

科技部補助專題研究計畫報告

性別歧視語意情緒分析與智慧諮商系統 (L01)

報告類別：成果報告
計畫類別：個別型計畫
計畫編號：MOST 110-2629-E-259-001-
執行期間：110年08月01日至111年07月31日
執行單位：國立東華大學資訊管理學系

計畫主持人：葉國暉

計畫參與人員：碩士班研究生-兼任助理：林宥均
碩士班研究生-兼任助理：李林發
大專生-兼任助理：楊冠彥

本研究具有政策應用參考價值：否 是，建議提供機關
(勾選「是」者，請列舉建議可提供施政參考之業務主管機關)
本研究具影響公共利益之重大發現：否 是

中華民國 111 年 07 月 20 日

中文摘要：數位科技的進步帶來了無所不在的網路社群環境，大大地延伸了人與人之間的互動，許多訊息在未經驗證下不脛而走，而人們的性別偏見所帶來的影響也跟在社群網路中逐漸被放大，如何抑制在網路社群中的性別偏見與負面情緒擴散逐漸為人們所重視。因此，本研究提出一套性別歧視語意分析與智慧預警系統，該系統將能透過自動化程序針對繁體中文文本進行語意情緒分析，且偵測該文本是否違反性別平等準則，並在使用者情緒較高昂或是違反性平時，給予提醒與警示，以提升其性平意識。本研究透過現存的性平個案文本進行系統有效性之驗證，實驗結果顯示本研究所提出之系統能有效判斷文本之情緒類別與是否違反性平準則。

中文關鍵詞：性別歧視、自然語言處理、BERT、情緒分析

英文摘要：The advancement of digital technology has brought a universal and ubiquitous circumstance in which online social networks have greatly extended the interactions among people. Many messages have been dramatically spread without verification and influence people's life in terms of the perspective of gender bias. Following by the amplification resulted from social networks, how to suppress the spread of gender prejudice and negative emotions in online social networks has promptly become more and more important to people. Therefore, in this study we propose a sentiment analysis and intelligent alert system for gender discrimination. The system is able to perform semantic sentiment analysis on traditional-Chinese-based texts through automated procedures, and to detect whether the text violates gender equality standards as well. Then, the system give reminders and warnings to users with highly negative emotion or with gender bias. The gender awareness has thus been enhanced. The validity of the proposed system is confirmed through existing texts with gender bias. The experimental results show that the proposed system can effectively judge the emotional category of the text and whether it violates the criteria of gender discrimination.

英文關鍵詞：Gender Discrimination, Natural Language Processing, Bidirectional Encoder Representations from Transformers (BERT), Sentiment Analysis

科技部補助專題研究計畫成果報告

(期中進度報告/期末報告)

性別歧視語意情緒分析與智慧諮商系統(L01)

計畫類別：個別型計畫 整合型計畫

計畫編號：MOST 110-2629-E-259-001

執行期間：110年8月1日至111年7月31日

執行機構及系所：國立東華大學資訊管理學系

計畫主持人：葉國暉

計畫參與人員：林宥均、李林發、楊冠彥

(第一年；110/08/01-111/07/31)

本計畫除繳交成果報告外，另含下列出國報告，共0份：

執行國際合作與移地研究心得報告

出席國際學術會議心得報告

中華民國 111 年 7 月 20 日

目錄

目錄.....	I
中文摘要.....	II
Abstract.....	III
一、前言.....	1
二、研究目的.....	3
三、文獻探討.....	3
四、研究方法.....	12
五、實驗設計與結果.....	23
六、結論與未來展望.....	67
七、參考文獻.....	68
八、附錄.....	74

中文摘要

數位科技的進步帶來了無所不在的網路社群環境，大大地延伸了人與人之間的互動，許多訊息在未經驗證下不脛而走，而人們的性別偏見所帶來的影響也跟在社群網路中逐漸被放大，如何抑制在網路社群中的性別偏見與負面情緒擴散逐漸為人們所重視。因此，本研究提出一套性別歧視語意分析與智慧預警系統，該系統將能透過自動化程序針對繁體中文文本進行語意情緒分析，且偵測該文本是否違反性別平等準則，並在使用者情緒較高昂或是違反性平時，給予提醒與警示，以提升其性平意識。本研究透過現存的性平個案文本進行系統有效性之驗證，實驗結果顯示本研究所提出之系統能有效判斷文本之情緒類別與是否違反性平準則。

關鍵字：性別歧視、自然語言處理、BERT、情緒分析

Abstract

The advancement of digital technology has brought a universal and ubiquitous circumstance in which online social networks have greatly extended the interactions among people. Many messages have been dramatically spread without verification and influence people's life in terms of the perspective of gender bias. Following by the amplification resulted from social networks, how to suppress the spread of gender prejudice and negative emotions in online social networks has promptly become more and more important to people. Therefore, in this study we propose a sentiment analysis and intelligent alert system for gender discrimination. The system is able to perform semantic sentiment analysis on traditional-Chinese-based texts through automated procedures, and to detect whether the text violates gender equality standards as well. Then, the system give reminders and warnings to users with highly negative emotion or with gender bias. The gender awareness has thus been enhanced. The validity of the proposed system is confirmed through existing texts with gender bias. The experimental results show that the proposed system can effectively judge the emotional category of the text and whether it violates the criteria of gender discrimination.

Keywords: Gender Discrimination, Natural Language Processing, Bidirectional Encoder Representations from Transformers (BERT), Sentiment Analysis

一、前言

近年來隨著資訊科技的進步，社群媒體的興起為生活帶來了不少改變，人們可以透過不同的社群媒體上面跟其他人互動或是表達自己對於事情的看法，例如：在 Facebook 上面發表跟心情有關的貼文、在有興趣的新聞貼文下面留言表達自己的看法，或是透過 LINE 跟朋友聊天等。

隨著社群媒體的興起，所帶來的負面影響也逐漸顯現，現在大多數的人都是透過社群媒體來了解生活周遭所發生的事，但是網路上的資訊不是真就是假，很多資訊在沒有經過驗證下不脛而走，因此網路社群中的性別偏見與負面情緒也隨著這樣的特性被廣而擴散且加大其影響幅度。然而近幾年較為明顯負面的影響就是假資訊的氾濫及性別偏見的擴散。

近來，2015-2019 年裡以性別偏見作為主要研究的主題正在逐年增長，代表近幾年來性別平等的意識漸漸的越來越受到注重[30]。李敖[6]論及與性別有關的歧視稱為性別歧視，指基於人的生理性別或社會性別而產生的歧視與偏見。Glick 等人[34]於 2000 年進行了一系列的跨國研究，指出性別刻板印象是具有跨國際的普同性。然而在兩性化中，對於不同的性別還是存在著不同的刻板印象(例如:典型的男主外女主內)，而性別刻板印象常會造成某個性別在社會中發展的阻礙，在生活或職場中也會受到對於性別的歧視。

現今社會追求的不再只是兩性的平等，而是更偏向於多元性別的平等。多元性別也稱為性別多元化，不只是單單的異性戀還涵蓋了女同性戀、男同性戀、跨性別和變性人簡稱為 LGBT[22]，近年來多元性別開始逐漸受到大家的關注[14]、[19])，但是被社會的可接受度還是很低，陳禹函[9]發現多元性別在職場會比兩性化性別受到更多的歧視與不公平的對待。

在台灣有一起較為受到大家關注且關於多元性別議題的校園事件為葉永鈺事件，他因與眾不同的性別氣質而遭到部分同學霸凌，最後在某一天下課前去廁所

後沒返回教室，被發現在廁所時已倒臥在血泊中，經搶救後隔天宣告不治。這件事情會受到大家的關注是因為他的死因在各方證詞與未經證實的猜測下有不同版本，部分人士認為此事肇因於學校未重視他的性別特質，也沒有教授多元性別教育有關，所以經過這件事情教育部將兩性平等教育委員會更名為性別平等教育委員會，教育政策從兩性教育延伸轉化成為多元性別教育，也讓教育部開始重視校園性別問題[23]。

從上述的文獻中可以看出不管是在職場或是校園中其實性別歧視都是存在的，甚至可能嚴重的充斥在我們所沒看到的生活，所以我們希望透過利用現有的資訊技術來改善校園中的性別歧視與偏見，讓校園的性別平等意識提升。

在面對校園中發生性別平等事件的時候，性別平等委員會便扮演著重要的角色，調查委員的背景與人生觀點會導致案件的處理方式和結果出現不同，然而在性平調查委員的短缺，且各調查委員的經歷背景皆不同的情況下，常常無法在每個事件中都給出一個令大家都滿意的結果。

再者，性別教育平等法(2004)公布後，學校開始落實性別平等教育，但在賴立淇[19]裡提到在研究有效樣本中，有 96%的研究對象至少經歷過一次以上不同形式的性別歧視，從此研究中可以發現儘管落實性別平等教育，可看出目前學生的性平意識與法規知識仍有所不足，加上社會上對性態度逐漸開放，使得部份行為人無從知悉對自身行為的嚴重性，雖然有些人認為只是開個玩笑，但是對於行為人或是被行為人所帶來的傷害可能是一輩子無法抹滅的創傷與回憶。

而數位科技的進步帶來了無所不在的網路社群環境，大大地延伸了人與人之間的互動，而人們的性別偏見所帶來的影響也跟著在社群網路中逐步放大，因此該如何抑制在網路社群中的性別偏見與負面情緒擴散，逐漸為人們所重視以及該如何預防性平案例的發生，便成為一個重要的課題。本研究欲建立一個自動化偵測性別案例情緒的系統，希望能夠在平常聊天中透過提醒的方式讓使用者可以逐漸增加性別平等的意識。

二、研究目的

載舟之水，亦可覆舟。數位科技的進步帶來了社群網路，提升人與人之間的情緒互動，卻也擴大傳遞了網路社群中的性別偏見與負面情緒。因此本研究的目標為預防並抑制網路社群中的性別偏見與負面情緒之擴散與影響，利用資訊技術提升無所不在社群網路環境下的性別平等意識與友善度。

建置一個社群網路之性別歧視語意自動偵測系統，主要為處理社群網路中以中文文本為主的疑似性別平等事件，希望能在事件發生初期即防止事件的惡化與擴散。系統主要針對文本進行情緒的語意分析和是否違反性平的偵測，會在判斷出情緒較高昂或是違反性平時給予使用者提醒並警示，以提升其性平意識。

三、文獻探討

論文的主旨在於利用資訊技術提升處於現今網路社群下的使用者知性平意識，那近幾年社群網路的快速發展其實也衍生出了不好負面的問題，像是社群網路上詐騙事件頻傳，由於社群網路具有匿名性，網路犯罪往往成為棘手的問題，恰好人們在網路上常留下許多資料，因此可以藉由這些資料做特徵分析，用以解決文本合法與來源識別性的問題。而要利用現有的資訊科技讓電腦理解並處理這些語言及文本則必須使用自然語言處理技術。

1. 自然語言處理

自然語言處理(Natural Language Processing, NLP)是人工智慧和語言學領域的分支，探討如何處理自然語言。什麼是自然語言，自然語言指的是人與人溝通時自然地發展出來的語言，例如：中文、英文等語言。而自然語言處理則是希望透過利用自然語言的方式讓電腦能夠理解自然語言並分析。

- 方法

- 基於規則的方式:

這是自然語言處理最一開始使用的方式，是以人工的方式先將全部的規則制定出來再讓電腦依照這些規則去學習並判斷。這種方式雖然是最簡單明瞭但是它需要的人力資源較高也較消耗時間，而且在人工判斷規則的時候會因為不同的人他們各自對於某些規則的理解能力不同而產生不同的判斷，可能會間接地導致訓練出來的結果不一樣。

➤ 基於機器學習的方式:

不再用程式語言命令電腦所有規則，而是建立演算法模型，讓電腦學會從訓練的資料中，尋找資料中所含的特徵。例如:支援向量機(Support Vector Machine, SVM)、最近鄰居法(K-Nearest Neighbors, KNN)、決策樹等。

➤ 基於深度學習(神經網路)的方式:

神經網路的出現讓自然語言處理的發展更進一步，因為在分析文本相關的資料時，除了分析字詞之外上下文的內容其實也很重要，上下文的不同會連帶影響字詞的意思。利用雙向的循環神經網路(Recurrent Neural Networks, RNN)可以得到上下文內容的意思，但是很容易梯度爆炸跟梯度消失，所以之後就被雙向的長短期記憶模型(Long Short-Term Memory, LSTM)所取代，而 LSTM 也解決了 RNN 所存在的問題。卷積神經網路的一個優點是具有更好的並行，卷積操作中每個時間步的狀態只依賴於局部上下文，而不是循環神經網路中那樣依賴於所有過去的狀態。卷積神經網路可以使用更大的捲積層涵蓋更廣泛的上下文內容。後來注意力機制(Attention)的出現，他可以透過權重的計算去找到輸入的哪些部份和特定的輸出有關。而自注意力機制(Self-Attention)則能獲取更多的上下文訊息。直到近幾年出現預訓練的語言模型，例如:ELMO、BERT。

● 應用

➤ 社群文本識別

近幾年，社群網路上詐騙事件頻傳，由於社群網路具有匿名性，網路犯罪往往成為棘手的問題，恰好人們在網路上常留下許多文字資料，因此可以藉由這些資料做特徵分析，用以解決文本合法與來源識別性的問題。首先，Abbasi 等人

[24]為了解決網路文本匿名問題提出了利用詞法、句法、結構和內容特異性做為分析識別的特徵，並採用支援向量機(SVM)與主成分分析(Principal Components Analysis)等作為檢測方法，其結果準確率達到 94%，接著，Brocardo 等人[28]提出作者身份鑑別研究分為三個不同的領域，作者身份識別、作者特徵描述、作者驗證，而作者身份識別和作者特徵描述已有許多研究，但在作者驗證的方面研究較少，為了解決作者驗證問題提出以詞法、句法做為特徵以及從 n 元語法(n-gram)分析中提取新特徵，並以資訊獲得(Information Gain)做為特徵選擇方法且採用支援向量機做為檢測方法，其檢測結果由樣本大小不同等誤差率(Equal Error Rate)從 9.98%變化到 21.45%。

此外，我們若將以作者寫作型態作為特徵分析發展方向，會發現每個作者的語言使用方式都有其獨特的句法詞語，且表達方式也有所不同，在相同的語言學習環境下，人們各自發展出獨特的寫作風格，從這些不同的寫作風格可以觀察及測量其特徵並進一步量化，再藉由這些特徵進行文本與情感分析。如 Zheng 等人[53]為了解決網路上留言訊息其作者身份的匿名性，提出了以詞語、句法、結構和特定內容做為特徵，並採取以決策樹、反向傳播神經網路(Backpropagation neural networks)和支援向量機為檢測方法，其驗證結果達到 70%到 95%的辨別率，再者，Barbon 等人[27]為了解決社群網站上惡意內容傳播，提出以文本資料驗證帳戶，以 Twitter 上的推文作為研究資料，並透過 N-gram 擷取特徵和最近鄰居法(KNN)辨識用戶，以及避免流行主題影響作者寫作風格，透過更新使用者內容來減少因應作者文章內容變化，其準確率達到 93%。目前大多文獻都以英文文本做研究，其他語言的研究較少，Albadarneh 等人[26]為了解決阿拉伯語言方面作者驗證問題，以阿拉伯語推特作為文本資料，提出以詞袋模型(Bag-of-words model)及 TFIDF 來計算特徵向量，接著將資料放置 Hadoop 分部是框架上進行樸素貝葉斯分類器(Naïve Bayes classifier)，其準確率為 61.6%。

在 2017，Vosoughi 等人[51]提出了一組謠言測量模型，用以偵測線上謠言，其主要利用三個語言特徵、使用者識別跟動態傳播(Propagation Dynamics)等共 17 個特徵來進行謠言的分析與預測，在方法下，謠言判定的準確率可以達到 75%。接著，Kesarwani 等人 [37]在 2020 年探討了如何利用 K-NN(K-Nearest Neighbor)演算法來進行線上假訊息的偵測，該方法可以達到 79%的判定準確率。Pandya[42] 則提出了一個結合 k-means 與 SVM 來提升垃圾郵件(Spam)的判定率，首先利用 k-means 演算法來移除資料庫中的極端值，然後，再將經過前處理的資料庫資料進行 SVM 分類，經過測試後，該方法可以達到 98%的高準確率。再者，Mansouri 等人[40]也在 2020 年間發展了一套機制，結合了 CNN(Convolutional Neural Network)演算法與半監督式的線性區別分析 (Semi-supervised Linear Discriminant Analysis；SLDA)演算法，用以偵測線上傳播的偽造或錯誤訊息，CNN 演算法是用來引導特徵，而 SLDA 演算法則用於資料分類，最終，判斷準確率也被提升到了 95%。最後，Freire 等人[32]考慮了群眾訊號與回饋，用以協助判定假訊息，主要為利用群眾對該訊息的反應與可能之傳播行為，來判定該訊息是否為疑似偽造之訊息。

➤ 文本情緒分析

黃裕晴[15]認為情緒是可以用來掌握小說流向的一種方式，透過自動化分析工具，從情緒這個切入點來分析小說內各角色或敘事內容的變化，探討劇情脈絡或是寫作風格。鄭建澤[17]利用 BERT 對餐廳的 google 評論進行情感分析，希望透過網站的方式提供使用者餐廳評論的深度見解，期許能過此方式幫使用者加速找到好的餐廳，而不是一一的瀏覽餐廳的評論才決定說是否選擇這家餐廳。Senthil Kumar 等人[49]把文本當成圖像，用卷積神經網路(CNN)中不同大小的卷積核，以便更好的學習文本分類的特徵，利用雙向 LSTM 對特徵做詞性標註，最後透過全連接層來預測文本的情緒類別。

Gao 等人[33]在 論文中所提出的 Target-Dependent BERT，跟一般的 BERT 不同的地方是它不是用 CLS 來做分類，而是用對目標詞做最大池化(Max Pooling)後和 CLS 進行拼接做 OR 運算，主要是因為池化層採用最大池化來保留關鍵特徵，減少特徵的維度，防止模型過度擬合。最後使用全連接層(Fully Connected Layer)來對輸入的句子做情感的分類。而陳震瑜等人[10]提出應用 Aspect-based sentiment analysis(ABSA)，分析社群網站對意見目標(Opinion Target)的情感類別，使用 BERT 做為字詞嵌入層的方法，實驗結果顯示，在 BERT 後面加入 Attention-over-Attention(AOA)注意力模型的效能優於 BERT 後面接 CLS 做情緒分析的效能。而 Xu 等人[52]認為問答系統對於電子商務領域來說很重要，所以提出了 review reading comprehension(RRC)，將問題和評論拼接起來作為輸入，並預測評論中的每個詞作為答案的開始或結束的機率，並對評論進行方面提取 (Aspect Extraction)和方面情感分類(Asspect Sentiment Classification)。

表一、自然語言處理文獻彙整

基於什麼方法	年分	作者	提出的方法	結果
基於機器學習	2017	Vosoughi 等人 [51]	提出了一組謠言測量模型， 用以偵測線上謠言	準確率 75%
基於機器學習	2020	Kesarwani 等人 [37]	利用 K-NN 來進行線上假訊 息的偵測	準確率 79%
基於深度學習 (神經網路)	2020	Mansouri 等人 [40]	利用 CNN+SLDA 用以偵測 線上傳播的偽造或錯誤訊息	準確率 95%
基於深度學習 (神經網路)	2020	Senthil Kumar 等人[49]	利用 CNN+雙向 LSTM 對文 本進行情緒分類	準確率 98.6%

2. 性別平等

性別平等是《世界人權宣言》的目標之一，其目的是營造性別平等的法律和社會環境。具體實踐中，性別平等的目標是使兩性在整個社會生活中平等對待，

而不只是在政治、工作或其他政策規定的領域[50]。而現今的性別平等不再只是兩性的平等而是對多元性別者的性別平等。

➤ 性別平等教育法

性別平等教育法在 2004 年公布，此法的立法目的：「促進性別地位之實質平等，消除性別歧視，維護人格尊嚴，厚植並建立性別平等之教育資源與環境」[21]。所以學校開始透過不論是正式課程還是非正式課程來讓學生增強對性別平等的觀念，但實際上透過這些課程學生就會對於性別平等的觀念都很了解嗎？

在曾佳雯等人[13]為了了解性別平等法實施之後在校園裡推動性別平等教育的成效如何，所以製作了問卷並以高雄市國小高年級的學生為研究對象進行調查，在研究結果中指出研究對象以接受性別平等教育的經驗皆為正向，然而從非正式課程中學到跟性別平等相關的知識與經驗和正式課程及潛在課程相比之下來的高，但學習成效會因其性別、主要照顧者等的不同而有所差異。

而何政光等人[8]利用情境式教學的方式來提升護生對性別平等認知的觀念，研究裡透過演練的方式來學習適當處理的方式、體驗不同性別的特質(例如:讓男同學體驗孕婦裝)和透過電影來一起討論跟性別平等相關的課題等情境式教學。在接受情境式教學的前後都利用作者設計的性別平等知識因應能力評量表對受測者進行前後測，研究結果顯示後測的平均分數較前測來得高(分數越高表示性別平等的知識越正確)，所以說明可以透過情境式教學的方式來讓學生增強對於性別平等的觀念。

➤ 性別刻板印象、性別歧視、性騷擾

性別角色是社會或文化期待人們表現適合於男性或女性的行為模式[1]，這些行為包括言行舉止、服裝、觀念等，而性別角色的影響連帶延伸出性別刻板印象和性別歧視，對於性別刻板印象最經典的例子就是傳統的「男主外，女主內」，從小被灌輸著男生要出門賺錢女生要在家做家務，而隨著時間人們的觀念開始慢慢改變，漸漸打破了「男主外，女主內」這種觀念，而朱蘭慧[2]提到談兩性，男生角色絕對無法缺席，兩性教育不是只有為了女性努力，所以作者採取生命史

研究法來探討研究對象在其生命歷程中影響其性別刻板印象形成的原因，從研究的三個對象中明顯的發現父母的教養、城鄉差距、學校課程、老師及同儕等都是影響男性性別刻板印象形成的原因。

杜宜展[7]編製了性別角色刻板印象量表對花蓮縣、台南縣、高雄縣為範圍中各選一所國立大學及國立和私立學院，對其大學生發放量表，研究結果中的背景變項結果指出不同學校、性別和年齡在性別角色刻板印象方面有顯著相關的存在，而和不同年級較無相關。李傲[6]提到與性別有關的歧視稱為性別歧視，指基於人的生理性別或社會性別而產生的歧視與偏見。對於性別歧視如果只有內心的想法並不會造成違法，但當性別歧視的觀念備付諸實施時，產生具體的傷害後(例如:因為性別歧視而在工作上受到不合理的對待)，即構成違法。賴立淇[19]探討大專院校場域的學生對於性別歧視知覺的差異以及不同性別角色取向其心理健康的差異，透過線上問卷的方式發現在有效樣本中有 96%的研究對象至少經歷過一次以上不同形式的性別歧視，研究結果顯示性別角色特質對性別歧視知覺與心理健康不具有調節效果。

陳秀峯[12]作者覺得性騷擾事件隨時隨地發生於社會上任何角落，但是發生在職場中的案例很多都存在著權力不平等的關係，所以透過課程的方式讓學生閱讀相關文獻及法規並以情境案例為教材透過如紀錄片、新聞報導等來讓學生了解這些議題訊息，並由老師來引導出問題並討論，希望透過這些方式讓學生學習到如果遭遇到性騷擾問題時該如何應對。汪淑娟等人[3]對中部某大學的大學生用問卷進行調查，結果顯示不分性別，大學生對性騷擾的認知仍然迷思，在性別平等的觀念具有性別上的差異，女生比男生具有較高的自我察覺。

1. 多元性別

近年來社會開始逐漸重視多元性別的權益，多元性別又稱為性別多元化，用來代表人類的社會性別特徵，目前較常見的為 LGBT，LGBT 是女同性戀者 (Lesbian)、男同性戀者 (Gay)、雙性戀者 (Bisexual) 與跨性別者

(Transgender) 的英文首字母縮略字，也可廣泛代表所有非異性戀者[22]。儘管台灣在 2019 年時將同性婚姻合法化，但還是有許多人無法接受同性結婚甚至對 LGBT 族群抱持著反感甚至排斥跟他們接觸，所以在生活中 LGBT 族群還是會莫名的遭受到許多不公平的對待。衛漢庭等人[18]對於希望能夠改善多元性別者在醫療場合中被歧視和不公平的對待，所以提出了從政策、空間、臨床實踐、教育這四個層面去解決這個問題，希望透過政策對多元性別友善的醫院做評鑑、提供多元性別友善的醫療場域和問診及讓醫護人員落實多元性別友善的教育。張曉佩等人[14]發現儘管多數人都知道多元性別，但被接受度還是很低，導致多元性別的人會比其他受到更多言語上面的謾罵，也發現許多心理諮商師對於多元性別不太了解可能無法給於較大的幫助，所以作者利用課程實驗的方式，讓課程裡的參與者透過課程能夠更加了解多元性別，實驗結果也顯示較多的人透過課程更加了解後，反而從不喜歡或厭惡轉成接受和不排斥。

可以從上述的文獻中看出其實大多數的人對於性別平等這件事情是知情但並沒有太過深入的了解，導致可能在不知情的情況下觸犯性平相關的法規，而也可能在遇到不公平的性別對待時並不清楚該如何保護自己的權利。但也可從文獻中看出可以透過課程或是不同的方法來輔助人們增強性平方面的意識。

表二、性別平等文獻彙整

年分	作者	提出的方法	結果
2007	杜宜展[7]	對大學生發放自製的量表	不同學校、性別和年齡在性別角色刻板印象方面有顯著的影響
2009	何政光等人[8]	情境式教學(例如:體驗不同性別的特質)	研究結果顯示後測的平均分數比前測高
2009	汪淑娟等人[3]	對某所大學的大學生進行問卷調查	結果顯示不分性別，大學生對性騷擾的認知仍然迷思，在性別平等的觀念具有性別上的差異，女生比男生具有較高的自我察覺。

2019	賴立淇[19]	透過線上問卷對大學生進行研究	在有效樣本中發現有 96% 的研究對象至少經歷過一次以上不同形式的性別歧視。
------	---------	----------------	--

3. 利用自然語言處理技術來分析性別平等之研究

Barbon 等人[30] NLP 和性別偏見的相互作用是雙重的。一方面，NLP 可以成為檢測不同社會環境中性別偏見的工具，例如在在線新聞或廣告中。另一方面，NLP 經常產生性別偏見系統，從而在社會中延續和放大性別偏見。Saha 等人[50] 對 304 篇關於自然語言處理中的性別偏見的論文進行了調查。作者發現關於性別偏見的研究存在四個核心限制。1) 大多數研究將性別視為二元變量。2) 大部分研究都是在英文或其他資源豐富的語言設置中進行的。3) 儘管有無數關於 NLP 方法中的性別偏見的論文，發現大多數新開發的演算法都沒有測試他們的模型是否存在偏見。

Li 等人[41]因為認為不同性別的身分和引用差異不僅對弱勢性別而且對整個研究領域都可能產生重大不利影響。所以透過研究觀察在自然語言處理領域的研究中，引用的部分作者排序的順序是否因為性別而受到不同對待。作者也希望透過此研究能提高人們對性別差距的認識，並提高研究的包容性和公平性，不只要重視男性與女性的差異，也應該重視跨性別者和不同種族的不同。EXIST 2021[54]是 IberLEF 2021 上關於社群網路中性別歧視的第一個共享任務。這個共享任務分為兩個子任務，資料集則為英文和西班牙文的文本。任務一是二元分類，判斷文本是否為性別歧視，任務二則是將被判定為性別歧視的文本再做細分為意識形態與不平等(Stereotyping and Dominance)、性暴力(Sexual Violence)等五類。而 Albadarneh 等人[29]利用數據增強來讓模型達到更好的效果。使用的是回譯，將英文和西班牙文文本翻譯成德文後再翻譯回英文和西班牙文，藉此增加資料量。使用多個分類器測試(例如:邏輯式回歸、支持向量機、長短期記憶、BERT 等)，最後顯示的成果為 BERT 效果較佳，跟其他分類器相比準確度較高。Rodríguez-Sánchez 等人[47]使用 MeTwo 資料集[46]來對數據做增強，使用 mBERT、XLM-

RoBERT 和 XLM-Twitter 三個模型進行微調，將模型改成單一任務模型和多任務模型，來比較兩者對於兩個子任務的準確度哪個較高，從結果可看出多任務模型運用在任務二上的結果較佳。

表三、文獻[29]使用方法之實驗結果彙整

Butt 等人[29]	任務一	任務二
Multilayer perceptron (MLP)	89.38	67.54
Bidirectional Encoder Representations from Transformers(BERT)	81.32	63.31
Support Vector Machine(SVM)	67.67	70.71

四、研究方法

1. BERT(Bidirectional Encoder Representations from Transformers)

這邊先介紹系統中有使用到的 BERT 架構，BERT 全名為(Bidirectional Encoder Representations from Transformers)，2018 年由 Google 提出。是由多個 Transformer 中 encoder 的部分堆疊而成的架構，每個 Transformer 的 encoder 裡面都包含兩個部分，一個自注意力層(self-attention layer)和一個前饋層(feed-forward layer)，而目前有兩種不同的 BERT 模型:BERT-base 和 BERT-large，如下。

- BERT-Base : 12-layer, 768-hidden, 12-heads, 110M parameters。
- BERT-Large : 24-layer, 1024-hidden, 16-heads, 340M parameters。

而 BERT 被視為近幾年來的一個強大的語言模型有兩個原因，一個原因為它使用從 BooksCorpus（8 億字）和 Wikipedia（25 億字）中提取的未標記數據進行預訓練。另一個則是因為 BERT 不僅從左到右，也從右到左從單詞序列中學習訊息，而且會根據上下文表達的單詞訊息不同而給單詞不同的詞嵌入。

BERT 的輸入可以為一個或兩個句子和文章，而一個輸入的最長長度為 512 個字，超過 512 個字的話要進行分段。在 BERT 中他會將輸入的句子依照語言的不同去做切割，利用 WordPieces 將英文的部分會切成詞，中文的部分會切成字，

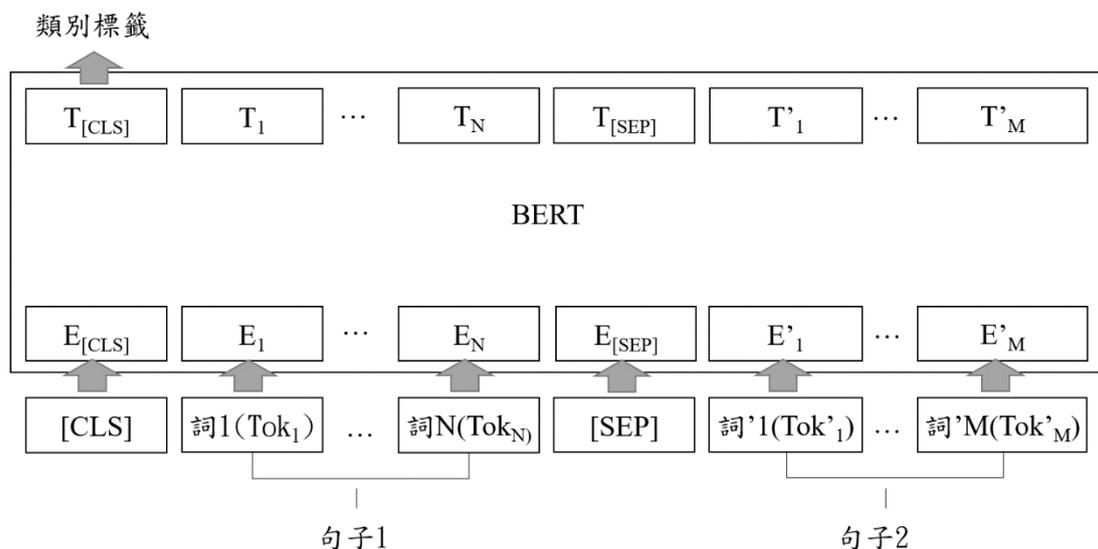
然後將每個詞或字分別轉成詞嵌入(Token Embeddings)、段落嵌入(Segment Embeddings)和位置嵌入(Position Embeddings)，再將三個向量加總起來變成一個新的向量輸入進去 BERT 模型中進行運算，如圖一。

輸入 (Input)	[CLS]	今	天	天	氣	很	好	[SEP]
詞嵌入 (Token Embeddings)	$E_{[CLS]}$	$E_{今}$	$E_{天}$	$E_{天}$	$E_{氣}$	$E_{很}$	$E_{好}$	$E_{[SEP]}$
段落嵌入 (Segment Embeddings)	E_A	E_A	E_A	E_A	E_A	E_A	E_A	E_A
位置嵌入 (Position Embeddings)	E_0	E_1	E_2	E_3	E_4	E_5	E_6	E_7

圖一、BERT 的輸入

(資料來源: Bert: Pre-training of deep bidirectional transformers for language understanding)[27]

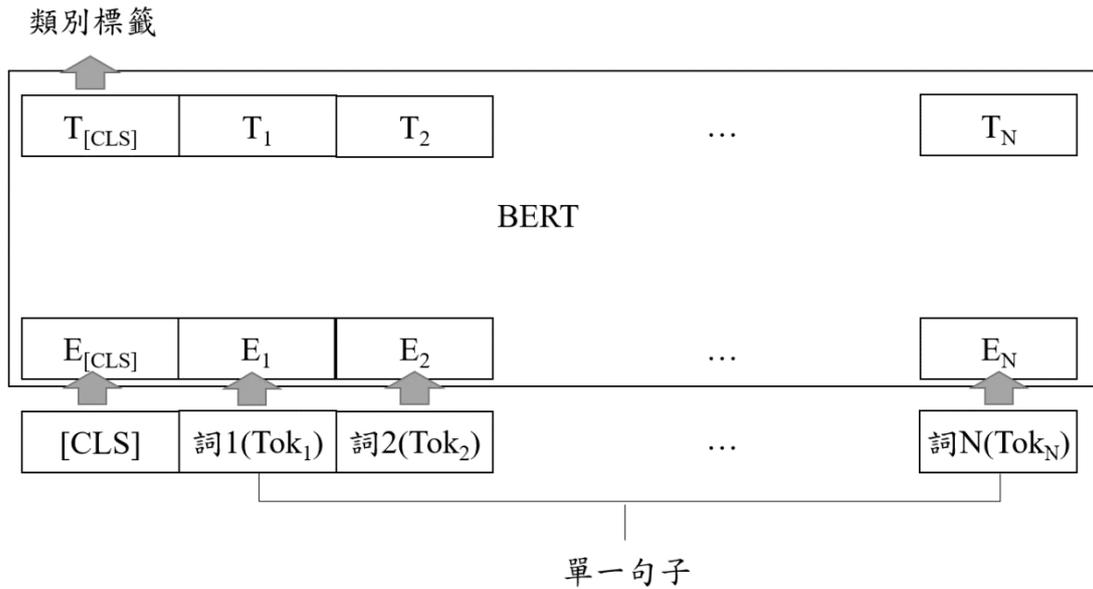
圖二至圖四為 BERT 中較常見的四個應用:單一句子、成對句子分類、單一句子標註和問答任務。利用 BERT 對文本進行分類時，在最後輸出文本的類別時，會抓取[CLS]這個 Token 的值作為分類的結果，選擇[CLS]的原因是因為和文本中的其他字詞相比，這個沒有明顯語意的 Token 會更公平的融合文本中各個詞的語意訊息。



圖二、BERT 常見的四個任務(1):

句子對分類任務(Sentence Pair Classification Tasks): MPRC、SWAG

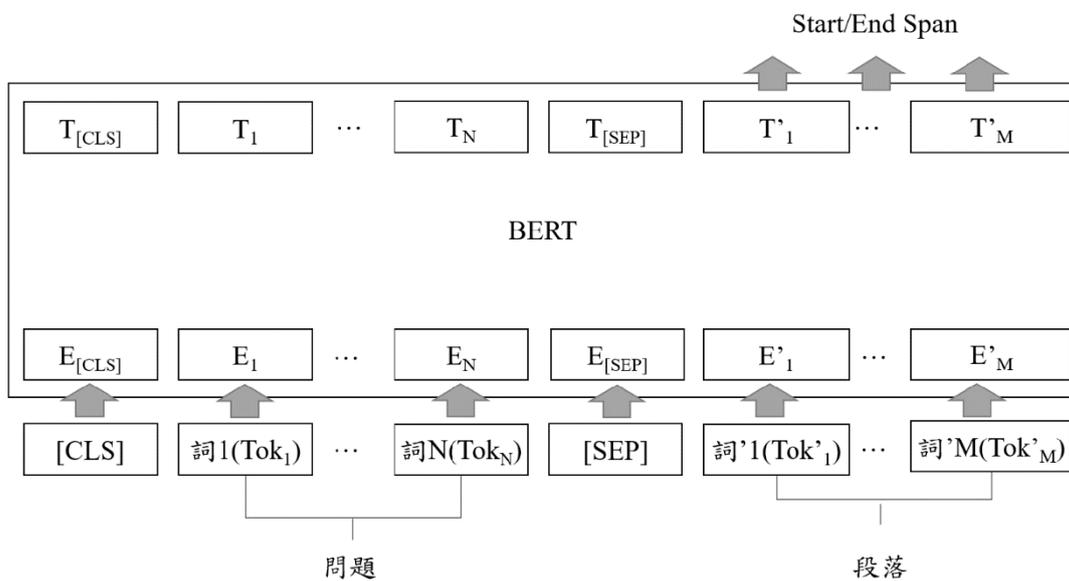
(資料來源: Bert: Pre-training of deep bidirectional transformers for language understanding)[27]



圖三、BERT 常見的四個任務(2):

單一句子分類任務(Single Sentence Classification Tasks): STT-2、CoLA

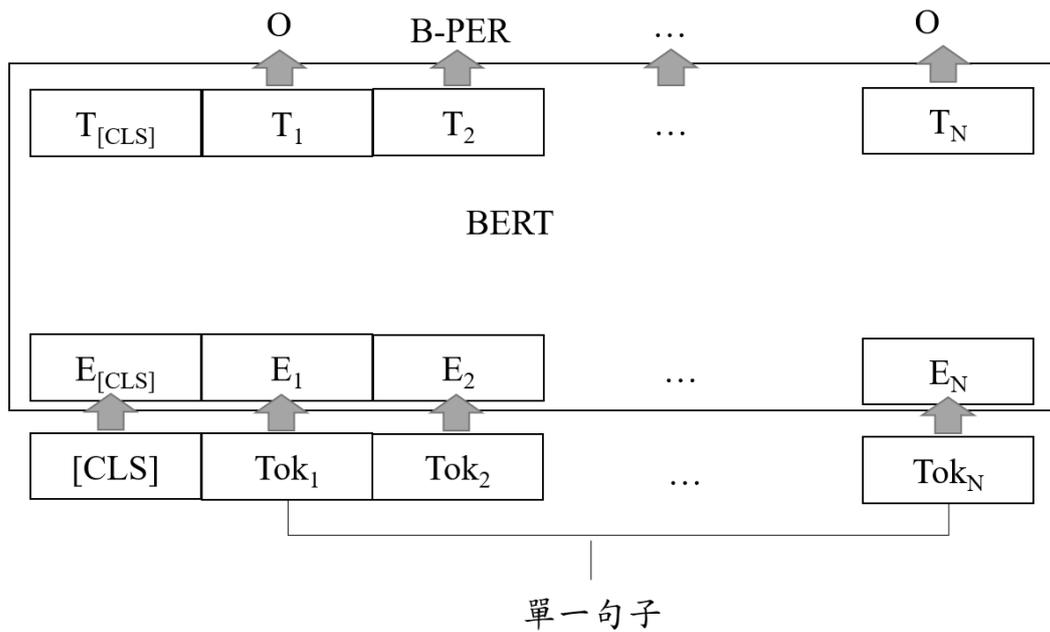
(資料來源: Bert: Pre-training of deep bidirectional transformers for language understanding)[27]



圖四、BERT 常見的四個任務(3):

問答任務(Question Answering Tasks): SQuAD v1.1

(資料來源: Bert: Pre-training of deep bidirectional transformers for language understanding)[27]



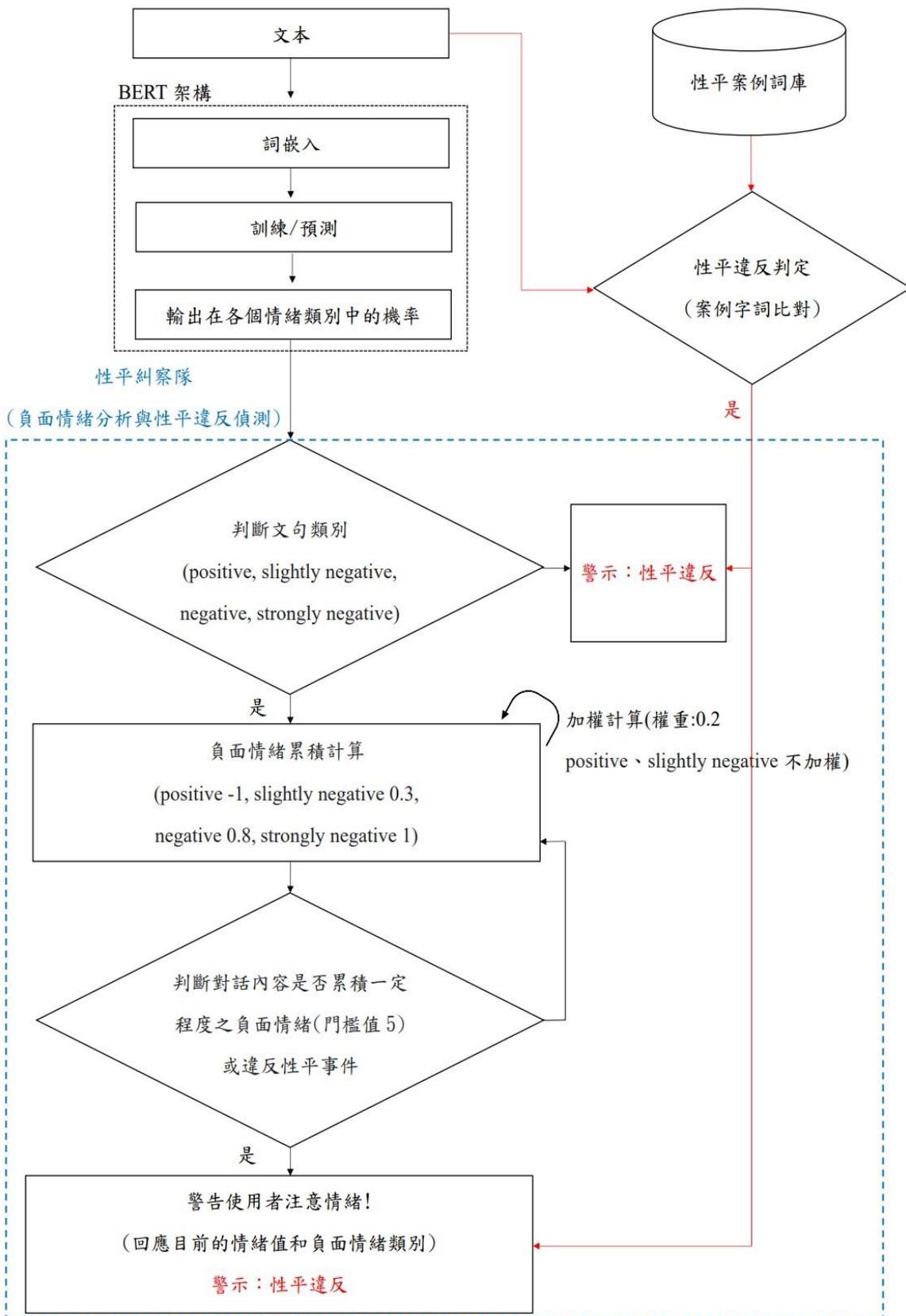
圖五、BERT 常見的四個任務(4):

單一句子標記任務(Single Sentence Tagging Tasks): CoNLL-2003 NER

(資料來源: Bert: Pre-training of deep bidirectional transformers for language understanding)[27]

2. 性別歧視語意分析與智慧預警系統

圖六是我們所提出的系統架構圖，從輸入文本到聊天機器人給出回應會經過兩個架構，分別是 BERT 架構(判別文本情緒類別)和性平糾察隊(負面情緒分析與性平違反偵測)。



圖六、性別歧視語意分析與智慧預警系統

- 文本

系統所輸入的文本為繁體中文的文章或是字句。

- BERT 架構

本研究所使用的是 BERT-Base-Chinese 模型，利用 BERT 判別所輸入文本的情緒類別，如 3-1 節所介紹的 BERT 常見的四大任務，系統所使用的是 BERT 中的單一句子分類任務，輸入為一篇文章或是一個句子，當輸入的文本進入 BERT 架構後會先利用 WordPieces 將中文文本切成字然後對每個字做詞嵌入，並透過已經訓練過的 BERT 模型對文本進行情緒類別的分析，去預測輸入的文本在各個情緒類別中的機率，並透過抓取機率值最大的類別做為文本的情緒並輸出。

- 性平糾察隊

當預測出文本的情緒後，性平糾察隊會去抓取預測的文本情緒類別後進行負面情緒的分析與性平違反的偵測。

在負面情緒的分析中第一步進行文本情緒類別的判斷，將抓取到的情緒依照不同等級分類到不同類別(例如:快樂(happy)這個類別會被分類到正面(positive))，第二步開始進行文本負面情緒的分數計算，每個類別的分數都不相同，出現正面(positive)扣 1 分，弱負面(slightly negative)加 0.3，這兩個類別的情緒不進行加權。而負面(negative)加 0.8 分、強負面(strongly negative)加 1 分，兩者權重皆為 0.2，權重要連續判斷為同個情緒才會進行疊加(例如:連續出現兩次負面(negative)時第一次加 0.8 分而第二次會加 1 分以此類推)。當負面情緒值累積超過門檻值 5 分會跳出通知提醒使用者注意情緒!並回應目前的情緒值和目前的負面情緒類別。

性平違反偵測則是將輸入的文本和性平案例詞庫去做字詞的比對，如果有比對到詞庫中的字詞就會判定違反性平，並不會受到情緒類別及情緒累積分數的影響，會立即跳出通知警示:性平違反，提醒使用者已經違反性平藉此提升使用者的性平意識。

- 判斷文句類別

當 BERT 輸出預測文本之情緒類別後，會依照表四中所對應到的情緒類別進

行轉換。

表四、文句類別情緒表

文句情緒類別	BERT 輸出之情緒類別
正面(positive)	快樂(happy)
弱負面(slightly negative)	難過(sad)
負面(negative)	生氣(angry) 諷刺(irony)
強負面(strongly negative)	侮辱(insult)

➤ 性平糾察隊之各情緒值計算

本系統會針對不同情緒對情緒值進行增減，不同情緒會有不同的加權值，當如果連續遇到兩個判定為一樣情緒類別時，情緒值會依照當前情緒類別的權重進行疊加。

表五、各情緒類別之情緒值

情緒類別	情緒初始值	動態權重
正面(positive)	-1	0
弱負面(slightly negative)	0.3	0
負面(negative)	0.8	0.2
強負面(strongly negative)	1	0.2

當連續遇到兩個判定為一樣情緒類別時，情緒值會依照當前情緒類別的權重進行疊加，而 positive(happy)的初始值為-1 以動態權重為 0 上升。



圖 七、positive(happy)的情緒值計算

當連續遇到兩個判定為一樣情緒類別時，情緒值會依照當前情緒類別的權重進行疊加，而 slightly negative(sad)的初始值為 0.3 以動態權重 0 上升。



圖 八、slightly negative(sad)情緒值計算

當連續遇到兩個判定為一樣情緒類別時，情緒值會依照當前情緒類別的權重進行疊加，而 negative(angry)的初始值為 0.8 以動態權重 0.2 上升。



圖 九、negative(angry)情緒值計算

當連續遇到兩個判定為一樣情緒類別時，情緒值會依照當前情緒類別的權重進行疊加，而 negative(irony)的初始值為 0.8 以動態權重 0.2 上升。



圖 十、negative(irony)情緒值計算

當連續遇到兩個判定為一樣情緒類別時，情緒值會依照當前情緒類別的權重進行疊加，而 strongly negative(insult)的初始值為 1 以動態權重 0.2 上升。



圖 十一、strongly negative(insult)情緒值計算

➤ 系統之提醒與警示

而圖十二為在對話中情緒值累積超過預設的門檻值 5 的時候，性平糾察隊會跳出提醒告知使用者要注意情緒，直到情緒降到門檻值以下才會停止提醒。

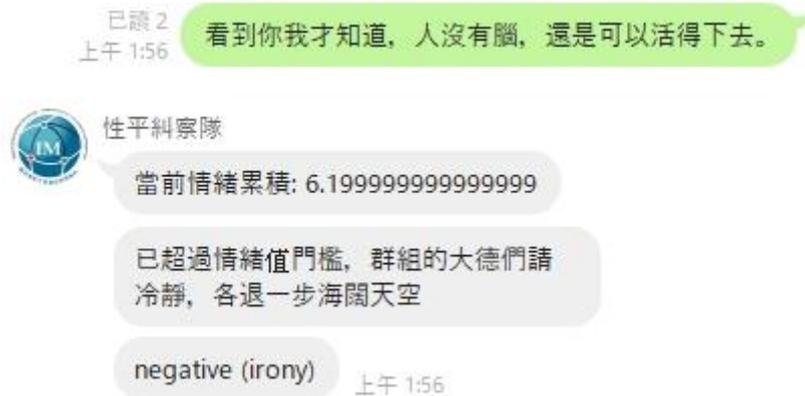


圖 十二、超過情緒門檻值之提醒

圖十三為當文本跟性平案例詞典中的字詞比對到時，會立即跳出違反性平！用來警示使用者已經觸犯性平藉此提升使用者的性別平等意識。



圖十三、違反性平之警示

➤ 性平案例詞庫

表六為系統中所使用的性平案例詞庫中的部分關鍵詞，詞庫中的字詞皆是從性平個案中和法院判決書中所抓取並彙整而成，完整的詞庫請參考附錄 A。

表六、性平案例詞庫

婊子	娘娘腔	賤人	臭賤貨
死台女	綠茶婊	娘砲	媽寶
婊子配狗	小白臉	破麻	死三八
死 gay	不男不女	死賤人	賤貨

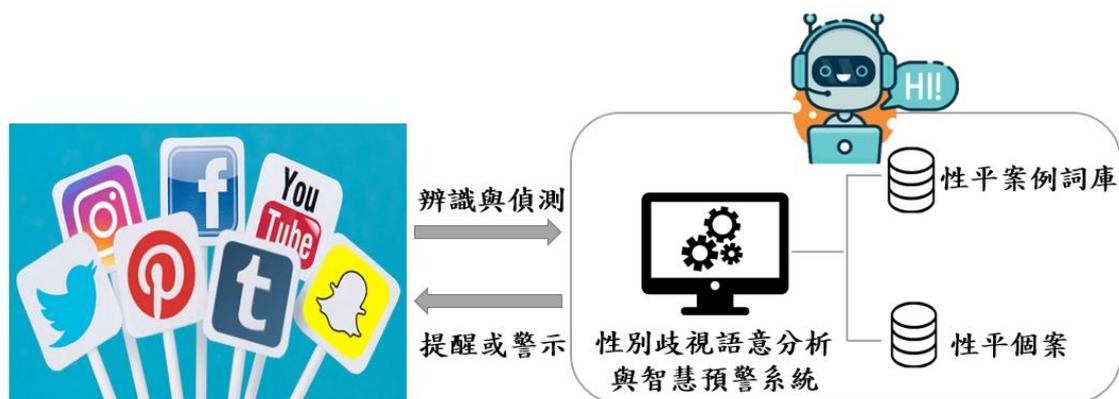
➤ 性平違反判定

系統對於性平違反的判定，所使用的是字詞的比對，只要文本中的字詞跟性平案例詞庫中的字詞比對到時，不會因為當下的情緒類或是情緒累積值是否達到門檻值，則會立即警示使用者違反性平，如圖十三。

五、實驗設計與結果

1. 實驗設計

建置一個系統，可以辨識與偵測社群媒體上的資訊並適時給予使用者提醒或警示。



圖十四、系統之實驗設計圖

2. 實驗環境

● 硬體環境

表七列出了系統整個環境中的硬體規格。

表七、硬體環境表

作業系統(OS)	Windows 10 64-Bit
處理器(Processor)	RTX 3060 Ti /AMD-5600x
記憶體(Memory)	Ram 32gb

● 軟體環境

表八列出了系統整個環境的軟體規格，包含 Python、Tensorflow、BERT、BOT 所使用的版本。

表八、軟體環境表

Python	3.8
Tensorflow	1.11.0
BERT	BERT-Base, Chinese (12-layer,768-hidden,12-heads,110M parameters)
BOT	LINE BOT

3. 系統驗證

在測試畫面中會顯示實際執行時所沒有的當前的情緒累積值跟文本的類別。
一開始會利用「清空」將性平糾察隊中的情緒累積值清零，避免影響後面的測試，
如圖十五。



圖十五、將系統情緒值歸零

- 以文句為基礎之測試結果

圖十六為一連串使用單一句子測試的結果，先把情緒值歸零後開始測試，途中性平糾察隊會因為情緒類別的不同導致情緒累積值的增減，但只要偵測到跟性平相關的字詞無論情緒值累積多少都會立即跳出警示。



圖 十六、單一文句之測試結果

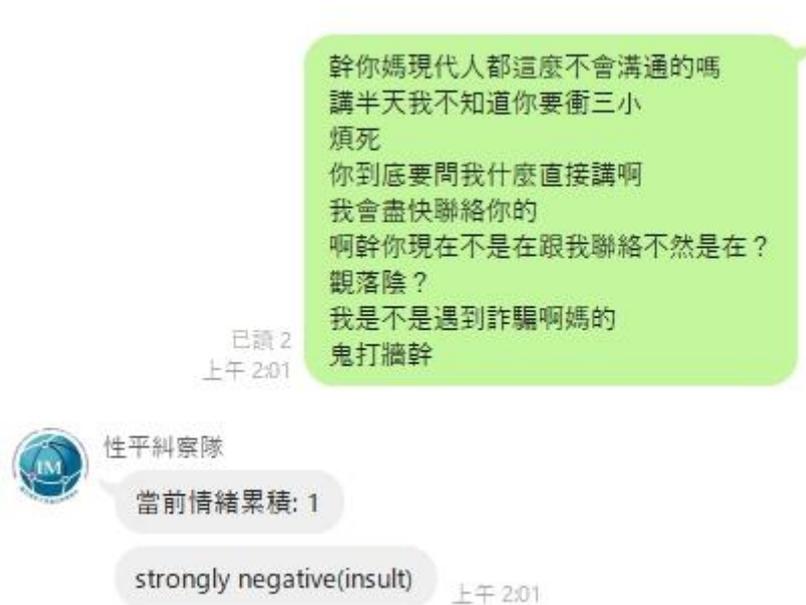
- 以文本為基礎之測試結果
 本小節將以五篇不同情緒的文本對系統進行有效性之認證。
- **文本一之結果**
 (文本一)為使用者尋求客服人員幫忙後，問題沒有被解決，於是憤而在社群網路上發文宣洩不滿情緒之情境。而圖十七為具相同脈絡之聊天會談內容結果，

因為是輸入整個文章，會依照上下文的不同而判定為不同情緒，所以文本一被判定為 strongly negative(insult)，是帶有一些不雅字詞的非常負面且具有侮辱意味之文章。

文本一之輸入：

幹你媽現代人都這麼不會溝通的嗎
講半天我不知道你要衝三小
煩死
你到底要問我什麼直接講啊
我會盡快聯絡你的
啊幹你現在不是在跟我聯絡不然是在？
觀落陰？
我是不是遇到詐騙啊媽的
鬼打牆幹

文本一之輸出:如圖十七



圖十七、文本一之測試結果

• 文本一細部字句分析:

圖十八至圖二十為將文本一拆成單一字句放入性平糾察隊後得到每個句子所表達的情緒，可以從中看出大部分的句子都被判定為非常負面的情緒，推斷當時使用者的情緒較高昂。可以看出這是一篇非常負面且具有侮辱意味的文章，而透過整篇文章的上下文也成功判定這篇文章是 strongly negative(insult)的情緒。



圖 十八、文本一之字句分析(1)

已讀 2
上午 2:02 我會盡快聯絡你的



性平糾察隊

當前情緒累積: 4.0

positive(happy) 上午 2:02

已讀 2
上午 2:02 啊幹你現在不是在跟我聯絡不然是？



性平糾察隊

當前情緒累積: 5.0

已超過情緒值門檻，群組的大德們請
冷靜，各退一步海闊天空

strongly negative(insult) 上午 2:02

已讀 2
上午 2:02 觀落陰？



性平糾察隊

當前情緒累積: 6.2

已超過情緒值門檻，群組的大德們請
冷靜，各退一步海闊天空

strongly negative(insult) 上午 2:03

圖 十九、文本一之字句分析(2)

已讀 2
上午 2:03

我是不是遇到詐騙啊媽的



性平糾察隊

當前情緒累積: 7.6000000000000005

已超過情緒值門檻，群組的大德們請
冷靜，各退一步海闊天空

strongly negative(insult) 上午 2:03

已讀 2
上午 2:03

鬼打牆幹



性平糾察隊

當前情緒累積: 9.2000000000000001

已超過情緒值門檻，群組的大德們請
冷靜，各退一步海闊天空

strongly negative(insult) 上午 2:03

圖二十、文本一之字句分析(3)

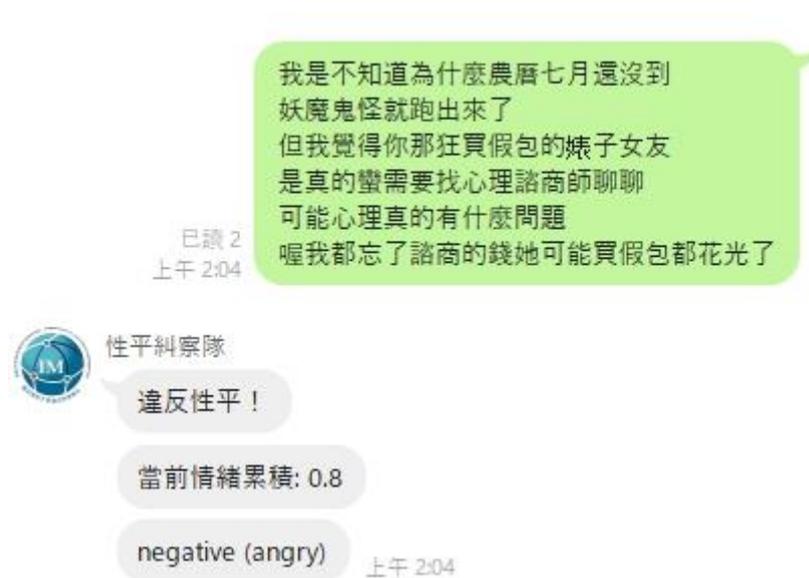
➤ 文本二之結果

(文本二)為法院判決書上所抓取的個案內容，為使用者因懷疑和前任交往期間 A 女介入了他們的感情而心生不滿，因此在社群媒體上公開發文侮罵 A 女，並加油添醋了一些不實的內容，而圖二十一在性平糾察隊中被判定為違反性平且是帶有生氣意味的一篇文章，而法院判決書的結果是以妨礙名譽結案。

文本二之輸入:

我是不知道為什麼農曆七月還沒到
妖魔鬼怪就跑出來了
但我覺得你那狂買假包的婊子女友
是真的蠻需要找心理諮商師聊聊
可能心理真的有什麼問題
喔我都忘了諮商的錢她可能買假包都花光了

文本二之輸出:如圖二十一



圖二十一、文本二之測試結果

• 文本二細部字句分析:

圖二十二和圖二十三為將文本二拆成單一字句放入性平糾察隊後得到每個句子所表達的情緒，可以從中看出七句裡面有六句不是被判定為生氣、諷刺不然就是侮辱並且還違反性平，可以看出這是一篇憤怒值偏高的文章，而透過整篇文章的上下文也成功判定這篇文章是 negative(angry)的情緒。



圖二十二、文本二之字句分析(1)

已讀 2
上午 2:06

可能心理真的有什麼問題



性平糾察隊

當前情緒累積: 4.9

negative (irony)

上午 2:06

已讀 2
上午 2:06

喔我都忘了諮商的錢她可能買假包都花光了



性平糾察隊

當前情緒累積: 6.5

已超過情緒值門檻，群組的大德們請
冷靜，各退一步海闊天空

negative (angry)

上午 2:06

圖二十三、文本二之字句分析(2)

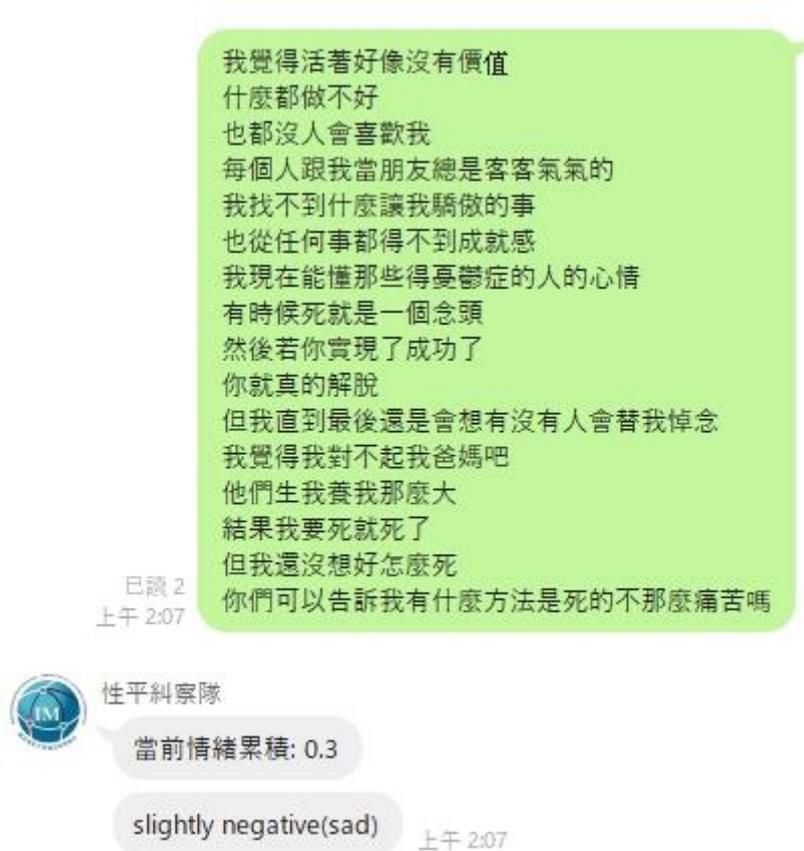
➤ 文本三之結果

(文本三)為使用者因為心情低落，所以在社群媒體上面發文抒發情緒，並希望看過這篇文章的人能夠給予他一些建議之情境。而圖二十四在性平糾察隊中被判定為 slightly negative(sad)，表示這是一篇帶有難過悲傷意味的一篇文章。

文本三之輸入：

我覺得活著好像沒有價值
什麼都做不好
也都沒人會喜歡我
每個人跟我當朋友總是客客氣氣的
我找不到什麼讓我驕傲的事
也從任何事都得不到成就感
我現在能懂那些得憂鬱症的人的心情
有時候死就是一個念頭
然後若你實現了成功了
你就真的解脫
但我直到最後還是會想有沒有人會替我悼念
我覺得我對不起我爸媽吧
他們生我養我那麼大
結果我要死就死了
但我還沒想好怎麼死
你們可以告訴我有什麼方法是死的不那麼痛苦嗎

文本三之輸出:如圖二十四



圖二十四、文本三之測試結果

- 文本三細部字句分析:

圖二十五至圖二十八為將文本三拆成單一字句放入性平糾察隊後得到每個句子所表達的情緒，可以從中看出大部分被判定為難過、生氣與諷刺。使用者透過諷刺自己的不好和生氣的自己不夠好間接的抒發他難過的心情，而使用者因為不斷的表達低落且負面的情緒，儘管 slightly negative(sad)情緒是不進行加權的，但是他還是超過了情緒門檻情，而性平糾察隊也可以適時的提醒使用者不要過度的負面，可以看出這是一篇情緒較為低落的文章，而透過上下文也成功判定這篇文章是 slightly negative(sad)的情緒。

已讀 2
上午 2:07

我覺得活著好像沒有價值



性平糾察隊

當前情緒累積: 0.3

slightly negative(sad)

上午 2:07

已讀 2
上午 2:07

什麼都做不好



性平糾察隊

當前情緒累積: 1.1

negative (irony)

上午 2:07

已讀 2
上午 2:07

也都沒人會喜歡我



性平糾察隊

當前情緒累積: 2.1

negative (angry)

上午 2:08

已讀 2
上午 2:08

每個人跟我當朋友總是客客氣氣的



性平糾察隊

當前情緒累積: 3.3000000000000003

negative (irony)

上午 2:08

圖 二十五、文本三之字句分析(1)

已讀 2
上午 2:08 我找不到什麼讓我驕傲的事

 性平糾察隊
當前情緒累積: 3.6
slightly negative(sad) 上午 2:08

已讀 2
上午 2:08 也從任何事都得不到成就感

 性平糾察隊
當前情緒累積: 3.9
slightly negative(sad) 上午 2:08

已讀 2
上午 2:08 我現在能懂那些得憂鬱症的人的心情

 性平糾察隊
當前情緒累積: 2.9
positive(happy) 上午 2:08

已讀 2
上午 2:08 有時候死就是一個念頭

 性平糾察隊
當前情緒累積: 3.7
negative (irony) 上午 2:09

圖 二十六、文本三之字句分析(2)



圖 二十七、文本三之字句分析(3)



圖二十八、文本三之字句分析(4)

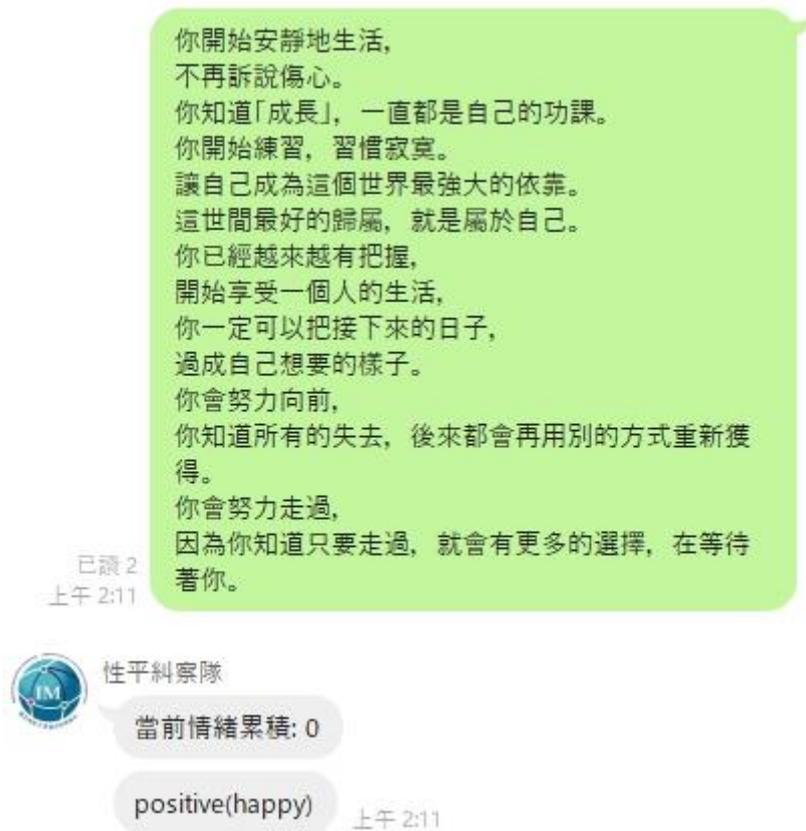
➤ 文本四之結果

(文本四)為一本勉勵自己不要害怕成長的心理勵志書籍裡的一篇文章。因為是心理勵志的書籍，所以可想而知裡面的文章一定都是正面的，而圖二十九在性平糾察隊中判定為 positive(happy)，表示這是一篇帶有正面情緒的文章，判定的結果跟書籍本身所要表達的涵義是一樣的。

文本四之輸入:

你開始安靜地生活，
不再訴說傷心。
你知道「成長」，一直都是自己的功課。
你開始練習，習慣寂寞。
讓自己成為這個世界最強大的依靠。
這世間最好的歸屬，就是屬於自己。
你已經越來越有把握，
開始享受一個人的生活，
你一定可以把接下來的日子，
過成自己想要的樣子。
你會努力向前，
你知道所有的失去，後來都會再用別的方式重新獲得。
你會努力走過，
因為你知道只要走過，就會有更多的選擇，在等待著你。

文本四之輸出:如圖二十九



圖二十九、文本四之測試結果

• 文本四細部字句分析:

圖三十至圖三十三為將文本四拆成單一字句放入性平糾察隊後得到每個句子所表達的情緒，可以從中看出雖然偶爾會有悲傷情緒但是大部分的字句還是被判定為是正面積極的，所以這是一篇帶有正面積極的文章，而透過上下文也成功判定這篇文章是 positive(happy)的情緒。



圖 三十、文本四之字句分析(1)

已讀 2
上午 2:12

讓自己成為這個世界上最強大的依靠。



性平糾察隊

當前情緒累積: 0

positive(happy)

上午 2:13

已讀 2
上午 2:13

這世間最好的歸屬，就是屬於自己。



性平糾察隊

當前情緒累積: 0

positive(happy)

上午 2:13

已讀 2
上午 2:13

你已經越來越有把握，



性平糾察隊

當前情緒累積: 0

positive(happy)

上午 2:13

已讀 2
上午 2:13

開始享受一個人的生活，



性平糾察隊

當前情緒累積: 0

positive(happy)

上午 2:13

圖 三十一、文本四之字句分析(2)



圖 三十二、文本四之字句分析(3)



圖 三十三、文本四之字句分析(4)

➤ 文本五之結果

(文本五)為使用者經過賣披薩的店，發現排隊的人潮眾多，因此在社群媒體上面發文提出疑問，希望能夠得到解答之情境。而圖三十四在性平糾察隊中被判定為 negative(irony)，表示這是一篇帶有諷刺意味的文章。

文本五之輸入:

pizza hut 今天排隊也太誇張了什麼活動？
經過人山人海，排隊比排快篩還誇張
阿是不怕群聚嗎？
一堆外送和排隊買的
吃到也起碼要排一個小時吧，
有這麼好吃嗎？還是有什活動？

文本五之輸出:如圖三十四

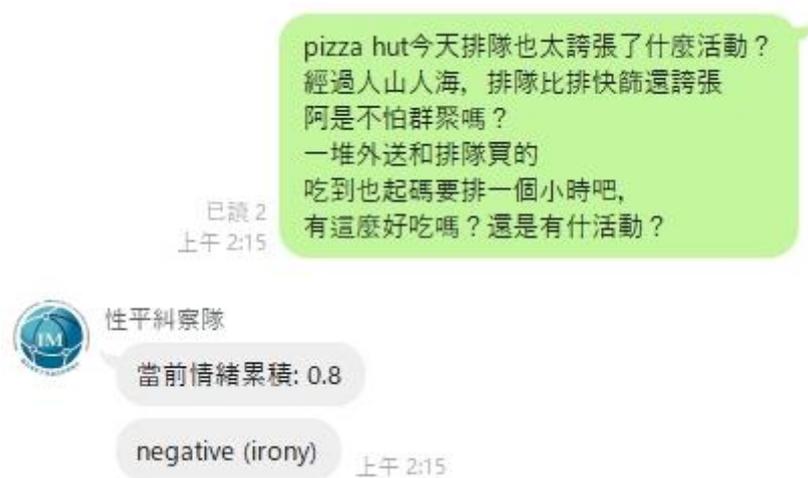


圖 三十四、文本五之測試結果

• 文本五細部字句分析:

圖三十五和圖三十六為將文本五拆成單一字句放入性平糾察隊後得到每個句子所表達的情緒，可以從中看出大部分被判定為生氣、諷刺和侮辱，而諷刺占大部分。使用者在發問的同時利用排隊這件事情來諷刺當下的時事，可以看出這是一篇帶有諷刺意味的負面文章，而透過上下文也成功判定這篇文章是 negative(irony)的情緒。



圖 三十五、文本五之字句分析(1)

已讀 2
上午 2:16

吃到也起碼要排一個小時吧,



性平糾察隊

當前情緒累積: 6.399999999999999

已超過情緒值門檻, 群組的大德們請
冷靜, 各退一步海闊天空

negative (irony) 上午 2:16

已讀 2
上午 2:16

有這麼好吃嗎? 還是有什活動?



性平糾察隊

當前情緒累積: 8.2

已超過情緒值門檻, 群組的大德們請
冷靜, 各退一步海闊天空

negative (irony) 上午 2:16

圖 三十六、文本五之字句分析(2)

- 以個案為基礎之測試結果

本小節將以三個學校的性平案例和兩個法院判決書的案例，總共五個個案對系統進行有效性之驗證。

- 個案一之結果

(個案一)圖三十七到圖四十一的個案中，對話中出現的聊天紀錄都是出自行為人，和傳訊息的對象原為情侶，因為吵架而導致雙方不愉快，所以行為人試著要約被行為人出來，被拒絕後逐漸開始情緒高昂甚至說出了不雅的字句，這時如果使用性平糾察隊的話，可以即時偵測出那句不雅的話違反性平並給予警示，且也可以在情緒累積值超過門檻值時給予提醒，而在對話中可以看出行為人的情緒從諷刺被行為人為什麼不考慮他的感受到後來直接說出不雅甚至違反性平的話來侮辱被行為人。從對話後半段可以看出男方傳的訊息所累積的情緒值都均超過情緒門檻值五，所以都會給予提醒，希望使用者能夠注意自己的情緒。

已讀 2
上午 2:21 開完會上來找我

性平糾察隊
當前情緒累積: 0
positive(happy) 上午 2:21

已讀 2
上午 2:21 我已經好了

性平糾察隊
當前情緒累積: 0
positive(happy) 上午 2:21

已讀 2
上午 2:21 上來說清楚

性平糾察隊
當前情緒累積: 0.8
negative (irony) 上午 2:21

已讀 2
上午 2:21 我就去唱歌

性平糾察隊
當前情緒累積: 0
positive(happy) 上午 2:22

圖 三十七、個案一之測試結果(1)



圖 三十八、個案一之測試結果(2)

已讀 2
上午 2:23

媽的又被你搞到不爽了



性平糾察隊

當前情緒累積: 6.2

已超過情緒值門檻，群組的大德們請
冷靜，各退一步海闊天空

strongly negative(insult) 上午 2:23

已讀 2
上午 2:23

看你媽雞巴啦



性平糾察隊

違反性平！

當前情緒累積: 8.2

已超過情緒值門檻，群組的大德們請
冷靜，各退一步海闊天空

strongly negative(insult) 上午 2:23

已讀 2
上午 2:23

你為什麼要去唱歌



性平糾察隊

當前情緒累積: 10.2

已超過情緒值門檻，群組的大德們請
冷靜，各退一步海闊天空

negative (irony) 上午 2:23

圖 三十九、個案一之測試結果(3)

已讀 2
上午 2:23

給我一個理由啊



性平糾察隊

當前情緒累積: 12.4

已超過情緒值門檻，群組的大德們請
冷靜，各退一步海闊天空

negative (irony)

上午 2:23

已讀 2
上午 2:23

我心情不好你跑去夜唱?



性平糾察隊

當前情緒累積: 14.8

已超過情緒值門檻，群組的大德們請
冷靜，各退一步海闊天空

negative (irony)

上午 2:23

已讀 2
上午 2:24

我心情不好你不來陪我?



性平糾察隊

當前情緒累積: 17.400000000000002

已超過情緒值門檻，群組的大德們請
冷靜，各退一步海闊天空

negative (irony)

上午 2:24

圖 四十、個案一之測試結果(4)

已讀 2
上午 2:24

我在秘密基地，你不上來我就不理你了



性平糾察隊

當前情緒累積: 20.200000000000003

已超過情緒值門檻，群組的大德們請
冷靜，各退一步海闊天空

negative (irony)

上午 2:24

圖 四十一、個案一之測試結果(5)

➤ 個案二之結果

(個案二) 圖四十二的個案中，對話中出現的聊天紀錄都是出自於行為人，在交友軟體上認識同校的被行為人，並在聊天內容中說出帶有性暗示意味的字眼。可以看出當行為人說出性暗示意味的對話之後，雖然沒有違反性平，但是可以看到在聊天內容中，只要是帶有性暗示意味的內容就會被性平糾察隊判定為非常負面的侮辱類別，雖然對話只有短短四句，但是因為判斷出來的情緒都較負面所以最後超出情緒累積值並提醒使用者要注意情緒。

已讀 2
上午 2:27 那妳有在約嗎



性平糾察隊

當前情緒累積: 1

strongly negative(insult) 上午 2:27

已讀 2
上午 2:27 用這個



性平糾察隊

當前情緒累積: 2.0

negative (irony) 上午 2:27

已讀 2
上午 2:27 啊會想試ㄇ



性平糾察隊

當前情緒累積: 3.4

strongly negative(insult) 上午 2:27

已讀 2
上午 2:28 還尸妳是有男朋友了ㄉ可愛女生哈哈



性平糾察隊

當前情緒累積: 5.0

已超過情緒值門檻，群組的大德們請
冷靜，各退一步海闊天空

strongly negative(insult) 上午 2:28

圖 四十二、個案二之測試結果

➤ 個案三之結果

(個案三) 圖四十三到圖四十五的個案中，對話中出現的聊天紀錄都是出自於行為人，而在對話中可以看出行為人明示和暗示被行為人如果不遵從他的意思就會通報其違反規定的事情。雖然對話中並沒有帶有違反性平的字詞出現，但是還是因為負面情緒的累積值超過門檻值所以這時性平糾察隊會給予使用者提醒要注意情緒，但從對話可以看出，儘管最後的幾句對話被判定為是正面的positive(happy)情緒，但是情緒累積值還是超過門檻值五，所以性平糾察隊還是會持續給予提醒。



圖 四十三、個案三之測試結果(1)



圖 四十四、個案三之測試結果(2)

已讀 2
上午 2:29

你有你的心情



性平糾察隊

當前情緒累積: 6.999999999999999

已超過情緒值門檻，群組的大德們請
冷靜，各退一步海闊天空

positive(happy) 上午 2:30

已讀 2
上午 2:30

我也有我的心情



性平糾察隊

當前情緒累積: 5.999999999999999

已超過情緒值門檻，群組的大德們請
冷靜，各退一步海闊天空

positive(happy) 上午 2:30

已讀 2
上午 2:30

你開心就好



性平糾察隊

當前情緒累積: 4.999999999999999

positive(happy) 上午 2:30

圖 四十五、個案三之測試結果(3)

➤ 個案四之結果

(個案四) 為法院判決書上所抓取的個案內容，圖四十六到圖四十九的個案中，對話中出現的聊天紀錄都是出自男方，和傳訊息的對象原為夫妻，但由於男方有家暴的行為，所以女方向法院申請了保護令，但是男方依舊會透過傳訊息的方式燒擾女方。可以從對話看出男方的情緒非常的差，因為判定的結果不是生氣就是非常負面，甚至還有一半的對話都判定為違反性平，而法院判決書的結果是以違反保護令結案。

The image shows a vertical sequence of chat messages and analysis blocks. Each analysis block includes a logo for the 'Gender Equality Mediation Team' (性平糾察隊) and a timestamp of '上午 2:43'. The analysis results are as follows:

- Message 1: "三千五千還扣除利息!他媽的八年看還不還的清! 幹!白癡!..."
Analysis: 當前情緒累積: 1, strongly negative(insult)
- Message 2: "被幹爽的還拿錢給對方!在教你還聽不進去!也沒自我檢討!..."
Analysis: 違反性平!, 當前情緒累積: 2.0, negative (angry)
- Message 3: "要孩子你去找別的男人幹你啦!...操!人在福中不知福!操你媽的王八蛋破麻"
Analysis: 違反性平!, 當前情緒累積: 3.4, strongly negative(insult)

圖 四十六、個案四之測試結果(1)

已讀 2
上午 2:43

你慢慢被騙啦！幹！破麻中的破麻！



性平糾察隊

違反性平！

當前情緒累積: 5.0

已超過情緒值門檻，群組的大德們請冷靜，各退一步海闊天空

strongly negative(insult) 上午 2:43

已讀 2
上午 2:43

林北教你怎麼打贏這官司！幹！...破麻就是破麻無藥可救



性平糾察隊

違反性平！

當前情緒累積: 6.8

已超過情緒值門檻，群組的大德們請冷靜，各退一步海闊天空

strongly negative(insult) 上午 2:43

已讀 2
上午 2:43

林北走後！我會向家事法院提出你的不適任母親！也會跟你社工講你還能被詐欺拿十多萬出來！我會讓審審去讓社會局安置！...



性平糾察隊

當前情緒累積: 8.6

已超過情緒值門檻，群組的大德們請冷靜，各退一步海闊天空

negative (angry) 上午 2:43



圖 四十七、個案四之測試結果(2)

已讀 2
上午 2:44

對我的態度是這樣子！連反省都不會！真可悲！別的人知道自己不對還會道歉！而你都不會！他媽的連狗都不如！可悲！這就是你



性平糾察隊

當前情緒累積: 10.6

已超過情緒值門檻，群組的大德們請冷靜，各退一步海闊天空

negative (irony) 上午 2:44

已讀 2
上午 2:44

要我講幾次啦！破麻！我自己要走的！幹！...



性平糾察隊

違反性平！

當前情緒累積: 13.0

已超過情緒值門檻，群組的大德們請冷靜，各退一步海闊天空

strongly negative (insult) 上午 2:44

已讀 2
上午 2:44

對我的態度如此不好～還敢講什麼？...會有報應的唷～欺騙！野蠻！霸道！這就是你！



性平糾察隊

當前情緒累積: 15.4

已超過情緒值門檻，群組的大德們請冷靜，各退一步海闊天空

negative (angry) 上午 2:44

圖 四十八、個案四之測試結果(3)

已讀 2
上午 2:45

城隍爺那我也會去！你的所作所為天地不容！你等著上天怎麼收



性平糾察隊

當前情緒累積: 18.0

已超過情緒值門檻，群組的大德們請冷靜，各退一步海闊天空

negative (irony) 上午 2:45

圖 四十九、個案四之測試結果(4)

➤ 個案五之結果

(個案五) 為法院判決書上所抓取的個案內容。圖五十和圖五十一的個案中，對話中出現的聊天紀錄都是出自於女方，女方和男方原為朋友關係，因故交惡，某天女方到男方在社群媒體中的發文下方留言並公然侮辱男方。從圖四十五和圖四十六中看出女方的留言判定的情緒不是非常負面就是帶有生氣意味，且全部都違反性平，而法院判決書的結果是以妨礙名譽結案。

已讀 2
上午 2:46

有某個娘娘腔這小人在旁邊煽風點火，造謠是非，加油添醋，很多都很难達成

性平糾察隊

違反性平！

當前情緒累積: 1

strongly negative(insult) 上午 2:46

已讀 2
上午 2:46

把娘娘腔刪除好友而且讓他去吃屎，我讓你這些今年都達成

性平糾察隊

違反性平！

當前情緒累積: 2.2

strongly negative(insult) 上午 2:46

已讀 2
上午 2:46

請娘娘腔不要按讚阿，怎麼這麼討人厭啊那個心機男，可以請他離開我們的生活嗎？臭三八，噁心鬼，不要臉，真是我看過最三姑六婆的娘娘腔

性平糾察隊

違反性平！

當前情緒累積: 3.6

strongly negative(insult) 上午 2:46

圖 五十、個案五之測試結果(1)

已讀 2
上午 2:46

把娘娘腔同性戀刪除好友拉，保證賴會好。他是大病毒~~~



性平糾察隊

違反性平！

當前情緒累積: 5.199999999999999

已超過情緒值門檻，群組的大德們請冷靜，各退一步海闊天空

strongly negative(insult) 上午 2:46

已讀 2
上午 2:47

...雖然我曾經被死娘砲狠狠傷害過但我不會因此就否決所有陰柔的男人...



性平糾察隊

違反性平！

當前情緒累積: 6.799999999999999

已超過情緒值門檻，群組的大德們請冷靜，各退一步海闊天空

negative (angry) 上午 2:47

已讀 2
上午 2:47

你是說娘娘腔+死娘砲+學人精+三姑六婆+不要臉+醜八怪嗎？



性平糾察隊

違反性平！

當前情緒累積: 8.799999999999999

已超過情緒值門檻，群組的大德們請冷靜，各退一步海闊天空

strongly negative(insult) 上午 2:47



圖 五十一、個案五之測試結果(2)

從上面的實驗結果可以看出系統不論是對文句、文本和個案都能夠有效判斷情緒類別、情緒值是否超過門檻值和是否違反性平準則，表九彙整了以文本為基礎的實驗結果，而表十則彙整了以個案為基礎的實驗結果，表十中實際狀況(違反性平?)的個案一到個案三皆為真實個案，結果具敏感性，所以暫不公開實際結果，而個案四跟個案五為法院判決書，此為公開判決，判決結果為騷擾、侮辱及具有性別刻板印象之歧視用語，所以違反性平成立。

表九、文本之實驗結果彙整

	是否超過負面情緒 門檻值	系統判定是否違反性平	系統判定之情緒類別
文本一	是	否	強負面
文本二	是	是	負面
文本三	是	否	弱負面
文本四	否	否	正面
文本五	是	否	負面

表十、個案之實驗結果彙整

	是否超過負面情緒 門檻值	系統判定是否違反性平	實際狀況(違反性平?)
個案一	是	是	因資訊具敏感性， 暫不公開
個案二	是	是	因資訊具敏感性， 暫不公開
個案三	是	是	因資訊具敏感性， 暫不公開
個案四	是	是	是
個案五	是	是	是

六、結論與未來展望

本研究提出了一套性別歧視語意分析與智慧預警系統，並透過文句、文本和個案驗證之有效性，從實驗結果可以看出本研究所提出之系統是能夠有效判斷文本之情緒類別與是否違反性平準則，並在使用者情緒較高昂或是違反性平時，給予提醒與警示，以利提升其性平意識。未來希望可以建置不同類型(如不同性別或不同個性)的智慧機器人並根據使用者的文本情緒與所處情境，來提供更適合使用者的性平諮商。

七、參考文獻

- [1] 王慶福、王郁茗(2007)。大學生性別角色取向與人際能力,關係適應之比較。*中山醫學雜誌*, 18(2), 217-230。
- [2] 朱蘭慧 (2003)。男性性別角色刻板印象之形成與鬆動。*應用心理研究*, (17), 85-119。
- [3] 汪淑娟、陳心怡、許韻婕、謝明鴻、李淑杏、陳家玉。(2009)。大學生性騷擾認知與性別刻板印象之性別差異。*臺灣性學學刊*, 15(1), 67-80。
- [4] 吳肇銘、金志聿、蔡毓霖(2019)。運用文字探勘於教學評鑑分析之研究-以中原大學資管系課程為例。*商管科技季刊*, 20(4), 437-468。
- [5] 李然、林政、林海倫、王伟平、孟丹。(2018)。文本情緒分析綜述。*計算機研究與發展*, 55(1), 30。
- [6] 李傲(2007)。性別歧視的界定。*河北法學*, 25(1), 68-72。
- [7] 杜宜展 (2007)。大學生性別角色刻板印象之研究。*實踐博雅學報*, (8), 111-148。
- [8] 何政光、戴文珠 (2009)。以情境式教學提升護生性別平等認知之研究。*馬偕學報*, (7), 21-47。
- [9] 陳禹函(2010)。藏鏡人的心內話—同性戀公務人員的職場處境。*政治大學公共行政研究所學位論文*, 1-116。
- [10] 陳震瑜、張嘉惠(2020)。基於 BERT-DAOA 的意見目標情感分析。*第三十二屆計算語言學與語音處理研討會論文集({ROCLING} 2020)*, 398-412。
- [11] 陳韋帆、古倫維 (2018)。中文情感語意分析套件 CSentiPackage 發展與應用。*圖書館學與資訊科學*, 44(1), 24-41。
- [12] 陳秀峯 (2019)。職場倫理課程中之性騷擾與多元性別議題探討—以案例式教學為例。*遠東通識學報*, 13(2), 1-31。

- [13] 曾佳雯、郭洪國雄 (2017)。校園推動性別平等教育學習成效之調查研究。
人文社會電子學報，13(1)，1-14。
- [14] 張曉佩、李宜蓉(2018)。多元性別議題諮商能力：訓練方案對知識、覺察與技能影響之研究。*中華輔導與諮商學報*，(52)，51-80。
- [15] 黃裕晴(2017)。文本作品的情緒探勘。*中山大學資訊管理學系研究所學位論文*。1-70。
- [16] 黃曬莉(2007)。性別歧視的多重樣貌。見黃淑玲、游美惠(編)。性別向度與台灣社會。台北:巨流。
- [17] 鄭建澤(2021)。透過 BERT 進行文本情感分析以建構餐廳推薦網站。國立清華大學。
- [18] 衛漢庭、陳牧宏、顧文瑋 (2015)。建構多元性別友善的醫療服務。*護理雜誌*，62(1)，22-28。
- [19] 賴立淇(2019)。大專院校場域學生性別歧視知覺、性別角色與心理健康之相關研究。*國立清華大學碩士論文*。
- [20] 性別平等. (2022, May 15). Retrieved from 維基百科, 自由的百科全書 : <https://zh.wikipedia.org/w/index.php?title=%E6%80%A7%E5%88%A5%E5%B9%B3%E7%AD%89&oldid=71646778>
- [21] 性別平等教育法. (2022, May 24). Retrieved from 維基百科, 自由的百科全書 : <https://zh.wikipedia.org/w/index.php?title=%E6%80%A7%E5%88%A5%E5%B9%B3%E7%AD%89%E6%95%99%E8%82%B2%E6%B3%95&oldid=71794783>
- [22] 多元性別. (2022, January 8). Retrieved from 維基百科, 自由的百科全書 : <https://zh.wikipedia.org/w/index.php?title=%E5%A4%9A%E5%85%83%E6%80%A7%E5%88%A5&oldid=69550957>
- [23] 葉永鋳事件. (2022, May 5). Retrieved from 維基百科, 自由的百科全

書 : <https://zh.wikipedia.org/w/index.php?title=%E8%91%89%E6%B0%B8%E9%8B%95%E4%BA%8B%E4%BB%B6&oldid=71484579>

- [24] Abbasi, A., & Chen, H. (2008). Writeprints: A stylometric approach to identity-level identification and similarity detection in cyberspace. *ACM Transactions on Information Systems (TOIS)*, 26(2), 1-29.
- [25] Agarwal, I., P Rana, D., & S More, A. (2018). Predicting the Veracity of Fake Information on Smart Media for Reducing Misinformation Diffusion. *International Journal of Computational Intelligence & IoT*, 1(2).
- [26] Albadarneh, J., Talafha, B., Al-Ayyoub, M., Zaqaibeh, B., Al-Smadi, M., Jararweh, Y., & Benkhelifa, E. (2015, December). Using big data analytics for authorship authentication of arabic tweets. *In 2015 IEEE/ACM 8th International Conference on Utility and Cloud Computing (UCC)* (pp. 448-452). IEEE.
- [27] Barbon, S., Igawa, R. A., & Bogaz Zarpelão, B. (2017). Authorship verification applied to detection of compromised accounts on online social networks. *Multimedia Tools and Applications*, 76(3), 3213-3233.
- [28] Brocardo, M. L., Traore, I., & Woungang, I. (2015). Authorship verification of e-mail and tweet messages applied for continuous authentication. *Journal of Computer and System Sciences*, 81(8), 1429-1440.
- [29] Butt, S., Ashraf, N., Sidorov, G., & Gelbukh, A. (2021). Sexism identification using BERT and data augmentation-EXIST2021. *In International Conference of the Spanish Society for Natural Language Processing SEPLN 2021, IberLEF 2021*.
- [30] Costa-jussà, M. R. (2019). An analysis of gender bias studies in natural language processing. *Nature Machine Intelligence*, 1(11), 495-496.
- [31] Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2018). Bert: Pre-training of

deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*.

- [32] Freire, P. M. S., & Goldschmidt, R. R. (2019, October). Fake news detection on social media via implicit crowd signals. *In Proceedings of the 25th Brazilian Symposium on Multimedia and the Web* (pp. 521-524).
- [33] Gao, Z., Feng, A., Song, X., & Wu, X. (2019). Target-dependent sentiment classification with BERT. *IEEE Access*, 7, 154290-154299.
- [34] Glick, P., Fiske, S. T., Mladinic, A., Saiz, J. L., Abrams, D., Masser, B., ... & López, W. L. (2000). Beyond prejudice as simple antipathy: hostile and benevolent sexism across cultures. *Journal of personality and social psychology*, 79(5), 763.
- [35] Gillis, N. B. (2021, August). Sexism in the Judiciary: The Importance of Bias Definition in NLP and In Our Courts. *In Proceedings of the 3rd Workshop on Gender Bias in Natural Language Processing* (pp. 45-54).
- [36] Kang, N., Singh, B., Afzal, Z., van Mulligen, E. M., & Kors, J. A. (2013). Using rule-based natural language processing to improve disease normalization in biomedical text. *Journal of the American Medical Informatics Association*, 20(5), 876-881.
- [37] Kesarwani, A., Chauhan, S. S., & Nair, A. R. (2020, June). Fake News Detection on Social Media using K-Nearest Neighbor Classifier. *In 2020 International Conference on Advances in Computing and Communication Engineering (ICACCE)* (pp. 1-4). IEEE.
- [38] Li, D., Guo, H., Wang, Z., & Zheng, Z. (2021). Unsupervised Fake News Detection Based on Autoencoder. *IEEE Access*, 9, 29356-29365.
- [39] Li, W., Gao, S., Zhou, H., Huang, Z., Zhang, K., & Li, W. (2019, December). The

- automatic text classification method based on bert and feature union. In *2019 IEEE 25th International Conference on Parallel and Distributed Systems (ICPADS)* (pp. 774-777). IEEE.
- [40] Mansouri, R., Naderan-Tahan, M., & Rashti, M. J. (2020, August). A semi-supervised learning method for fake news detection in social media. In *2020 28th Iranian Conference on Electrical Engineering (ICEE)* (pp. 1-5). IEEE.
- [41] Mohammad, S. M. (2020). Gender gap in natural language processing research: Disparities in authorship and citations. *arXiv preprint arXiv:2005.00962*.
- [42] Pandya, D. (2019, September). Spam Detection Using Clustering-Based SVM. In *Proceedings of the 2019 2nd International Conference on Machine Learning and Machine Intelligence* (pp. 12-15).
- [43] Qi, P., Cao, J., Yang, T., Guo, J., & Li, J. (2019, November). Exploiting multi-domain visual information for fake news detection. In *2019 IEEE International Conference on Data Mining (ICDM)* (pp. 518-527). IEEE.
- [44] Rasool, T., Butt, W. H., Shaukat, A., & Akram, M. U. (2019, February). Multi-label fake news detection using multi-layered supervised learning. In *Proceedings of the 2019 11th International Conference on Computer and Automation Engineering* (pp. 73-77).
- [45] Reis, J. C., Correia, A., Murai, F., Veloso, A., & Benevenuto, F. (2019). Supervised learning for fake news detection. *IEEE Intelligent Systems*, *34*(2), 76-81.
- [46] Rodríguez-Sánchez, F., Carrillo-de-Albornoz, J., & Plaza, L. (2020). Automatic classification of sexism in social networks: An empirical study on twitter data. *IEEE Access*, *8*, 219563-219576.
- [47] Rodríguez-Sánchez, F., Carrillo-de-Albornoz, J., & Plaza, L. (2021). A Multi-Task

and Multilingual Model for Sexism Identification in Social Networks.

- [48] Saha, T., Jayashree, S. R., Saha, S., & Bhattacharyya, P. (2020). BERT-caps: A transformer-based capsule network for tweet act classification. *IEEE Transactions on Computational Social Systems*, 7(5), 1168-1179.
- [49] Senthil Kumar, N. K., & Malarvizhi, N. (2020). Bi-directional LSTM–CNN combined method for sentiment analysis in part of speech tagging (PoS). *International Journal of Speech Technology*, 23, 373-380.
- [50] Stanczak, K., & Augenstein, I. (2021). A Survey on Gender Bias in Natural Language Processing. *arXiv preprint arXiv:2112.14168*.
- [51] Vosoughi, S., Mohsenvand, M. N., & Roy, D. (2017). Rumor gauge: Predicting the veracity of rumors on Twitter. *ACM transactions on knowledge discovery from data (TKDD)*, 11(4), 1-36.
- [52] Xu, H., Liu, B., Shu, L., & Yu, P. S. (2019). BERT post-training for review reading comprehension and aspect-based sentiment analysis. *arXiv preprint arXiv:1904.02232*.
- [53] Zheng, R., Li, J., Chen, H., & Huang, Z. (2006). A framework for authorship identification of online messages: Writing-style features and classification techniques. *Journal of the American society for information science and technology*, 57(3), 378-393.
- [54] EXIST2021:<http://nlp.uned.es/exist2021/>

八、附錄

● 附錄 A、性平案例詞庫

婊子	臭婊子	婊子配狗	娘炮	娘砲
娘娘腔	約砲	約炮	爸寶	媽寶
賤人	賤貨	死賤人	臭賤人	死賤貨
臭賤貨	賤女人	小白臉	破麻	台女
死台女	死 gay	死三八	臭三八	長得很胸
長得好胸	男人婆	不男不女	被摸一下是會 怎樣	渣男
綠茶婊	臭機掰	死破麻	肉便器	摸胸
吸奶	摳逼	下賤	幹你娘	老雞掰
塞你娘	賤雞掰	狗男女	操雞巴	操你媽
死婊子	酒店妹	畜牲	畜生	賤婊子
破婊子	妓女	鮑魚又臭	死娘砲	死娘炮
吃軟飯	娘砲系	很娘砲	死爸寶	死媽寶
雞巴	雞掰	小娘炮	死八婆	嘴賤
禽獸	豬狗不如	耍三八	死娘娘腔	龜兒子
翹小指	幹卡稱	老賤貨	雞巴毛	靠爸靠媽
爛貨	死賤女人	很騷	多無恥	養小白臉
死 38	臭 38	死母狗	姦夫淫婦	下賤女人
妓女茶妹	娼婦	下三濫	被幹	爛人
臭雞巴	老機掰	老婊子	下三爛	女表
狗婊子	小婊子	作酒店	雜種	公車
酒店飯局妹	北港香爐	蕩婦	很破	石皮
下海	婊子牌坊	破婊	BITCH	洪幹
娼妓	犯賤	騷貨	被人玩過	狐狸精
二手貨	討客兄	淫賤	臭梨子	臭雞歪
爛穴	臭貨	變性人	濺種	慾女

很好幹	躺著賺	援交妹	齷齪	爛梨裝蘋果
公車女	公車騷女	人渣	母狗	爛男人
小雜種	牛郎	騷樣	男妓	機掰小
衣冠禽獸	脫衣陪酒	性服務	欠幹	裝聖女
做半套	臭 B	賤胚	貝戈戈	破格
操你媽的 B	我禽你媽的	狗娘養的	萬人騎	欠人幹
輪奸	輪姦	狗雞巴禽	兼職坐檯	洞癢
跟別人睡覺	38 雞	臭婊	流鶯	王八畜生
母狗瘤	騷禍	賣身	拎娘雞掰	賤禍
賤逼	破鞋	公交車	下賤貨色	禽妳媽
欠禽	爛 B	淫盪	老鮑魚	臭鮑魚
臭懶叫	淫亂	人人上	機八癢	臭幾歪
爛破麻	臭雞掰	臭機巴	人人幹	任人脫
任人騎	查某體	瘋女人	臭卒仔	娘砲剛
沒帶 LP	幹你媽雞巴	破格的人	臭女人	眾人幹
被狗幹	四處被人幹	淫蕩	死畜牲	臭豬
下垂奶	香爐	肥母豬	死肥母豬	臭爛擊敗
性飢渴	母豬騎士	男豬	母豬賽貂蟬	雞巴女人
母豬教	賤婦	豬母	小女豬	小母豬
女豬	奸夫淫婦	坐檯小姐	被幹卡稱	好賤
小賤人	臭鮑	姘頭	鄙下低級	去給狗幹
騷包	傻 B	臭 B	死 B	癌 B
瘤 B	妳幹屌	酒店小姐	婊人	賤小人
狗嫖	SB	PB	婊子還錐子	爛梨仔裝蘋果
潑婦	給別人幹	兩腳開開	伎女	特殊性關係
淫婦	爛婊子	臭賤女人	給人幹	破雞巴
被幹死	女賤	男賤	賣淫	瘋母狗
下賤女	騷屌	賤男人	醜逼	淫女

援交	小狼狗	狼狗	你媽個 B	去給人幹炮
幹炮	賣春的	破雞麥	賤破麻	賤查某
死破母	沒覽教	假鮑魚	雞麥癢	死甲
吹喇叭	操你媽逼	老母狗	零號	吸ㄋㄟㄋㄟ
操雞掰	小人妖	賣屁眼	狗兒子	死人妖
捅屁屁	網路淫蟲	淫蟲	金剛芭比	酒店母狗
性無能	我禽你媽	鴻幹免錢	老機八	當牛郎
偷漢子	媛交	破屌	小破麻	爛雞巴
破麻火爐	臭機歪	破雞八	幹恁娘機掰	愛修幹
做 S	破麻成性	心機婊	基八	眾人用
幹破嘛	破女人	相幹	垃圾女人	臭爛雞
破麻雞	禽老禽雞掰	包養	當香爐	轟幹
破麻禽雞掰	豬生狗養	高級應召女郎	表子	淫蛋女
石皮麻	死破拔	破幹	有性病	女表子
賤破麻	幹恁祖母	破機掰	逼養	破少年
老娼	水性楊花	龜孫子	恐龍妹	操妳媽的 B
去轟幹	好淫蕩	騷浪賤	破雞掰	屎穴
機掰癢	眾人插	下面癢	大黑鮑	美紅幹
勾引男人	蕩婦婊子	陪睡	賤種	賣肉體
有夠破	被包養	臭查某	亂搞男女關係	躺著賺錢
濫 B	濫女人	騷貨母狗	臭雞麥	男盜女娼
下賤妓女				

110年度專題研究計畫成果彙整表

計畫主持人：葉國暉		計畫編號：110-2629-E-259-001-			
計畫名稱：性別歧視語意情緒分析與智慧諮商系統 (L01)					
成果項目		量化	單位	質化 (說明：各成果項目請附佐證資料或細項說明，如期刊名稱、年份、卷期、起訖頁數、證號...等)	
國內	學術性論文	期刊論文	0	篇	準備中 1. 2022年第十五屆崇越論文大賞 2. 2022年台灣教育研究學會學術研討會
		研討會論文	1		
		專書	0	本	
		專書論文	0	章	
		技術報告	1	篇	本研究提出一套性別歧視語意分析與智慧預警系統，該系統將能透過自動化程序針對繁體中文文本進行語意情緒分析，且偵測該文本是否違反性別平等準則，並在使用者情緒較高昂或是違反性平時，給予提醒與警示，以提升其性平意識。
		其他	0	篇	
國外	學術性論文	期刊論文	0	篇	
		研討會論文	0		
		專書	0	本	
		專書論文	0	章	
		技術報告	0	篇	
		其他	0	篇	
參與計畫人力	本國籍	大專生	1	人次	該為大學生主要協助系統建置
		碩士生	2		兩位碩士生主要負責系統建置與期末報告撰寫
		博士生	0		
		博士級研究人員	0		
		專任人員	0		
	非本國籍	大專生	0		
		碩士生	0		
		博士生	0		
		博士級研究人員	0		
		專任人員	0		
其他成果 (無法以量化表達之成果如辦理學術活動、獲得獎項、重要國際合作、研究成果國際影響力及其他協助產業技術發展之具體		正在準備上述提及之論文投稿稿件，此外，計畫研發之性平糾察隊系統亦正在準備參賽(共計三場競賽)，目前已參加"當數據遇上商業—來自數據長的挑戰"競賽，並榮獲ESG獎項(正在等主辦單位的獎盃發放			

效益事項等，請以文字敘述填列。)

，<https://meet.jobs.kktix.cc/events/dataideathon>)。