中文評論面向之應用，其中面向以一項商品或服務評論者用法說明商品服務好壞的切入角度，針對服務或商品評論者用來敘述此商品的觀點（高亞得，2019）。不同於上述論文以商家知曉網友對於行銷方式與商品的滿意度，本專題以 YouTuber 為出發點，正所謂知己知彼，百戰百勝，除了透過用戶數據分析，推薦用戶客製化的推薦影片，YouTuber 也需要了解自己的影片成效。但目前 YouTuber 僅能從觀看次數、觀看時間、點擊次數等使用者行為數據來追蹤其影片成效，而缺乏最直接表現使用者情緒的留言分析，所以本專題想幫助 YouTuber 更加了解其觀眾對於他影片的感受。

圖 1- 3YouTube 影響台灣民眾的購物決策

資料來源：Darren Freeman, Client Officer, Ipsos in Taiwan。台灣人首選的影音平台是什麼?。Ipsos 益普索市場研究。取自 https://www.ipsos.com/sites/default/files/ct/publication/documents/2019-10/n108_bht_taiwanese_turn_to_youtube_for_online_video_content.pdf

目的

目前 YouTuber 後台管理頁面，僅針對瀏覽次數、觀看時間、點擊次數、分享次數等數據追蹤其影片成效，但這些數據缺乏使用者的留言分析，無法讓 YouTuber 真正瞭解觀眾對於影片的真實回饋，因此本專題針對留言數據進行輿情分析分別為以下五個指標，對 YouTuber 的喜好程度、對影片的喜愛程度、留言內容激動程度、留言內容諷刺程度、留言內容腥羶色程度，希望從留言的評論中得到觀眾所重視的淺在因素，YouTuber 並可根據不同面向進行改善，拍出符合市場需求的影片。

專有名詞定義

YouTuber：以影音網站 YouTube 為主要活動據點的網路紅人或在 YouTube 投稿之影片創作者。

自媒體：指一般民眾藉由網路手段，向不特定的大多數人或者特定的單個人傳遞規範性及非規範性資訊的新媒體，或稱「草根媒體」、「個人媒體」、「公民媒體」。因部落格、共享協作平台與社群網路（如：臉書、Instagram、微博等）的興起，使每個人都具有媒體、傳媒的功能。

機器學習（ML）：透過從過往的資料和經驗中學習並找到其運行規則，最後達到人工智慧的方法。機器學習包含透過樣本訓練機器辨識出運作模式，而不是用特定的規則來編程。這些樣本可以在資料中找到。換句話說，機器學習是一種弱人工智慧，它從資料中得到複雜的函數來學習以創造演算法，並利用它來做預測。

FastText 模型介紹：FastText 是 Facebook AI Research 推出的文字分類和詞訓練工具，其原始碼已經託管在 Github 上。FastText 最大的特點是模型簡單，只有一層的隱層以及輸出層，因此訓練速度非常快，在普通的 CPU 上可以實現分鐘級別的訓練，比深度模型的訓練要快幾個數量級。同時，在多個標準的測試資料集上，FastText 在文字分類的準確率上，和現有的一些深度學習的方法效果相當或接近。

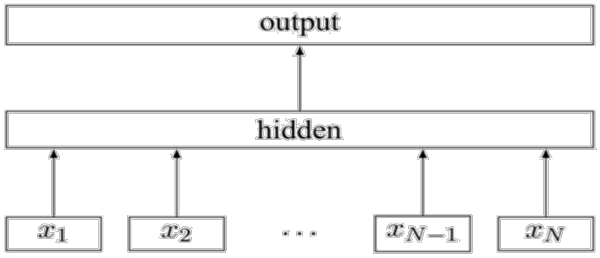

圖 1-4 FastText 架構圖

資料來源：Figure 1- available via license: CC BY

取自 https://www.researchgate.net/figure/Model-architecture-of-FastText-for-a-post-with-N-n-gram-features-x-1-x-N-The_fig1_332287010

Flask：Flask 為一個使用 Python 撰寫的輕量級 Web 應用程式框架，能提供一個平台來呈現資料。當使用者請求該網址之後，根據函數運算之後回傳給使用者資料，也就是網頁的資料。

Ngork：內網伺服器與外界溝通的一個服務。

Kappa 統計量：判斷不同的方法進行實驗時，其信度差異，進行一致性的校驗，本專題利用 kappa 統計量，來判斷每個人標記資料的差異。κ 計算的結果介於-1~1，但通常 κ 是落在 0~1 間，並可分為五組來表示不同等級的吻合度:0.0~0.20 極低的吻合度(slight)、0.21~0.40 一般的吻合度 (fair)、0.41~0.60 中等的吻合度(moderate)、0.61~0.80 高度的吻合度 (substantial)和 0.81~1 幾乎完全吻合(almost perfect)。kappa 值只適用於類別尺度(nominal scale)和序位尺度(ordinal scale)的資料。

研究方法與步驟

研究步驟

1. 選擇影片

本研究資料擷取自十位知名的 YouTuber，分別為這群人、蔡阿嘎、木曜 4 超玩、HOW FUN、STR NETWORK、滴妹、鍾明軒、愛莉莎莎、黃大謙、博恩站起來。將爬取每個頻道從 2020/05/31 前的十部影片，共計 100 部影片底下之留言。留言取自影片發佈至 2020/08/08，共計 253,191 筆留言。再由每部影片隨機抽取 200 則留言進行標記，總計標記 18,651 筆留言。

影片名稱	留言數	平均留言字數	留言字數標準差	平均留言字數(含標點符號)	留言字數標準差(含標點符號)
HowFun / 人生第一口的神秘威士忌=3=	520	14.434615384615400	13.057585288654100	16.31346153846150	15.336426707062500
【狗屎寫手】阿滴真的被玩壞了	1755	14.661538461538500	15.949445393183600	16.427350427350400	17.30811871077950
《一日系列第一百二十集》嗶嗶!!紅線內不准停車!!部哥坤達來開單啦!!-一日派出所員警	2932	23.639154160982300	35.43412215647180	26.145293315143200	38.09581424242700
這群人 TGOP｜婚禮的經典語錄feat.江美琪 劉爵金 達伶Classic Quotations for Wedding	2444	15.245908346972200	23.190195899803300	17.328559738134200	25.48845915215340
【博恩在脫口秀的前一天爆炸】歌唱節目	47	18.46808510638300	19.232955258945500	20.382978723404300	21.160415455247300
學習如何用鏈花開鎖【魔藥學第二課】	1463	14.788106630211900	14.877022401122700	16.816131237183900	16.445065441088900
[破百萬了！?全部Youtuber朋友給我的祝福??！這部片是阿圓送我的百萬禮物??｜愛莉莎莎Alisasa	1078	18.608534322820000	25.143576600695700	22.11873840445270	28.1973217197546
斥資上百萬! 滴妹的飲料店創業之路!？滴妹	2602	18.422751729438900	22.35719750582230	21.13105303612610	24.320747807176900
【嘎奇麥唬爛#54】盲測10家高檔飯店便當！竟這麼貴，有價值嗎？	431	17.842227378190300	14.700203404835200	20.22969837587010	17.424538955616200
台灣人請勿壞習慣！	6289	44.54348863094290	70.637156518513	49.058514867228500	77.35074734549100

圖 2-1　十位 YouTuber 隨機挑選影片的敘述統計

2. 設定標記準則

以一位使用者的留言為一個單位進行資料標記，分別標記對 YouTuber 喜好、對影片內容喜愛、留言激動程度、留言諷刺程度、留言腥羶色程度。其中對 YouTuber 喜好標記的分數為 1 到 5 分，1 分為非常不喜歡，2 分為有點不喜歡，3 分為沒感覺，4 分為有點喜歡，5 分為非常喜歡；對影片內容喜愛標記的分數為 1 到 5 分，1 分為非常不喜歡，2 分為有點不喜歡，3 分為沒感覺，4 分為有點喜歡，5 分為非常喜歡；留言激動程度標記的分數為 1 到 5 分，1 分為沒有情緒波瀾，2 分為有點些微情緒起伏，3 分為有情緒起伏，4 分感到驚嘆，5 分感到不可思議或有髒話；留言諷刺程度標記的分數為 1 到 5 分，1 分為沒有諷刺程度，2 分為輕微諷刺，3 分為類比諷刺，4 分為反諷，5 分為反諷且具有攻擊性；留言腥羶色程度標記的分數為 1 到 5 分，1 分為沒有任何腥羶色字眼，2 分為隱晦暗喻，3 分為明示提及，4 分為講出具體行為，5 分為講出侵犯他人的具體行為。使用機器學習分別對五個標準的評分做出模型，自動計算出評分。

3. 建置資料庫

根據步驟 1 將所選定的 youtuber 名稱及類型存入”YOUTUBER”資料表，並將爬取頻道中的所有影片名稱及相關資訊，存入”video”資料表，而影片底下的留言則存入”total_review”資料表中。步驟 2 為資料標記，其中包含四位標記者的使用資訊及各自的標記結果，分別存入”user”及”total_review”資料表。如下圖：

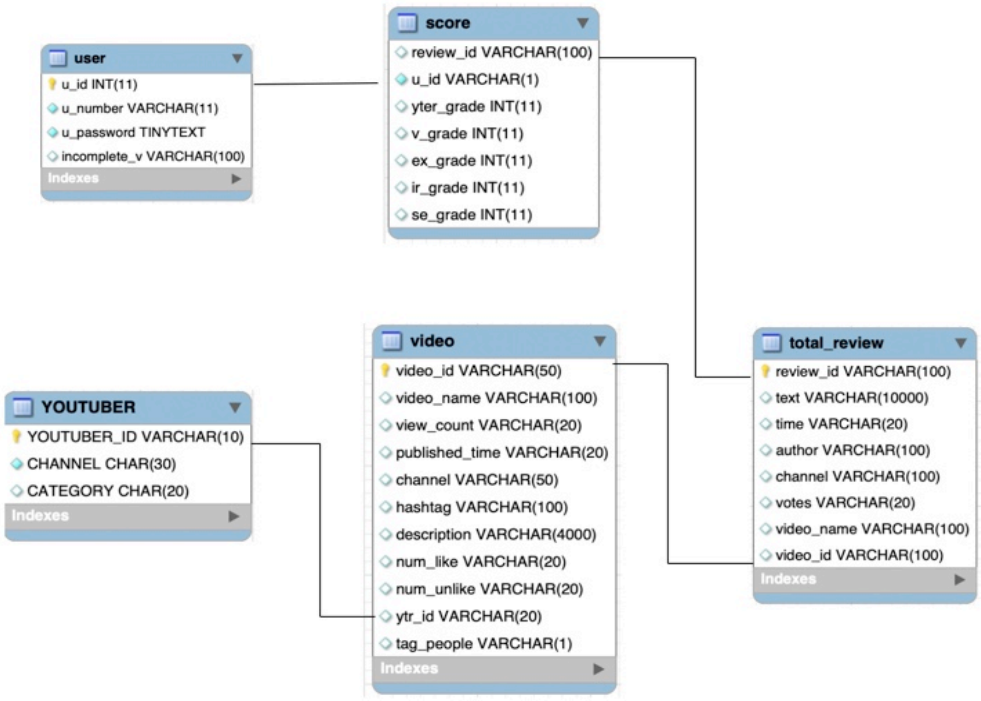

圖 2-2 資料庫實體聯絡模式圖

表 2-1 資料庫各表與其欄位說明

資料表名稱	內容	說明
YOUTUBER	擷取 10 位知名 YouTuber 的名稱及類型	類別包含搞笑、談話、 生活
video	紀錄影片資訊	影片 ID、影片名稱、觀看次數、發布時間、頻道、hashtag、影片描述、按讚人數

user	四位標記者資訊	紀錄標記網頁的使用者的帳號及密碼
total_review	標記資料的影片來源	留言 ID 及內容、發布時間、留言者名稱及頻道、留言按讚數、留言對應的影片名稱及 ID
score	標記者的標記分數	使用者 ID、各個使用者的標記分數

4. 模型訓練

此專題參考 GitHib 上 Chinese-Text-Classification-pytorch 此網站的模型。選擇 7 種模型進行訓練，其中包含 TextCNN、TextRNN、TextRNN_Att、TextRCNN、FastText、 DPCNN、Transformer。訓練模型的數據集來自組內的標記資料，並將所有標記資料的 80%、10%、10%作為訓練集、測試集及驗證集，透過訓練不同的模型及調整參數，比較各個模型在驗證集的準確度及模型的效能提升程度，最終選用 FastText 該組模型進行留言預測及分析。

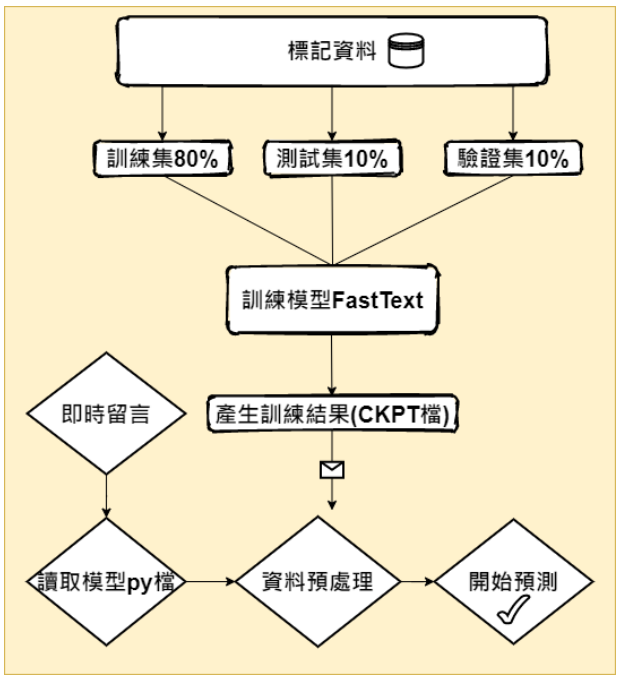

圖 2-3 模型流程圖

研究方法

針對分析對象的影片留言，對影片內容的正負向、激動、諷刺、腥羶色來進行資料標記。

1 研究資料獲取：

1.1 運用 Python 的 Selenium 技術爬取每位 YouTuber 的影片網址及影片 ID。

1.2 參考 Egbert Bouman, (2020, August 7). Re: youtube-comment-downloader. Retrieved from https://github.com/egbertbouman/youtube-comment-downloader 抓取 YouTube 影片資訊及全部留言。影片資訊包含影片標題、發行時間、觀看次數、頻道名稱、主題標籤、內容介紹、喜歡人數及不喜歡人數。

1.3 將上述資料寫入資料庫，分別為 YOUTUBER、video、total_review 三張資料表。資料表 YOUTUBER 放入 YouTuber 資訊，資料表 video 放入影片資訊，資料表 total_review 放入影片留言。

2 設計標記網站：運用 HTML、PHP、MySQL、CSS、JavaScript，並將網站架設在 Heroku 上，標記網站 https://data-marker.herokuapp.com/index.php。

2.1 首頁：總共四位標記者，每人有自己的帳號登入標記頁面，以便資料庫紀錄標記資料。

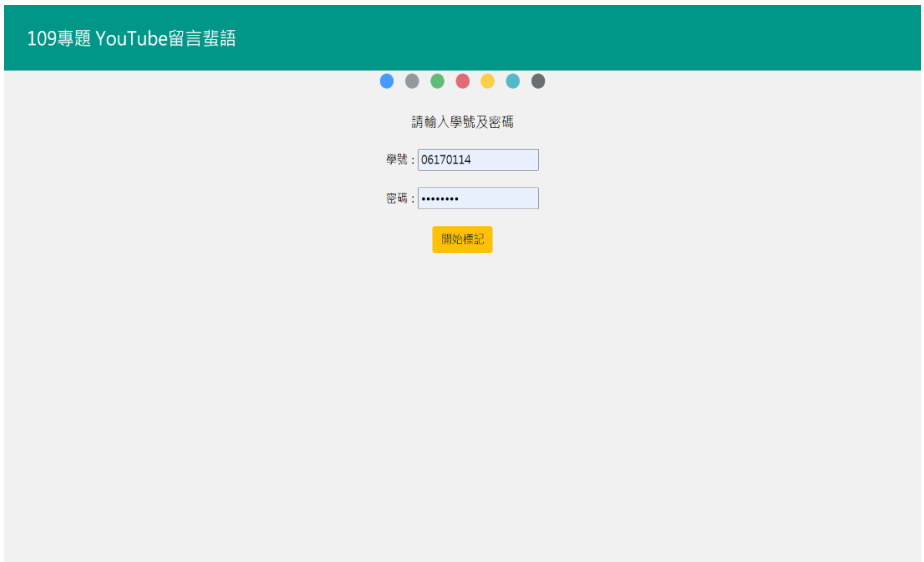

圖 2-4 109 專題 YouTube 留言蜚語登入頁面

2.2　標記頁面：使用 PHP 連接資料庫獲取欲標記資料，將資料呈現在網站上，包含影片名稱、頻道名稱、影片介紹。一頁標記 20 筆留言，標記完成後將結果回傳資料庫。

圖 2-5 109 專題 YouTube 留言蜚語標記頁面

3　標記資料：

3.1　訂定標記準則

表 2-2 標記定義

	1	2	3	4	5
對 YouTuber 的喜好程度	酸民（發表尖酸苛薄的評論，不在乎事情對錯）	有理由的討厭，發表自己的言論或是負面評價	客觀陳述論點	對 youtuber 有正面評價，但未表明喜好	有出現護航 YouTuber 之言論或直接表明喜愛(喜歡 xxx)
對影片的喜愛程度	強烈譴責	對於部分內容討厭(討厭>喜愛)	客觀陳述論點	對影片有正面評價(喜愛>討厭)	最愛、大推
留言內容激動程度	不帶情緒之言論(建議、好像、不太)	些微情緒起伏、哈哈	有情緒起伏、沒想到、哈哈哈	驚嘆，太、！、真的	感到不可思議，超(級)、千萬、髒話、！！！

標記分佈留言內容諷刺程度	沒有諷刺	當事人可接受的玩笑諷刺	類比諷刺(用愛發電，用屁發電)	反諷（好棒棒、外國的月亮比較圓）	反諷且具有攻擊性
留言內容腥羶色程度	沒有腥羶色	暗示提到性特徵、性暗示	直接表明性特徵或提到相關形容詞(ex:淫蕩)	講出具體的腥羶色的行為	講出對別人想做腥羶色具體行動

4 標記結果評估：

4.1 標記資料敘述性統計

	yter_count	v_count	ex_count	ir_count	se_count
1	49	8	541	5052	5043
2	68	123	1923	107	106
3	4087	4207	1901	29	28
4	877	773	597	1	10
5	108	78	227	0	2

圖 2-6 標記者 1 標記分佈

	yter1	v1	ex1	ir1	se1
count	5189.000000	5189.000000	5189.000000	5189.000000	5189.000000
mean	3.178647	3.152245	2.623434	1.032376	1.038543
std	0.520844	0.464575	0.967493	0.208974	0.253081
min	1.000000	1.000000	1.000000	1.000000	1.000000
25%	3.000000	3.000000	2.000000	1.000000	1.000000
50%	3.000000	3.000000	3.000000	1.000000	1.000000
75%	3.000000	3.000000	3.000000	1.000000	1.000000
max	5.000000	5.000000	5.000000	4.000000	5.000000

圖 2-7 標記者 1 敘述性統計

	yter_count	v_count	ex_count	ir_count	se_count
1	29	18	823	3159	3221
2	27	88	1226	57	31
3	2657	2522	503	25	12
4	472	618	606	24	7
5	88	27	115	8	2

圖 2-8 標記者 2 標記分佈

	yter2	v2	ex2	ir2	se2
count	3273.000000	3273.000000	3273.000000	3273.000000	3273.000000
mean	3.172013	3.167430	2.377941	1.064467	1.025665
std	0.515694	0.492687	1.148706	0.385954	0.229172
min	1.000000	1.000000	1.000000	1.000000	1.000000
25%	3.000000	3.000000	1.000000	1.000000	1.000000
50%	3.000000	3.000000	2.000000	1.000000	1.000000
75%	3.000000	3.000000	3.000000	1.000000	1.000000
max	5.000000	5.000000	5.000000	5.000000	5.000000

圖 2-9 標記者 2 敘述性統計

	yter_count	v_count	ex_count	ir_count	se_count
1	13	15	818	3212	3321
2	34	105	1250	82	16
3	2971	2728	765	12	24
4	268	464	331	29	3
5	79	53	201	30	1

圖 2-10 標記 3 標記分佈

	yter3	v3	ex3	ir3	se3
count	3365.000000	3365.000000	3365.000000	3365.000000	3365.000000
mean	3.108767	3.129272	2.360178	1.093016	1.022883
std	0.432821	0.482995	1.127890	0.500263	0.213430
min	1.000000	1.000000	1.000000	1.000000	1.000000
25%	3.000000	3.000000	2.000000	1.000000	1.000000
50%	3.000000	3.000000	2.000000	1.000000	1.000000
75%	3.000000	3.000000	3.000000	1.000000	1.000000
max	5.000000	5.000000	5.000000	5.000000	5.000000

圖 2-11 標記者 3 敘述性統計

	yter_count	v_count	ex_count	ir_count	se_count
1	40	21	1639	5433	5599
2	15	22	1792	39	12
3	5012	5145	888	11	13
4	208	215	686	52	8
5	371	243	641	111	14

圖 2-12 標記者 4 標記分佈

	yter4	v4	ex4	ir4	se4
count	5646.000000	5646.000000	5646.000000	5646.000000	5646.000000
mean	3.151435	3.112823	2.450584	1.117074	1.020900
std	0.554796	0.465103	1.323877	0.631280	0.251667
min	1.000000	1.000000	1.000000	1.000000	1.000000
25%	3.000000	3.000000	1.000000	1.000000	1.000000
50%	3.000000	3.000000	2.000000	1.000000	1.000000
75%	3.000000	3.000000	3.000000	1.000000	1.000000
max	5.000000	5.000000	5.000000	5.000000	5.000000

圖 2-13 標記者 4 敘述性統計

4.2 標記組間、組內差異

	p12	p13	p14	p23	p24	p34	p123
yter	0.648235	0.457350	0.411481	0.406093	0.265054	0.120493	0.663866
video	0.535283	0.477610	0.130716	0.393188	0.157016	0.320863	0.449743
agitate	0.249718	0.173917	0.226127	0.414981	0.329263	0.404061	0.260204
irony	0.223402	0.199560	0.166899	0.047440	0.041783	0.070697	0.475524
sex	0.419029	0.176495	0.220084	-0.009615	0.083013	0.149265	-0.020833

圖 2-14 Kappa 組間差異

	kao_group_in	liu_group_in	hsu_group_in	chou_group_in
yter	0.858290033065659	0.910913140311804	0.7665732959850610	0.9373629815220800
video	0.8177178271965000	0.8989286436223970	0.7319034852546920	0.9318801089918260
ex	0.623149394347241	0.8714818146767770	0.604221635883905	0.9354630525976120
irony	0.5397008055235900	0.6575342465753420	0.6563573883161510	0.7071742313323570
sex	0.48892674616695100	0.31623931623931600	1.0	1.0

圖 2-15 Kappa 組內差異

5 模型選用

5.1 模型訓練環境：分別以下列兩種環境進行模型訓練。

5.1.1 Google Colab 環境：Tesla P100-PCIE-16GB GPU

5.1.2 Linux 伺服器：Geforce GTX 1060 6GB *2 GPU

5.2 模型資料預處理：

模型訓練資料集主要包含留言及對應的五個標記分數。其中分數的數值來自四位使用者的標記結果。從資料庫取出留言及標記分數，依照留言 id 統整相同留言不同使用者的標記分數，最終以四位使用者的標記分數加總除以標記者人數作為最終的分數依據。若統整分數為小數型態，則無條件進位至下位整數，透過以上的計算方法，可得到單一數值作為後續訓練模型的留言對應分數。留言的統整方式則是整理文字型態，並進行斷詞作為訓練模型的文本輸入。本專題嘗試使用 jieba 及 CKIT 兩種斷詞方式，有鑒於 jieba 的結果較準確且大幅縮短運算的時間，故最終選擇此方法作為模型輸入的文字前處理，藉由上述方法，使資料符合模型之格式進行資料切割，隨機擷取 80% 、10%、10%作為訓練集、測試集及驗證集。

5.3 模型訓練過程：

從訓練資料中的斷詞結果建立模型中的字典及索引。測試集將根據字典，將留言轉成索引。透過計算標記資料中的留言字數平均數及眾數，來選擇模型輸入的文字長度，由於部分留言長度過長影響留言平均數，因此改取留言字數的眾數來作為模型輸入的文字長度依據。

標記資料的過程中，發現多數留言在判斷 Youtuber 喜好、影片喜好、諷刺及腥羶色程度這些指標的分數並無明顯差異，以至於標記結果分佈過於懸殊，使模型無法達到良好的訓練效果，對於此困難本專案使用了兩種解決方法。第一，針對不平衡數據，進行 Oversample(隨機採樣)來增加樣本數，透過隨機複製樣本的方式，使少樣本的類別數量能與樣本數較多的標記資料數量達到一致，再重新進行模型訓練。第二，修正研究方法，將諷刺程度、腥羶色程度、YouTuber 喜好、影片內容喜好的四個指標準則改為二分法，只判斷是與否的程度分析。例如：諷刺和腥羶色，將原本標記為 1 分，不帶有諷刺、腥羶色程度的留言，維持 1 分的標記分數，而原先標記 2-5 分帶有諷刺、腥羶色程度的留言，合併成 2 分。Youtuber 喜好、影片內容喜好，同樣修改為判斷留言是否為中立程度的分析，因此將原本標記為 3 分的中立選項改成 1 分，其餘帶有情緒的分數 1,2,4,5 分更改為 2 分，根據新的統整分數作為模型判斷的指標依據。實驗修正如下：

表 2-3 第一次留言指標修正

分數說明	Youtuber 喜好	影片喜好	諷刺程度	腥羶色
1	對 Youtuber 無明顯喜愛偏好	對影片內容無表示意見	留言中無諷刺語句	留言中無腥羶色字眼
2	對 Youtuber 表示喜愛程度	對影片內容提出看法	留言中有諷刺語句	留言中有腥羶色字眼，甚至攻擊他人的語句

5.4 模型調整

5.4.1 第一次修正留言輿情指標的模型訓練：

藉由上述調整分數的結果，重新進行模型訓練，結果發現腥羶色及諷刺程度這兩項指標採用二分法的分類效果有改善，正確率及 F1 Score 的分數都有顯著的提升。相反的，對於 Youtuber 喜好及影片喜好的模型訓練結果，F1 score 的分數雖然有進步，但是效果提升的程度不明顯。結果如下表：

表 2-4 影片喜好程度的模型訓練結果對比

	標記分數 1-5 分	標記分數 1、2 分
Test Accuracy	84.64%	85.26%
Macro Average F1 score	0.1834	0.5387
Confusion Matrix	Confusion Matrix 0 0 2 0 0 0 0 21 0 0 0 0 1223 0 0 0 0 173 0 0 0 0 26 0 0	Confusion Matrix 1212 11 202 20

表 2-5 YouTuber 喜好程度的模型訓練結果對比

	標記分數 1-5 分	標記分數 1、2 分
Test Accuracy	20%	86.37%
Macro Average F1 score	0.0668	0.6471
Confusion Matrix	Confusion Matrix 0 0 9472 149 0 0 0 9621 0 0 0 0 9620 1 0 0 0 9621 0 0 0 0 9621 0 0	Confusion Matrix 1190 10 187 58

5.4.2 第二次修正留言輿情指標的模型訓練：

有鑒於上述模型對於 Youtuber 喜好、影片內容喜好程度的訓練結果. 不盡理想，因此重新將這兩個指標的實驗進行修正，將二分法修正為三種指標的分類方式。原本資料中 1-2 分，無明顯喜愛的分數合併為 1 分，再將原本 3 分的中立分數變更為 2 分，最後則是將 4-5 分，具有喜好程度的分數合併為 3 分，利用新產生的 3 個指標進行模型訓練。實驗修正如下：

表 2-6 第二次留言指標修正

分數說明	Youtuber 喜好	影片喜好
1	對 Youtuber 具有負面態度	對影片內容表示負面想法
2	對 Youtuber 態度中立	對該影片內容無表示意見
3	對 Youtuber 具有正向態度	對影片內容有正面肯定

5.4.3 最終選用的留言輿情指標：留言指標中的激動程度保持五種指標，YouTuber、影片內容的喜愛則做為三類型的指標判斷，而留言諷刺、腥羶色程度則改為二分法。以上述的資料前處理將留言及標記分數輸入模型，以做為模型的訓練資料集。

5.5 模型結果評估：

利用"dropout"及"hidden_size"調整參數來評估模型準確率，dropout 分別 0.1 及 0.3，hidden_size 分別有五組數據，訓練結果如下表：

YouTuber 喜好程度

表 2-7 YouTuber 喜好程度參數調整一

固定參數 Dropout=0.1，更動 hidden_size 和 hidden_size2 參數，比較訓練結果。

	hidden_size=2000 hidden_size2=1000	hidden_size=2500 hidden_size2=1250	hidden_size=3000 hidden_size2=1500
Test Accuracy	94.35%	94.42%	94.49%
Dev Accuracy	95.31%	95.45%	95.45%
Macro Average F1 score	0.8458	0.8503	0.8695
Confusion Matrix	Confusion Matrix 18 5 1 4 1166 22 3 47 185	Confusion Matrix 17 7 0 4 1162 26 1 43 191	Confusion Matrix 18 6 0 2 1162 28 1 43 191

	hidden_size=5000 hidden_size2=2500	hidden_size=10000 hidden_size2=5000
Test Accuracy	94.76%	94.21%
Dev Accuracy	95.86%	95.38%
Macro Average F1 score	0.8719	0.8512
Confusion Matrix	Confusion Matrix 18 6 0 2 1169 21 1 46 188	Confusion Matrix 18 6 0 5 1160 27 1 45 189

表 2-8 YouTuber 喜好程度參數調整二

固定參數 Dropout=0.3，更動 hidden_size 和 hidden_size2 參數，比較訓練結果。

	hidden_size=2000 hidden_size2=1000	hidden_size=2500 hidden_size2=1250	hidden_size=3000 hidden_size2=1500
Test Accuracy	92.21%	92.69%	92.42%
Dev Accuracy	93.31%	94.21%	93.52%
Macro Average F1 score	0.7976	0.7836	0.7772
Confusion Matrix	Confusion Matrix 13 11 0 2 1155 35 0 65 170	Confusion Matrix 11 13 0 1 1150 41 1 50 184	Confusion Matrix 11 13 0 1 1158 33 1 62 172

	hidden_size=5000 hidden_size2=2500	hidden_size=10000 hidden_size2=5000
Test Accuracy	92.83%	92.9%
Dev Accuracy	93.87%	94.34%
Macro Average F1 score	0.8053	0.816
Confusion Matrix	Confusion Matrix 14 10 0 3 1160 29 1 61 173	Confusion Matrix 14 10 0 1 1161 30 1 61 173

分析結果：透過正確率(Dev Accuracy)及混淆矩陣(Confusion Matrix)判斷模型好壞，了解各指標預測情形。發現 dropout=0.1 時，不論 hidden_size 和 hidden_size2 如何調整其預測結果會稍微優於 dropout=0.3 的預測能力。

影片喜好程度

表 2-9 影片喜好程度參數調整一

固定參數 Dropout=0.1，更動 hidden_size 和 hidden_size2 參數，比較訓練結果。

	hidden_size=2000 hidden_size2=1000	hidden_size=2500 hidden_size2=1250	hidden_size=3000 hidden_size2=1500
Test Accuracy	93.45%	93.25%	93.04%

Dev Accuracy	92.90%	92.83%	91.87%
Macro Average F1 score	0.8039	0.8118	0.7875
Confusion Matrix	Confusion Matrix 16 14 2 1 1170 34 1 43 170	Confusion Matrix 19 12 1 5 1163 37 1 42 171	Confusion Matrix 15 15 2 2 1164 39 1 42 171

	hidden_size=5000 hidden_size2=2500	hidden_size=10000 hidden_size2=5000
Test Accuracy	93.11%	92.14%
Dev Accuracy	92.97%	90.70%
Macro Average F1 score	0.819	0.7779
Confusion Matrix	Confusion Matrix 19 11 2 2 1165 38 1 46 167	Confusion Matrix 16 15 1 2 1174 29 1 66 147

表 2-10　影片喜好程度參數調整二

固定參數 Dropout=0.3，更動 hidden_size 和 hidden_size2 參數，比較訓練結果。

	hidden_size=2000 hidden_size2=1000	hidden_size=2500 hidden_size2=1250	hidden_size=3000 hidden_size2=1500
Test Accuracy	92.21%	91.87%	91.45%
Dev Accuracy	91.39%	90.90%	91.11%
Macro Average F1 score	0.7529	0.7648	0.7526
Confusion Matrix	Confusion Matrix 12 18 2 1 1168 36 0 56 158	Confusion Matrix 14 17 1 1 1172 32 0 67 147	Confusion Matrix 12 18 2 1 1168 36 0 56 158

	hidden_size=5000 hidden_size2=2500	hidden_size=10000 hidden_size2=5000
Test Accuracy	92.49%	92.01%
Dev Accuracy	91.66%	90.35%
Macro Average F1 score	0.798	0.7736
Confusion Matrix	Confusion Matrix 18 12 2 4 1156 45 1 45 168	Confusion Matrix 15 16 1 2 1158 45 1 51 162

分析結果：透過正確率(Dev Accuracy)及混淆矩陣(Confusion Matrix)判斷模型好壞，了解各指標預測情形。發現 dropout=0.1 時，不論 hidden_size 和 hidden_size2 如何調整其預測結果會稍微優於 dropout=0.3 的預測能力。

是否有腥羶色

表 2-11 腥羶色程度參數調整一

固定參數 Dropout=0.1，更動 hidden_size 和 hidden_size2 參數，比較訓練結果。

	hidden_size=2000 hidden_size2=1000	hidden_size=2500 hidden_size2=1250	hidden_size=3000 hidden_size2=1500
Test Accuracy	99.79%	99.45%	99.45%
Dev Accuracy	99.10%	99.10%	99.10%
Macro Average F1 score	0.9554	0.881	0.9034
Confusion Matrix	Confusion Matrix 1433 0 3 15	Confusion Matrix 1430 3 5 13	Confusion Matrix 1426 7 1 17

	hidden_size=5000 hidden_size2=2500	hidden_size=10000 hidden_size2=5000
Test Accuracy	99.59%	99.59%
Dev Accuracy	99.17%	99.10%

Macro Average F1 score	0.9156	0.9156
Confusion Matrix	Confusion Matrix 1430 3 3 15	Confusion Matrix 1430 3 3 15

表 2-12 腥羶色程度參數調整二

固定參數 Dropout=0.3，更動 hidden_size 和 hidden_size2 參數，比較訓練結果。

	hidden_size=2000 hidden_size2=1000	hidden_size=2500 hidden_size2=1250	hidden_size=3000 hidden_size2=1500
Test Accuracy	99.59%	99.52%	99.31%
Dev Accuracy	99.24%	99.17%	99.10%
Macro Average F1 score	0.9107	0.8927	0.8594
Confusion Matrix	Confusion Matrix 1431 2 4 14	Confusion Matrix 1431 2 5 13	Confusion Matrix 1428 5 5 13

	hidden_size=5000 hidden_size2=2500	hidden_size=10000 hidden_size2=5000
Test Accuracy	99.31	99.31
Dev Accuracy	99.17%	98.90%
Macro Average F1 score	0.8594	0.8512
Confusion Matrix	Confusion Matrix 1428 5 5 13	ConfusionMatrix 1429 4 6 12

分析結果：由於判斷腥羶色程度的資料中，不具腥羶色佔大多數，因此在不同的 dropout 參數底下，預測分數都能表現的很好，幾乎接近 100%，但混淆矩陣可以說明，當 dropout=0.1 時，即使具有腥羶色的留言只有 20 筆，模型也能正確的判斷出實際具有腥羶色，因此最終選擇 dropout=0.1 為參數。

激動程度

表 2-13 激動程度參數調整一

固定參數 Dropout=0.1，更動 hidden_size 和 hidden_size2 參數，比較訓練結果。

	hidden_size=2000 hidden_size2=1000	hidden_size=2500 hidden_size2=1250	hidden_size=3000 hidden_size2=1500
Test Accuracy	80.43%	80.91%	81.32%
Dev Accuracy	79.81%	80.77%	81.32%
Macro Average F1 score	0.7958	0.7976	0.8077
Confusion Matrix	Confusion Matrix 252 39 5 1 2 34 439 43 7 3 6 52 252 9 5 10 13 16 151 8 6 7 11 7 73	Confusion Matrix 256 34 6 1 2 32 442 38 13 1 8 50 243 18 5 9 7 13 161 8 5 7 8 12 72	Confusion Matrix 250 42 5 2 0 29 452 37 7 1 7 56 246 10 5 10 11 13 158 6 4 9 8 9 74

	hidden_size=5000 hidden_size2=2500	hidden_size=10000 hidden_size2=5000
Test Accuracy	78.77%	80.84%
Dev Accuracy	79.53%	82.01%
Macro Average F1 score	0.7892	0.8074
Confusion Matrix	Confusion Matrix 216 70 6 4 3 32 437 46 10 1 8 47 257 10 2 6 17 12 157 6 5 7 9 7 76	Confusion Matrix 245 41 7 2 4 45 424 45 10 2 10 34 264 12 4 6 10 15 159 8 5 5 8 5 81

表 2-14 激動程度參數調整二

固定參數 Dropout=0.3，更動 hidden_size 和 hidden_size2 參數，比較訓練結果。

	hidden_size=2000 hidden_size2=1000	hidden_size=2500 hidden_size2=1250	hidden_size=3000 hidden_size2=1500
Test Accuracy	71.54%	73.95%	72.78%
Dev Accuracy	73.26%	73.54%	73.54%
Macro Average F1 score	0.7016	0.7256	0.7166
Confusion Matrix	Confusion Matrix 204 85 9 1 0 35 423 59 8 1 6 77 224 12 5 9 16 32 130 11 2 17 14 14 57	Confusion Matrix 224 66 8 1 0 42 415 58 10 1 8 62 235 13 6 10 13 26 139 10 5 7 15 17 60	Confusion Matrix 216 78 4 1 0 42 422 54 8 0 7 76 226 11 4 11 17 28 135 7 5 13 15 14 57

	hidden_size=5000 hidden_size2=2500	hidden_size=10000 hidden_size2=5000
Test Accuracy	75.6%	75.74%
Dev Accuracy	74.57%	75.05%
Macro Average F1 score	0.7491	0.7519
Confusion Matrix	Confusion Matrix 232 60 6 1 0 43 426 48 8 1 6 69 232 13 4 11 16 22 142 7 5 10 11 13 65	Confusion Matrix 239 55 4 1 0 49 416 51 9 1 9 66 230 12 7 11 15 25 139 8 6 9 9 9 71

分析結果：當 dropout=0.1 時，模型的準確率可達 80%，而 dropout=0.3 時，模型的準確率只有到 75%，故最終選擇 dropout=0.1 作為後續的參數選擇。

諷刺程度

表 2-15 諷刺程度參數調整一

固定參數 Dropout=0.1，更動 hidden_size 和 hidden_size2 參數，比較訓練結果。

	hidden_size=2000 hidden_size2=1000	hidden_size=2500 hidden_size2=1250	hidden_size=3000 hidden_size2=1500
Test Accuracy	99.28%	99.28%	98.21%
Dev Accuracy	98%	98.21%	97.66%
Macro Average F1 score	0.8868	0.8829	0.8793
Confusion Matrix	Confusion Matrix 1381 10 15 45	Confusion Matrix 1383 8 17 43	Confusion Matrix 1382 9 17 43

	hidden_size=5000 hidden_size2=2500	hidden_size=10000 hidden_size2=5000
Test Accuracy	98.48%	98.9%
Dev Accuracy	98.21%	98.14%
Macro Average F1 score	0.8996	0.9244
Confusion Matrix	Confusion Matrix 1383 8 14 46	Confusion Matrix 1388 3 13 47

表 2-16 諷刺程度參數調整二

固定參數 Dropout=0.3，更動 hidden_size 和 hidden_size2 參數，比較訓練結果。

	hidden_size=2000 hidden_size2=1000	hidden_size=2500 hidden_size2=1250	hidden_size=3000 hidden_size2=1500
Test Accuracy	98.21%	98.21%	98.41%
Dev Accuracy	97.93%	98.07%	98.21%
Macro Average F1 score	0.8772	0.8793	0.8923
Confusion Matrix	Confusion Matrix 1383 8 18 42	Confusion Matrix 1382 9 17 43	Confusion Matrix 1384 7 16 44

	hidden_size=5000 hidden_size2=2500	hidden_size=10000 hidden_size2=5000
Test Accuracy	98.41%	98.55%
Dev Accuracy	98.28%	92.28%
Macro Average F1 score	0.8959	0.8981
Confusion Matrix	Confusion Matrix 1382 9 14 46	Confusion Matrix 1387 4 17 43

分析結果：諷刺程度的原始資料中，具有諷刺意味的資料筆數偏少，故單看準確率會誤判模型的訓練成效，因此需檢查混淆矩陣來作為模型判斷的依據，結果可以發現當 hidden_size=10000 與 hidden_size2=5000 的預測能力最佳，但 dropout=0.1 的準確率會優於 dropout=0.3，故最終選擇 dropout=0.1。

6 模型成果：

選擇 5 個指標分別對應的模型最佳參數，以測試集、驗證集的正確率和 Marco Average F1 Score 三個指標去選擇。最後選擇以下五組參數，作為此專題預測的模型參數。

表 2-17 模型參數選擇

指標名 / 模型參數	Youtuber 喜好	Video 喜好	腥羶色
	dropout=0.1 hidden_size=3000 hidden_size2=1500	dropout=0.1 hidden_size=5000 hidden_size2=2500	dropout=0.1 hidden_size=5000 hidden_size2=2500

指標名 / 模型參數	激動程度	諷刺程度
	dropout=0.1 hidden_size=10000 hidden_size2=5000	dropout=0.1 hidden_size=10000 hidden_size2=5000

7 預測成果：

訓練完成時將結果儲存成 checkpoint 檔(ckpt)。預測時只需將 CKPT 檔讀取回來，並選擇其對應參數的 py 檔案，即可獲得分析結果。由於不需重複訓練資料，因此大幅節省等待時間。

8 網站：運用 Flask、HTML、MySQL、CSS、JavaScript 製作呈現互動式網站，除了電腦的網頁版之外另外為手機使用者設計手機版網頁。

8.1 首頁：輸入欲分析影片網址，同時提供簡單網頁介紹。先辨別是否為 YouTube 影片網址，再將影片 ID 回傳至後台資料庫，若影片資訊已存在於資料庫中便直接從資料庫 result 資料表抓去結果回傳到前端呈現分析頁面。若影片資訊不存在於資料庫中便開始執行爬取留言資料，資料爬取完成後將影片 ID 告知模型伺服器並執行模型預測、寫入資料庫 result 資料表，模型伺服器再將執行完成訊息回傳，便可從資料庫 result 資料表抓去結果至前端呈現分析頁面。

8.2 分析頁面：總表及五個指標的分析呈現。透過後端得到 5 項指標預測分數呈現於網站上，內容以數字和圖表視覺化呈現。

8.2.1 總表：可以在此頁面看到該輸入影片的畫面連結方便直接查看及影片相關資訊。下方呈現 Youtuber 喜好程度、影片喜好程度、激動程度三指標平均雷達圖、Youtuber 喜好程度、影片喜好程度、激動程度儀表板及諷刺、腥羶色留言所佔比例，以及五項指標各個留言分佈。

8.2.2 五個指標：可以在此頁面看到影片相關資訊、留言總數、平均分數儀表板。並以長條圖及圓餅圖呈現分佈情況，最下方有每個分數的留言內容方便查看各個分數實質留言內容。

8.2.3 指標說明頁面：提供 Youtuber 喜好程度、影片喜好程度、激動程度、諷刺、腥羶色標準說明。

9 LINE Bot：製作 LINE Bot 方便手機使用者，LINE Bot 帳號@600ypyxn，使用者只要輸入影片網址便可以獲得精簡版分析圖，另外得到完整分析網站之網址。

9.1 開啟聊天對話後，輸入 YouTube 影片網址。

9.2 機器人自動回覆 Youtuber 喜好程度、影片喜好程度、激動程度三指標平均雷達圖及諷刺、腥羶色留言所佔比例簡易分析。另外提供詳細分析網站之網址提供使用者查看。網頁版分析特別設計手機版，用手機也能有完整的觀看享受。

9.2.1 若輸入內容不屬於 YouTube 影片網址，則機器人自動回覆「請輸入 YouTube 網址！」

9.2.2 若輸入的 YouTube 影片不存在資料庫則進行 5.1.1 的處理步驟。若模型伺服器回傳失敗，則機器人自動回覆「YouTube 忙線中，請再試一次！」

結果呈現

實際成果

網站架構

使用者輸入網址至網站首頁，網站伺服器接收 video_id 後，判斷是否進行爬蟲。將爬蟲完成訊息傳送至模型伺服器，模型伺服器接收訊息後則開始進行分析，並將結果回傳至資料庫，後續有相同影片需分析時，可快速從資料庫抓取結果並呈現至網頁前端。分析結果後，網頁伺服器將統計並呈現圖表至網頁前端。詳細請見圖 3-1。

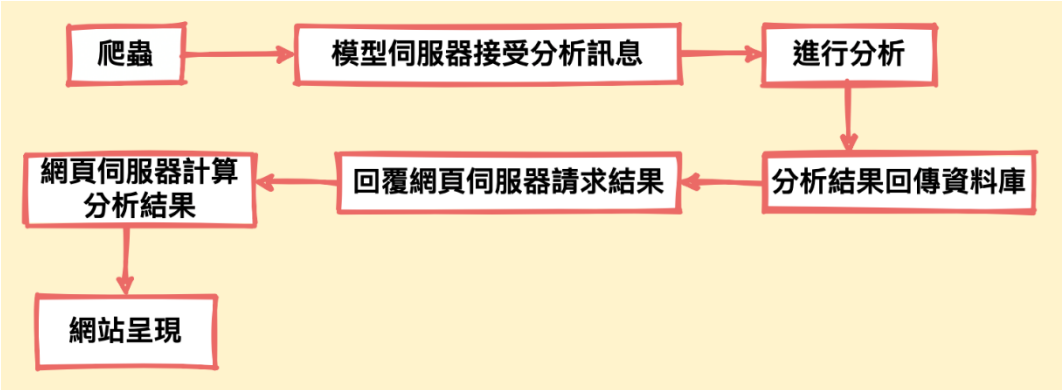

圖 3-1 網站架構流程圖

網站呈現

最終結果以互動式網站和 LINE Bot 呈現，其中以 Flask 作為網站開發的框架，並搭配 html 與 css 語法使用。使用者輸入影片網址就能即時抓取影片留言，並進行五類指標的分數預測，分析結果以視覺化圖表呈現，使用者可以得知留言的各項分數。

首頁：

1. 內含影片輸入、介面使用及操作步驟、網頁的熱門影片排行榜。當使用者輸入影片網址，將判斷影片資訊是否存在資料庫，若影片不存在將進行即時爬蟲，若存在則會從資料庫抓取資料。而網站可判斷三種網址的輸入形式，電腦版網址、app 網址、手機網頁版網址。詳細請見圖 3-2。

圖 3-2 網站首頁

內頁：

1. 內含分析結果（儀表板、長條圖、雷達圖、圓餅圖），及影片相關資料（背景資料、影片播放）。詳細請見圖 3-3 和 3-4。

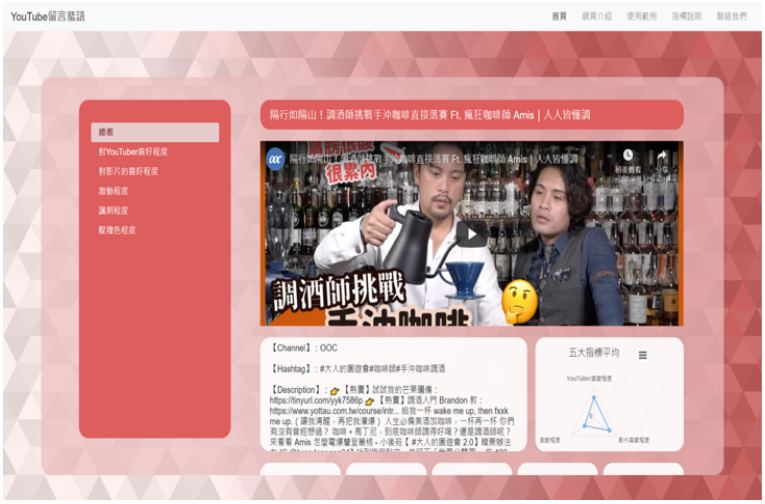

圖 3-3 網站總表內頁分析

圖 3-4 網站個別指標分析頁面

網站功能

1. 視覺化分析：以五大指標分數以儀表板、長條圖、雷達圖、圓餅圖等方式呈現給使用者，易於瀏覽結果。

2. 完整性分析：在個別指標分析內頁，可完整查看所有留言及對應的分數指標。詳細請見圖 3-5。

圖 3-5 完整留言內容呈現

3. 排行榜：排行榜根據使用者的搜尋次數，進行次數統計，可讓用戶知道哪部影片得到較高的關注程度。詳細請見圖 3-6。

留言蜚語熱門影片排行榜

排名	影片名稱	點擊數
1	【千千進食中】忠孝東路吃九間？...	42
2	《一日系列第一百五十集》蛤～阿...	31
3	非整人！跟三原的戀愛故事❤️	28
4	【蔡桃貴成長日記#78】第一次...	22
5	【開麵時刻】沒人敢說的真相！台...	22
6	隔行如隔山！調酒師挑戰手沖咖啡...	21
7	「君の名は。」スパークル/RA...	17
8	(여자)아이들((G)I-DL...	15
9	(여자)아이들((G)I-DL...	14
10	【蔡阿嘎老屋改建計畫#1】幫阿...	13
11	黃明志 Ft. 澎恰恰 202...	13
12	《木曜在幹嘛》第二屆YARTI...	12
13	【小吳】直接裁源‼️『權志琳假...	12
14	【黃阿瑪的後宮生活】24小時陪...	12
15	是網紅是明星！	12
16	【蔡桃貴】訓話蔡波能！不能帶女...	12
17	#61【台阿莫Life】學電影...	12
18	The last episod...	11

圖 3-6 熱門影片排行榜

4. 社群連結：提供 Line@條碼，讓網站可以更容易被推廣，也能讓使用者在第一時間透過 Line 得到簡易分析結果。

圖 3-7 Line@條碼

Line Bot 結果呈現

因應社群軟體興盛，本專題也設計出 Line Bot 的功能，使用者也可藉由 line 進行分析，得到簡易的分析雷達圖，並附上網址，讓使用者藉由點擊來觀看完整的分析結果。

Line Bot 的操作和網站的使用方式一樣，使用者只需輸入 YouTube 影片網址即可產生簡易分析，若輸入的影片不存在資料庫，也同樣會進行即時爬蟲。當中處理回覆訊息的方式是使用 reply_message，但此函數須在三十秒內進行回覆，超時則不予回覆，但由於本專題需進行即時爬蟲，時間會受到留言數量影響，故 LINE Bot 使用 push_message 自動推播的方式做回覆，因此可以在完成爬蟲、分析後進行回覆分析結果。

回覆內容為呈現影片留言數量、網頁詳細分析結果網址；YouTuber 喜好程度、影片喜好程度、激動程度三指標的簡易雷達圖；諷刺、腥羶色留言佔比。詳細請見圖 3-8。

圖 3-8 Line Bot 帳號
YouTube 留言蜚語

圖 3-9 Line Bot 聊天情境

附錄

參考文獻

許愷洋（2019）。跨代收視習慣大不同，長者網路收視率大漲。iSURVEY 東方線上。取自 https://www.smartm.com.tw/article/36303632cea3

Darren Freeman, Client Officer, Ipsos in Taiwan。台灣人首選的影音平台是什麼?。Ipsos 益普索市場研究。取自 https://www.ipsos.com/sites/default/files/ct/publication/documents/2019-10/n108_bht_taiwanese_turn_to_youtube_for_online_video_content.pdf

鄭伊庭（2017）。2018 影音行銷怎麼做？先記住 YouTube、FB、IG「5、1、10」準則！。SmartM。取自 https：//group.dailyview.tw/article/detail/142

高亞得（2019）。利用機器學習分析中文評論面向之應用 - 以 YouTube 影片評論為例。大同大學資訊經營研究所碩士論文，未出版，台北市。

玩轉 Fasttext。ITREAD 01。取自 https://www.itread01.com/content/1542089709.html

KD Chang（2017-06-03）Python Web Flask 實戰開發教學 - 簡介與環境建置 取自 https://blog.techbridge.cc/2017/06/03/python-web-flask101-tutorial-introduction-and-environment-setup/

量表信度的測量: kappa 統計量之簡介 崔懷芝。 取自 Ims.ctl.cyut.edu.tw › blog › webhd_read_file

網站圖表參考 https://www.highcharts.com

網站 loading gif 參考 https://loading.io

YouTube 爬蟲參考 https://github.com/egbertbouman/youtube-comment-downloader

網站設計 CSS 參考 https://getbootstrap.com