郝沛毅、歐仁彬、黃天受、楊盛琮（2018），『網路直播聊天室情緒探勘－使用模糊支持向量機』，*中華民國資訊管理學報*，第二十五卷，第二期，頁 185-218。

網路直播聊天室情緒探勘－使用模糊支持向量機

郝沛毅*
國立高雄科技大學資訊管理系

歐仁彬
國立高雄科技大學資訊管理系

黃天受
國立高雄科技大學資訊管理系

楊盛琮
國立高雄科技大學資訊管理系

摘要

「網路直播平台（live streaming video）」是近年不可被忽視的新興社群媒體平台。相較於一般影音平台乃是透過事先拍攝後進行上傳至網路分享，直播多了「即時性」與「互動性」兩種特性，在直播過程中，實況主能依據觀眾在聊天室的反應，立即做出回應，創造出更多的差異化內容，但觀眾踴躍的發言時，實況主就有可能遺漏觀眾的訊息。因此本研究之目的是希望透過情緒探勘技術，探勘聊天室的內容後，以較為簡單方式呈現觀眾想表達的意見，希望有助於實況主能以較輕鬆的方式得知觀眾反應，並可做為內容調整之參酌。

本研究提出之情緒探勘系統主要分為三大步驟，首先透過「網路聊天室爬蟲」建立 Socket 連線至 Twitch-IRC（Server），即時自動擷取聊天室內容，並將取得的留言內容進行「網路用語正規化」的步驟後，再透過模糊支持向量機進行情緒「正面」以及「負面」分類，透過模糊理論可以更精確的計算正面（負面）情緒的歸屬程度。最後以「文字雲」、「情緒波動圖」、「情緒雷達圖」、「情緒直方圖」、「情緒盒鬚圖」等各式圖表進行結果呈現。

關鍵詞：線上直播、情緒分析、網路用語、支持向量機、模糊理論

* 本文通訊作者。電子郵件信箱：haupy@cc.kuas.edu.tw
2017/09/11 投稿；2017/10/02 修訂；2018/03/14 接受

Hao, P.Y., Ou, J.B., Huang, T.S. and Yang, S.C. (2018), 'Sentiment analysis and opinion mining in live streaming by using fuzzy support vector machine', *Journal of Information Management*, Vol. 25, No. 2, pp. 185-218.

Sentiment Analysis and Opinion Mining in Live Streaming by Using Fuzzy Support Vector Machine

Pei-Yi Hao*
Department of Information Management, National Kaohsiung University of Science and Technology

Jen-Bing Ou
Department of Information Management, National Kaohsiung University of Science and Technology

Tien-Shou Huang
Department of Information Management, National Kaohsiung University of Science and Technology

Sheng-Cong Yang
Department of Information Management, National Kaohsiung University of Science and Technology

Abstract

Purpose－Live streaming video is an emerging community media platform in recent years. Compared to the traditional video, live streaming video is more instant and interactive. According to the response of the audience in the chat room, the live streamer can immediately respond and create different content. But when too many messages are generated in the chat room, the live streamer is likely to omit the audience's message. Therefore, the purpose of this research is to mining the contents of audience's message in the chat room through the sentiment mining technology and to present the results in a more simple way.

* Corresponding author. Email：haupy@cc.kuas.edu.tw
2017/09/11 received; 2017/10/02 revised; 2018/03/14 accepted

Design/methodology/approach － The proposed sentiment analysis system consists of three major models. First, we create a Socket connection to Twitch-IRC (Server) of chat room via the web crawler and capture the message from the chat room at predefined intervals. Second, we normalize the internet slang in audience's message into the standard format that can be analyzed. Finally, we propose a fuzzy support vector machine to classify the audience's message into positive or negative emotion.

Findings－On the average, the proposed approach yields satisfactory performance with accuracy rate of 98.88% on internet slang normalization and 87.72% on live streaming sentiment analysis. Questionnaires also demonstrate the efficacy and effectiveness of the resulted sentiment statistics.

Research limitations/implications － This study focused only on accurately classifying the audience's sentiment on live streaming. In our future work, we plan to investigate the correlation between audience's sentiment and audience rating.

Practical implications－The results of sentiment analysis are presented with text clouds, line graph, radar chart, histogram and box plot, etc. This can help the streamer to know the audience's response in a more relaxed way and to adjust the content of the live streaming video according to the audience's response.

Originality/value－This study is, to the best of our knowledge, the first attempt to apply fuzzy support vector machine and word embedding cluster features for the live streaming sentiment analysis in Taiwan.

Keywords: live streaming video, sentiment analysis, web crawler, internet slang, support vector machine

壹、緒論

隨著網際網路的蓬勃發展以及社群平台的興起，人們透過網路社群媒體取得資訊、進行娛樂、參與時事已成為主流，而「網路直播平台（Live Streaming Video）」更是一股不可被忽視的新興平台。不論是國內外直播業者如「Twitch」、「Live.me」、「17 App」以及「LiveHouse.in」，網站流量以及用戶數皆不斷上升，除此之外，全球最大的社群平台「FaceBook」、影音平台「YouTube」皆紛紛發展直播功能，開發新的互動模式，想進一步搭上這波直播熱潮。直播能夠吸引觀眾的主要因素為「即時互動性」，比起一般影音視頻，實況主能依據觀眾在聊天室的反應，立即做出回應，創造出更多的差異化內容，且讓觀眾共同融入情境、營造氣氛，成為內容製造的一份子，觀眾更會覺得因為有他之參與而讓當下直播更具特色。實況主的直播爆紅，除了實況主的個人風格外，最重要的是參與的觀眾，不論是在任何的娛樂產業，「觀眾」都是相當重要的存在，觀眾的感想、反應代表著表演的成功與否，若能獲得認同、肯定，就有機會獲得觀眾轉變為「粉絲（Fans）」，除了持續追蹤觀看外，也會在各自生活圈進行分享或是「打賞（Donate）」，而實況主得到觀眾的認可後，往往會有更好的表現，因此對於實況主來說，如何留住觀眾一直都是相當重要的議題。

由上述內容提到，直播最主要能夠吸引觀眾的因素為「即時互動性」，而在直播進行時，粉絲若想與實況主進行互動，皆是透過留言方式，希望實況主看到留言後可以做出相對應的回饋，但當觀看人數眾多，觀眾又踴躍的發言時，就會發生聊天室快速刷新，而實況主必須顧及拍攝內容進度以及觀眾反應，在分身乏術的情況下，可能時常遺漏觀眾的訊息，來不及觀看，造成無法即時反饋訊息或動作給粉絲。本研究為了改善當實況主在進行直播時，因聊天室內容快速刷新，而造成來不及觀看之問題，希望透過本研究建置之聊天室探勘系統，分別統計觀眾想表達的關鍵字，以及分析當下觀眾的情緒，幫助實況主在不影響直播的進行之下，能以較為簡單的方式得知觀眾情緒反應。

情緒分析（Sentiment Analysis），也稱為意見挖掘（Opinion Mining）、評論挖掘、評價提取、或態度分析，它是進行檢測、提取和分類關於不同主題的意見、情緒和態度的任務（Ravi 2015）。情緒分析可以完成各種預測任務，例如，根據 Amazon 上的產品評論中的情緒，可以預測產品的銷售量（Ghose 2011），根據 Twitter 上電影評論中的情緒，可以預測電影票房（Rui 2013），在財經新聞或社群推文上對於某家公司的情緒，可以用來推測該公司股價的漲跌（Bollen 2012; Li 2014），根據 Twitter 推文中的情緒，甚至可以用來預測美國總統選舉的結果（Jahanbakhsh 2014; Singh 2016）。本研究認為觀眾的情緒與收看人數是有相關

的，然而，針對熱門直播，由於觀眾踴躍發言，聊天室快速刷新導致實況主無法全盤掌握所有訊息與觀眾的整體情緒，因此，本研究主要的目的，是要正確地分類粉絲留言的情緒，並且使用圖形化統計資訊，幫助實況主在不影響直播的進行之下，即時地瞭解觀眾整體的情緒反應，並且適當地做出回應。

在網路的世界，任何人皆可隨時隨地輕易的發表各自的言論、想法以及意見，對任何人來說，面對如此巨量的文字訊息，若單以人工方式觀看與理解網友們想表達的意見以及情緒，顯然有失效率。因此，如何透過文字探勘的技術，自動化從大量的雜亂無章的文字訊息中，以機器識別網路上觀眾的留言中所隱含的正面或負面情緒，已經是在這資訊爆炸的時代相當重要的議題。但自然語言分析在現階段難以相當精準正確地了解分析留言中的情緒與意見，其主要原因為訊息出現當下的「時空背景、環境」所增加的歧異性，再加上網路中常出現「垃圾訊息」等等因素，皆會影響分析之成效。為此本研究將採取字典法（Dictionary-Based Methods），以國立臺灣大學情感詞典 NTUSD（NTU Sentiment Dictionary）以及語文探索辭典 C-LIWC 做為基礎（黃金蘭 2012; Ku 2007），建構情緒以及否定詞詞庫，但若只使用字典法，有可能因為同一句留言中出現多種情緒的詞彙，而無法分析該句留言的情緒極性。因此本研究配合機器學習法（Machine Learning）建置分類器，此分類器透過不斷加入新的資料特徵讓機器自動學習，將文字內容進行情緒分類，希望降低誤判的機率。

本研究的目的是建立一個即時情緒分析系統，透過網路爬蟲程式不斷抓取 Twitch 實況主聊天室新的留言後，借助 Jieba 斷詞系統[1] 進行字詞拆解，將網路用語如注音文、表情符號、特殊用語正規化為可分析的中文格式後，萃取留言中的情緒與語義特徵，透過模糊支持向量機分類該留言為「正面」或「負面」情緒。傳統情緒分類僅考慮情緒的極性為正面或負面，而未考慮到正面與負面情緒之間曖昧的灰色地帶，明顯地，情緒不是只有絕對的正面或負面，而是有一種相對的隸屬程度，本研究將提出使用模糊支持向量機進行使用者留言的情緒分類，模糊支持向量機的輸出為[0,1]區間內的數值，代表該留言屬於正面情緒的隸屬程度，如此更能夠捕捉到現實世界中情緒表達時曖昧與不精確的特性。本研究最後將以「文字雲」、「情緒波動圖」、「情緒雷達圖」、「情緒直方圖」以及「情緒盒鬚圖」等方式呈現結果，希望有助於實況主以較輕鬆方式得知當下內容是否符合觀眾需求，並可做為內容調整之參酌。

1 Jieba 斷詞，Available: https://github.com/fxsjy/jieba

貳、文獻回顧

一、情緒分析與應用

隨著網際網路應用的普及，網路社群已經成為生活中不可或缺的一部份，如早期的「無名小站」、「痞客邦」，一直到現在的「Facebook」、「Instagram」等等，越來越多人喜歡在網路上以文字、圖片或影音等各種形式抒發自己的心情，或者分享生活週遭的點點滴滴，而這些內容其實就代表著使用者的主觀意見，以及個人的情緒和想法。因此在這透過網路就能輕易取得資訊的時代，若能從網路社群平台的大量資料中，分析出有用的資訊，是件相當值得被重視的事情。

情緒分析（Sentiment Analysis）又稱為意見探勘（Opinion Mining），已有十餘年的發展（Ravi 2015），該技術是建立在文字探勘之下，從早期透過顧客消費後所填寫的意見表進行滿意度調查，進而改善產品或服務品質，到近年利用網路社群與論壇中對於某項主題的不同意見與評論等資料進行研究，探勘該文字訊息中想描述的觀點之情緒極性為「正面」或「負面」，分析人們對產品、服務、組織、個人、問題、事件等等的意見、評價、態度以及情緒（Li 2013; Liu 2012; Kang 2013; Rui 2013）。本研究希望透過探討在直播中觀眾的留言內容的情緒歸屬，以留言內容的情緒詞數量以及語義資訊，分析出觀眾的情緒波動與分布，以圖 1 為例，可發現發表該留言的觀眾情緒極性為正面。

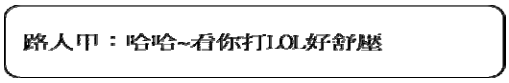


圖 1：Twitch 觀眾留言範例

意見是影響人類做決定的核心，在生活週遭每當需要做決定時，總是想參考他人意見，企業和組織總是想找到關於自身產品與服務的輿論，個人消費者也想知道現有用戶購買產品的意見，甚至政治家也會參考選民的意見來進行政策的決定。本研究根據過往相關研究將其目的歸納整理為三種：

1. 推薦：顧客在網路上尋找某一種服務或產品時，透過網路關鍵字搜尋的結果，通常相當大量與繁雜，若是單以人工過濾來找到符合自身需求之產品，顯然有失效率，因此透過「Opinion Mining」找出眾多品牌中擁有良好口碑之產品，能夠為使用者省下大量時間。例如有關於手機 APP 使用後意見的情緒探勘（林彩雯 2015），以及旅遊經驗的意見探勘（甯格致 2010）。
2. 改善：對於廠商與服務提供者而言，若產品銷售狀況不佳，也能透過對顧

客使用產品之後的意見進行探勘，來找出銷售不如預期之原因，或是參考其他家廠商成功之因素來改善產品缺點或是制定策略。例如在電影場景中研究，找出是「劇情」、「特效」或是「演員」哪種因素更能獲得觀眾好評（邱鴻達 2010），若能找到成功因素，就能改善產品品質，進而增加成功的機率。

3. 預測：預測對企業、政客甚至個人的重要性是不言可喻的，尤其對於企業而言，因為科技進步太快，導致產品週期越來越短，若能善用各種工具來預測未來趨勢，就能協助企業提早準備。例如在音樂領域研究，透過情緒分析之技術探勘從網路爬蟲搜集而來的歌曲相關資訊和音樂排行榜名單資料，就能大致了解音樂閱聽人對於歌詞以及情感的偏好，進而掌握音樂發展的趨勢（廖偉帆 2015）。

情緒分析是一種新興的研究領域，目前在發展上，使用的方法與技術大致分為三大主流（Feldman 2013），依照方法和工具的不同，以下列簡介描述：

1. 機器學習法（Machine Learning）:「監督式學習（Supervised Learning）」必須事先透過人工方式整理好訓練資料（輸入資料和預期輸出結果），再透過訓練資料來學習出預測模型（Learning Model），藉著該模型預測其他資料的類別歸屬。其準確率主要依賴訓練資料整理的品質而決定，優點是預測結果較為精準，但人工整理資料方式需要耗費較長的時間。
2. 詞彙基礎法（Lexicon Based Approach）：在詞彙基礎法中最常被使用的是「字典法（Dictionary-Based Methods）」，字典法是先以人工方式蒐集、整理各式已知的情緒詞，最後再透過比對詞庫的方式，來判斷該句子極性，但若同一句子內同時擁有正、負兩面詞彙，判斷準確率可能不佳，如圖 2 所示，「輸了」屬於負面，「正常」屬於正面詞彙，因有正面詞彙出現，可能被判斷為正面語句，或正負相抵為中立語句，但該句留言以肉眼方式判斷有嘲諷之意，應屬負面語句。

王小明：這場遊戲**輸了很正常**

圖 2：正負詞彙同時出現範例

3. 混合法（Hybrid Methods）：若單純使用「機器學習法」或者「詞彙基礎法」，無法得到令人滿意的結果，因此混合法就是結合上述兩種方法，透過詞彙基礎法偵測關鍵字後，再進行機器學習法的判斷，更能明顯提高分類準確率。本研究為了使得情緒分類結果更加準確，因此採用「混合法」

進行系統實作，以 NTUSD 情緒辭典以及 C-LIWC 語文探索辭典（黃金蘭 2012; Ku 2007）做為詞庫基礎，再透過網路搜尋以及 Twitch 留言爬蟲，找出常用網路用語如「魯蛇、溫拿」等等，進行完整的情緒詞彙建置，之後將 Twitch 留言爬蟲所取的留言進行特徵選取與結果標記，利用「模糊支持向量機」進行分類模型的訓練，詳細方法會在第參節研究方法中介紹。

二、網路用語

網路用語泛指在網路上為了交流所產生與創造之慣用語，或含有特殊意義之諧音以及特殊符號，依照學者楊亨利（2015）的整理，網路用語組成元素包含「中文」、「英文」、「數字」、「注音」、「特殊符號」，且對網路用語分為五大類別，以下列條列方式說明：

1. 表情符號：表情符號又稱為顏文字，以中、英、注音字母以及各式符號所組成，試圖以文字模擬人們臉部表情「^_^」（開心）或肢體動作「Orz」（被打敗）。
2. 英數符號文：以英文字母以及阿拉伯數字所組成，使用原因為加快打字速度，或者想以更詼諧與幽默的方式表達，如「TMD」（他媽的）、「BJ4」（不解釋）等等。
3. 注音文：注音文表達主要分為僅以一個注音符號取代一個中文字，如「ㄅ要」（不要）、「ㄅㄔ」（白癡）、「ㄅㄅ」（掰掰），或是以拼音方式述說難以用中文形式表達的語句，如「ㄅㄡ必」（保佑）、「ㄍㄧㄥ」（死撐）等等。
4. 諧音文：為了使文字解讀起來更有感情，因此將中文以其他類似同音的中文字取代，如「偶愛你」（我愛你）。
5. 特殊用語：意旨各式外來語、或是改編熱門電影、動漫、生活時事等等，如「蘿莉」（可愛的小女孩）、「卡哇依」（可愛）、「魯蛇」（輸家）、「溫拿」（贏家）等。

參考上述文獻整理，本研究將透過人工搜尋以及網路爬蟲蒐集的各種網路用語，依照本研究的需求，將網路用語更精簡分為三大類，其結果如表 1 所示。

表 1：網路用語分類

本研究分類	參考（楊亨利 2015）文獻的分類	定義
表情符號	表情符號	各式文字組成圖像
注音文（單一注音）	注音文	一個注音取代一個中文
特殊用語	英數符號、注音文（拼音）、諧音文、特殊用語	連續英數、注音符號組成諧音、以及各式外來用語

三、網路用語正規化

資訊擷取（Information Extraction）是從自然語言內容中，取出特定主題、事件如「人、事、時、地、物」，因此在文字探勘中是相當重要的環節（Kaiser 2005），但無論是何種資料來源，或多或少都可能出現錯誤如「錯字」。因此，為了避免「垃圾進垃圾出（Garbage in, Garbage out）」的情況，必須在創造各式有價值的資訊應用之前，將取得的資訊進行篩選並更正。對於本研究而言，網路用語就是種錯誤資料，由於網路用語幾乎並未收錄於各式中文詞庫中，因此造成無法準確判斷是否為詞彙以及其代表含意，也就代表著很有可能嚴重影響情緒分析之成效。且因本研究探勘資料來源為網路聊天室之內容，在經過事先分析透過網路爬蟲所擷取的資訊中，發現網路用語如「表情符號」、「注音文」、以及其他「特殊用語」因觀眾年齡層不同，所占比例高達「20%」至「40%」左右，因此必須將這些網路用語正規化為可分析之中文格式，才能繼續後續研究。過往研究的解決方法大多以事先收集各式網路用語，並建立資料庫用於對照與轉換，在上一小節探討的網路用語中，「表情符號」與「特殊用語」由於出現種類有限，透過建立對照表就能解決，但在於注音文部分，一個注音符號就代表著多個中文字，如「ㄅ」就有可能代表著「掰、不、抱、吧、白、啵」，況且組成的文字排列方式眾多，難以將其 37 個注音符號全部所可能代表之中文收集完成。因此本研究參考（Liu 2006）之研究，將注音文（單一注音）資料進行特徵擷取後，利用基於統計學的「多數法則（Majority Rule）」將文字特徵轉換為數值向量，並透過支持向量機（Support Vector Machines; SVM）建置分類器，詳細方法會在第參節研究方法中介紹。

四、支持向量機 SVM

支持向量機（Support Vector Machine; SVM）是一種以 Vapnik 的統計學習理論為基礎的分類技術（Vapnik 1995），並且具有極優良的推理能力（Generalization Ability）。SVM 不像傳統的機器學習（Machine Learning）技術以最小化經驗風險（Empirical Risk）為目標－即使得訓練資料的分類誤差最小化，SVM 以最小化結構風險（Structural Risk）為目標－即使得未知的資料（測試資料）的分類誤差在一個機率上界以下。根據統計學習理論，支持向量機建構出一個具有最大邊界（Margin）的最佳超平面來分割兩個類別的資料，Vapnik 亦證明了最大化邊界等同於最小化推理誤差的上界（Vapnik 1995）。本研究將導入新穎的模糊支持向量機，作為核心的分類器，並用它來判別使用者留言中攜帶的情緒。本研究預期模糊理論能從兩個方向改善情緒分類的能力，首先，文字語言本身就是模糊的，例如高低、大小、快慢等，同時還有同義字與一字多義的問題，

因此，需要模糊理論能夠處理曖昧不精確的現實世界的特性的能力（Zadeh 1965）。此外，在情緒分類問題中，正面與負面情緒之間存在一個模糊的邊界，而非是簡單的二分法，因此需要模糊理論適合處理複雜與不確定問題的優點，判別使用者留言屬於正面與負面情緒的歸屬程度。

參、研究方法

本研究分別編寫「網路爬蟲（Web Crawler）」、「網路用語正規化」以及「模糊支持向量機（SVM）為基礎之情緒分類器」程式，結合「Jieba 斷詞」，整合為「即時直播留言內容情緒探勘」系統，進行情緒分析，其系統流程如圖 3。

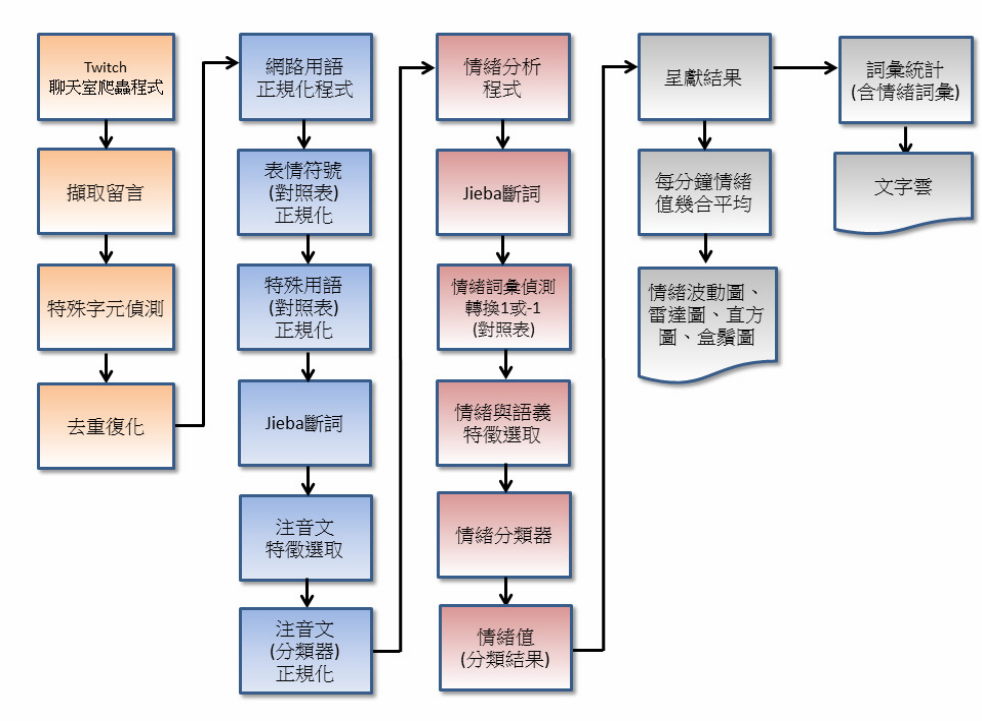


圖 3：系統流程圖

本系統流程可分為三個主要的部份：

1. 使用網路爬蟲程式取得直播留言資料後，進行網路用語正規化。
   (1) 表情符號與特殊用語是以對照表進行正規化。
   (2) 注音文是以 SVM 分類器進行正規化。
2. 將正規化後的資料，進行情緒分析。

(1) 借助 Jieba 斷詞，將句子轉換成數個有意義的詞彙。
(2) 萃取情緒與語義特徵值後，以模糊 SVM 分類器進行句子情緒分類，判斷該句子屬於正面與負面情緒的歸屬程度。

3. 結果呈現
(1) 以文字雲方式，視覺化呈現系統所有統計含有「情緒意義」之詞彙。
(2) 以折線圖、盒鬚圖、直方圖以及雷達圖方式，視覺化呈現觀眾「情緒波動」與「情緒分布」。

一、網路用語正規化

本研究在網路用語正規化中，主要分為兩個部分，首先在「表情符號」與「特殊用語」部分，依照事先建立好的對照表進行正規化，「注音文」則需先選取特徵值再透過 SVM 分類器進行處理，在特徵值選取方面，為了避免取得之特徵值為「表情符號」或「特殊用語」的情況發生，因此必須先進行「表情符號」與「特殊用語」正規化，盡可能減少影響分類結果之因素。本研究提出之網路用語正規化程式架構如圖 4 所示

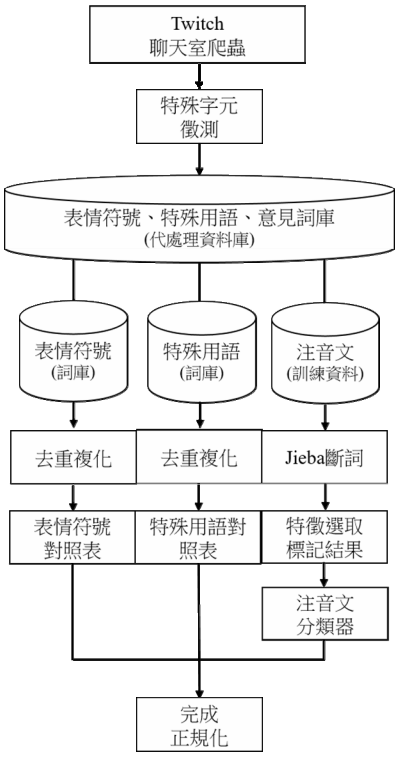


圖 4：網路用語正規化－程式架構

在進行正規化前，系統需先找出可能為網路用法之特殊字元集合如「英文」、「注音」、「數字」、「各式符號」以及少數中文字元如「日、口、皿、囧」，從字串中取出，其方法如下：

（一）特殊字元偵測

設計一迴圈讀取字串中每一字元，判斷字元 Unicode 不屬於中文(u4e00 - u9fff)，除少數中文字元為例外狀況，皆可能為本研究所要正規化之網路用語組成元素，舉例而言，有一筆留言為「厂厂玩得很 OP 喔:D :D :D」，經過程式過濾就能取得「厂厂」、「OP」以及「:D :D :D」。

```
Exception = ["日","口","皿","囧"] #宣告例外清單
if (not u'\u4e00' <= Character<= u'\u9fff') or Character in Exception
#字元不屬於中文範文或屬於例外狀況則為特殊字元
```

（二）去重複化

在使用電腦進行聊天或留言的情境下，能夠很容易透過複製、貼上或其他方式輸入重複詞句來表達意見如「:D :D :D」，在一般使用肉眼辨識可以相當容易辨識為「笑臉」，但在電腦處理自然語言的情況下，想正確轉換「:D」的意思已經是不容易，加上連續重複字元更是難上加難，因此在正規化網路用語之前，去重複化是必要過程，其方法如下。

1. 先判斷偵測到的特殊字元是否大於等於 2，任何表情符號與特殊用語唯有大於兩個（包含）字元組成才有意義。
2. 先將字串「:D :D :D」去除空白符號「:D:D:D」
3. for i in range(1, len(Text)/2) #設計一迴圈次數為字串長度整除於 2

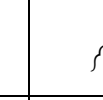
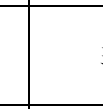
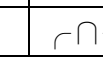
```
Temporary = Text[0:i] #暫存字串為完整字串第 0 至第 i 字元
if Temporary * (len(Text)/i) == Text #如符合則為重複
Text = Text. Replace(Text, Temporary) #將原本字串取代為暫存字串
```

（三）表情符號正規化

由於「表情符號」本身就代表著一種「情緒」，再加上許多表情符號「只能意會，難以形容」如「╰(◕‿◕)╯」、「つ ◕‿◕)つ」，因此本研究不再將表情符號轉換成中文後再判斷情緒的極性，而是參考知名顏文字網站 [2] 以正面（1）、負面（-1）方式標記表情符號的情緒，如表 2 所示。在上述範例中，系統將「:D :D :D」去重複化後，與對照表進行比對，發現與對照表中「:D」相同時，系統則將「:D」轉換成對照表中所對應的情緒「1」。

2 顏文字，Available: http://facemood.grtimed.com/

表 2：表情符號對照表－範例

表情符號對照表			
表情符號	情緒（0、1、-1）	表情符號	情緒（0、1、-1）
:D	1	ＸＤＤＤ	1
	1		1
	-1		1
(ˆ ω ˆ)，	1		-1

（四）特殊用語正規化

「特殊用語」大多以日常諧音或外來語轉變而來，常以固定形式出現，在使用上不易變化，故只需透過資料庫之對照表進行比對轉換即可，如上述例子，系統偵測到「OP」與對照表（如表 3）進行比對，轉換成對應的「厲害」。

表 3：特殊用語對照表－範例

特殊用語對照表（本研究整理）			
特殊用語	中文含意	特殊用語	中文含意
OP	厲害	94 狂	就是厲害
TMD	他媽的	98	走吧
OMG	我的天阿	3Q	謝謝
WTF	見鬼了	BJ4	不解釋
9487	就是白癡	8+9	八嘎囧

「表情符號」與「特殊用語」的正規化方式，完全採用比對對照表的方式進行轉換，因此發生正規化失敗，其原因一定是資料庫中無該筆資料，為了改善與解決此問題，系統將在轉換失敗時，將偵測到的特殊字元進行儲存，再透過人工方式判斷是否新增至「表情符號」及「特殊用語」資料庫。

（五）注音文正規化

注音文的使用在台灣已經是一ㄍ常見ㄉ現象（注音文的使用在台灣已經是一個常見的現象），從 BBS 的盛行，因為新注音輸入法的緣故，使用者們為求方便，懶得將字完整打出來，逐漸轉變成網路文化的一種新興語言。「注音文」是網路用語中「最難」正規化為中文之用語，這是因為相同的注音可以代表的中文

字相當繁多，例如表 4 中的範例，「ㄅ」就有可能代表著「不、吧、爸、八、白、掰、啵、叭、抱、寶」等意思，有時就算是以肉眼觀看也難以理解，且組合方式也難以找到規律，如連續注音「ㄅㄅ、ㄌㄅ」、注音在正中間「要ㄅ要、醜ㄅ怪、打ㄅ球」等等排列方式，皆代表著不同解讀意思，就算是完全相同的注音文，也可能代表著不同的意思，例如「ㄌㄅ」就可能代表著「喇叭」或「瞭吧」。

表 4：注音文範例

注音文	中文翻譯
要ㄅ要	要不要
還好ㄅ	還好吧
ㄎㄅ	哭爸
醜ㄅ怪	醜八怪
ㄅㄔ	白癡
ㄅㄅ	掰掰
ㄅㄧ個	啵一個
ㄌㄅ	喇叭
ㄅ歉	抱歉
ㄅ貝	寶貝

為了正確將注音文正規化為可分析的中文格式，本研究採用監督式機器學習技術中的 SVM 進行資料訓練以及分類（Cortes 1995），然而 SVM 輸入的特徵值必須為相關的數值資料，但自然語言的符號特徵值難以向量化，因此參考（Liu 2006）所提出的特徵值定義方式，其概念如下：「依據字串外表特徵，而不考慮其文字意義。基於統計學上的多數法則（Majority Rule），對字串特徵計算其出現的百分比。假設相同主題資料其特徵值百分比會大致相同的或相當類似（Liu 2006），在本研究中，共定義八個字串特徵如下：

1. Text_Full(T_F)：完整字串（所佔資料庫百分比）
2. Text_Quantity(T_{FQ})：字串長度（數值）
3. Text_First(T_{Ft})：字串第一個是否為要翻注音（bool）
4. Text_Last(T_L)：字串最後一個是否為要翻注音（bool）
5. Text_Continous(T_C)：要翻譯的注音是否連續（bool）
6. Text_OtherPhonetic(T_O)：是否接續其他注音（bool）
7. Text_Start(T_S)：字串開頭詞彙，若為該翻譯注音為 0（所佔資料庫百分

比）

8. Text_End(T_E)：字串結尾詞彙，若為該翻譯注音為 0（所佔資料庫百分比）

為達到注音文正規化，本研究先藉由 Jieba 將字串進行切割，並透過 Unicode 所屬範圍（u3105-u3129）進行注音符號的偵測，綜合以上特徵值定義進行特徵選取，舉例來說，以範例「剛剛那句話可ㄅ可以再說一次」為例。Jieba 斷詞將句子切割為「'剛剛', '那句', '話', '可', 'ㄅ', '可以', '再', '說', '一次'」，偵測到注音符號「ㄅ」後，擷取該注音符號的前後項詞彙，並取得特徵值如表 5 所示。

表 5：字串特徵選取範例

T_F	T_Q	T_{Ft}	T_L	T_C	T_O	T_S	T_E
可ㄅ可以	4	0	0	0	0	可	可以

若要使用 SVM 進行分類，必須將特徵值以數值的格式進行輸入，因此需要將特徵 T_F、T_S 與 T_E 進行向量轉換，為了更加清楚描述特徵轉換方式，在此以表 6 內的資料為範例，進行實際試算。

表 6：字串特徵範例

T_F	T_Q	T_{Ft}	T_L	T_C	T_O	T_S	T_E
可ㄅ可以	4	0	0	0	0	可	可以
要ㄅ要	3	0	0	0	0	要	要
ㄅㄅ	2	1	1	1	0	0	ㄅ
才ㄅ要	3	0	0	0	0	才	要
還好ㄅ	3	0	1	0	0	好	0
可ㄅ可以	4	0	0	0	0	可	可以
可ㄅ可以	4	0	0	0	0	可	可以
ㄅ要	2	1	0	0	0	0	要
ㄅㄅ	2	1	0	1	0	0	ㄅ
可ㄅ可以	4	0	0	0	0	可	可以
來ㄅ	2	0	1	0	0	來	0

根據「該特徵在資料庫筆數 / 資料庫總筆數」換算出所占百分比後，將 T_F、T_S 與 T_E 轉換成數值特徵向量，接著，需要人工標記期望結果才能將資料進行訓練（如表 7 所示），並得到分類模型，以便透過該模型進行其他資料的分類。

表 7：字串特徵範例－續

待轉換	T_F	T_Q	T_{Ft}	T_L	T_C	T_O	T_S	T_E	Label
可勹可以	4/11	4	0	0	0	0	4/11	4/11	不
要勹要	1/11	3	0	0	0	0	1/11	3/11	不
勹勹	2/11	2	1	1	1	0	3/11	2/11	掰
才勹要	1/11	3	0	0	0	0	1/11	3/11	不
還好勹	1/11	3	0	1	0	0	1/11	2/11	吧
可勹可以	4/11	4	0	0	0	0	4/11	4/11	不
可勹可以	4/11	4	0	0	0	0	4/11	4/11	不
勹要	1/11	2	1	0	0	0	3/11	3/11	不
勹勹	2/11	2	1	0	1	0	3/11	2/11	掰
可勹可以	4/11	4	0	0	0	0	4/11	2/11	不
來勹	1/11	2	0	1	0	0	1/11	2/11	吧

特徵 T_F、T_S 與 T_E 透過字串佔資料庫百分比來預測注音文類別的概念如下，假設注音文「勹」用於表示「不」佔有固定的百分比，基於展現方式相互獨立的假設，「可勹可以」字串佔資料庫的百分比，與「勹」表示「不」佔資料庫的百分比，應該是有關聯的，所以可以用字串百分比來預測「勹」的類別，舉例來說，如果「可勹可以」字串佔資料庫的百分比很高，則其中的勹就不可能屬於百分比很低的類別（例如「巴」）。此處，本研究假設「勹」用來表示「不」時，出現在不同展示法（例如「可勹可以」或「勹要」）的百分比是一致或相同的。

二、情緒分類

本研究的情緒分析方法主要分為二大步驟，首先對使用者的留言擷取情緒特徵與語義特徵，接著透過模糊支持向量機進行情緒分類，步驟如下：

（一）取得情緒特徵值

將資料進行正規化後，如「厂厂玩得很 OP 喔:D :D :D :D」正規化為「呵呵玩得很厲害喔」以及[+1]（:D 已標記為正面情緒表情符號），進行 Jieba 斷詞為「呵呵\玩得\很\厲害\喔」[+1]，進行比對「詞彙情緒對照表（如表 8）」，找出訊息中含有情緒字眼的詞彙，並利用表 9 的規則進行情緒轉換，其範例結果為[P,P][+1]。

表 8：詞彙情緒對照表

詞彙	情緒含意（Positive, Negative）
呵呵	P
厲害	P
魯蛇	N
八嘎囧	N

表 9：情緒轉換規則

公式定義	範例	情緒
否定詞+P = N	不喜歡	N
否定詞+N = P	不討厭	P

為了分析使用者留言的情緒，本研究定義五個情緒特徵值，其定義如下：

1. Mood_First(M_F)：字串中第一個出現的情緒（正面、負面或沒情緒）
2. Mood_Positive (M_P)：有幾個正面情緒（int）
3. Mood_Negative (M_N)：有幾個負面情緒（int）
4. QuestionMark(M_{QM})：字串中是否有問號的存在（bool）
5. Mood_Emoticons (M_E)：所有表情符號情緒之平均值，若無則為 0

依照上述例子「厂厂玩得很 OP 喔:D :D :D :D」轉換成 [P,P][+1]，進行特徵值選取並標記結果如表 10 所示。

表 10：情緒特徵選取範例

M_F	M_P	M_N	M_{QM}	M_E	情緒類別
P	2	0	0	1	正面

表 10 中的情緒類別是由三位專家人工裁定的結果，當專家意見分歧時，則採取多數決，專家判定結果的 Fleiss' Kappa 一致性量測分數為 0.8628，一致性屬於理想的。

（二）取得語義特徵（詞嵌入群集特徵，Embedding Cluster Features）

由於直播平台上留言的長度限制與特殊用語等因素，直播平台上的句子比其他領域的文章（例如科學出版物）表現出更高的變異性，大量增加的可變性將導致資料的稀疏性，亦即是在測試樣本中，通常會存在幾個從未見過，或是很少出現的詞彙，因此分類器可能被未見的、或很少發生的字詞干擾其判斷，尤其是當

餵給分類器處理的留言只有數個詞彙時，資料稀疏的問題更為嚴重。近年來，深度神經網路（Deep Neural Networks）和表徵學習（Representation Learning）為解決資料稀疏問題提出了新的思路（Hinton 2006; Bengio 2013），並提出了許多用於學習詞表徵的神經語言模型（Collobert 2011; Mikolov 2013）。word2vec 詞向量表示方法把文字資料從高維度稀疏的形式變成了類似圖像、語音的連續稠密數據，其基本思想是透過 Skip-gram（從當前詞預測上下文）或 Continuous Bag of Words（CBOW，從上下文來預測當前詞）神經語言模型將每個單詞從稀疏的 1-V 編碼（這裡 V 是詞彙量大小），經由隱藏層投影到較低維度的向量空間上，Skip-gram 和 CBOW 兩種模型都可以學習單詞和上下文之間的基本關係。word2vec 最有價值的元素是網路的隱藏層。它包含比詞彙大小更少的神經元，並為單詞提供向量編碼機制，以捕捉單詞和上下文之間的關係，一個詞在神經網路隱藏層上的表現稱為詞嵌入向量（Embedding Vector），它是一個實數值的向量。詞嵌入向量使我們能夠透過簡單地使用兩個嵌入向量之間的歐幾里德距離或餘弦距離，來測量對應字詞之間的語義相關性。透過對詞嵌入向量進行群集分析，以解決詞語空間中資料稀疏性的問題，並將它應用在情緒分析的任務上，是近年來嶄新的研究方向（e.g., Nikfarjam 2015; 張冬雯 2016; Alshari 2017），本研究將建立一組基於語義相似度的詞嵌入群集特徵（Embedding Cluster Features），來表示字詞之間的語義相似性（Nikfarjam 2015）。本研究使用 word2vec 工具[3]來生成 64 維的向量代表詞嵌入特徵，word2vec 將根據字詞在不同句子的上下文內容來學習詞嵌入特徵向量，經常出現在相似內容情境的字詞，將被映射成彼此接近的兩個詞嵌入特徵向量。接著，本研究對詞嵌入特徵向量執行 *K*-means 聚類分析，將語料庫中的詞彙分成 n=150 個不同的群集，其中 n 是使用者設定的整數。如此將可以把大量的字詞映射成數量明顯較小的群集 ID，將頻繁出現與罕見的字詞之間建立鏈接，並解決稀疏性的問題。表 11 顯示部分群集所包含的詞彙列表，如表 11 所示，日期被分配至相同的群集，國家也被分類到相同的群集。根據所生成的詞彙群集，我們可以定義詞嵌入群集特徵，這些特徵透過將相似的詞彙賦予相同的群集編號，來對特徵空間添加更高層級的抽象概念。本研究的詞嵌入群集特徵為 150 個二元的變數 $Em_1, Em_2, \ldots, Em_{150}$，若使用者在直播平台上留言中的字詞出現在詞彙群集 c_i 中，則 Em_i 的特徵值為 1，否則 Em_i 的特徵值為 0。

本研究對於詞嵌入向量（Word Embedding Vector）進行分群，建構詞嵌入群集特徵（Embedding Cluster Features），主要目的是希望：(1)根據語義資訊來擴充情緒辭典；(2)根據語義資訊來捕捉句子的情緒，即便該句子中都沒有出現情緒詞彙。首先，大多數的情緒分析任務，是透過情緒辭典來評估字詞所攜帶的情緒，

3 word2vec. https://code.google.com/p/word2vec/.

然而，根據 Yu（2013）學者的研究，情緒字典是領域相關的，某個領域的情緒字典，可能沒有涵蓋另一個領域的情緒詞彙，例如 Soar（飆升）在財經領域中是屬於正面情緒，但是一般的情緒辭典並不會收錄這個字，因此，Yu（2013）學者透過分析詞彙出現的上下文內容相似度，來找出與某情緒詞彙語義相似的新詞彙，並且將此新詞彙擴充到情緒辭典中。而詞嵌入向量（Word Embedding Vector）則是透過詞彙的上下文內容對詞彙進行編碼，使得我們可以根據兩個詞彙的所對應的嵌入向量之間的距離，來計算這兩個詞彙的語義相似度（上下文內容相似度），在本研究中，我們對於詞嵌入向量（Word Embedding Vector）進行分群，使得具有相似語義的詞彙能夠聚集成為一群，則跟情緒詞彙位於同一群集的詞彙，即便沒有被情緒辭典收錄，但是應該有攜帶相似的情緒訊息，所以跟情緒詞彙代表相同的詞嵌入群集特徵。舉例而言，「輸掉」有被情緒辭典 C-LIWC 收錄為負面詞彙，但是與它語義相似的詞彙如「落敗」、「出局」、「敗陣」，則都未被 C-LIWC 收錄為負面詞彙。但由於上述詞彙都與「輸掉」在相同的詞嵌入群集內，所以在判定情緒極性時，它們有相同的影響力。其次，現今的情緒分析，大多是依據情緒辭典來分析短句中的情緒，如果短句中沒有出現情緒詞彙，則無法判定其相關情緒的極性，本研究使用詞嵌入向量（Word Embedding Vector）分群的結果作為語義特徵，就是希望透過詞嵌入群集特徵（Embedding Cluster Features）來捕捉到這些雖然沒有出現情緒關鍵字，但是有在闡述情緒的語義資訊，例如，「玩遊戲」、「落敗」;「女朋友」、「分手」;「手機」、「摔壞」;「騎車」、「雷殘」，上述詞彙都沒有出現在情緒辭典 C-LIWC 之中，但是其中內含的語義都攜帶有負面的情緒。如表 11 所示，並非所有詞嵌入群集特徵都與情緒有關，因此本研究透過資訊獲益（Information Gain）來對 150 個詞嵌入群集特徵進行篩選，找出其中 30 個與情緒最具相關性的詞群集作為語義特徵。

表 11：詞嵌入群集（Embedding Cluster）的範例

Cluster#	Examples of Clustered Words
c1	輸掉、輸給、敗給、落敗、慘敗、敗陣、一敗塗地、出局、不敵……
c2	女朋友、男朋友、女友、男友、未婚妻、未婚夫、老公、老婆……
c3	分手、離婚、攤牌、絕交、撕破臉皮、決裂、劃清界線、反目成仇、脫離關係……
⋮	⋮
c149	台灣、香港、新加坡、南洋、日本、英國、美國、上海、內地……
c150	四月、八月、十月、九月、十一月、正月、六月……

（三）模糊支持向量機

本研究的核心分類器為嶄新的最大邊界模糊支持向量機，由於支持向量機擁有絕佳的預測能力，對於文字探勘的任務能夠得到理想的結果，而模糊理論很適合處理文字探勘任務，因為文字本身就是模糊的（Zadeh 1965），例如高低與胖瘦，本研究使用的最大邊界模糊支持向量機，則同時融合支持向量機絕佳的推理能力與模糊理論善於處理雜訊與不精確資料的能力（Hao 2016）。本研究認為模糊理論是很適合應用在情緒分析的任務，因為正面與負面情緒之間，存在一條模糊的灰色邊界，非是簡單的二分法，而是有相對的歸屬程度。給定一組訓練資料$\{\mathbf{x}_i, y_i, \mu_i\}$, i=1,…,N, $y_i \in\{-1,1\}$, $\mu_i \in(0,1]$，其中 $\mathbf{x}_i \in R^n$ 表示第 i 筆訓練樣本，y_i 為其對應的類別標籤。每一筆訓練樣本皆指派一個模糊歸屬程度（Fuzzy Membership），以 μ_i 表示，它表示訓練樣本 $\mathbf{x}_i$ 屬於所對應類別的信心強度，模糊支持向量機試圖找出一個具有最大邊界（Margin）的模糊超平面$\langle \mathbf{W}\cdot\mathbf{x}\rangle+\mathbf{B}=\Theta$來分割正負類別，權重向量 **W** 中的元素與偏移量 **B** 皆設定為三角形模糊數字（Triangular Fuzzy Number）。模糊權重向量定義為 **W**=(**w**, **c**)，其中 $\mathbf{w}=[w_1,...,w_n]^t$ 與 $\mathbf{c}=[c_1,...,c_n]^t$，表示近似於 **w**，模糊度為 **c**。模糊偏移量定義為 **B**=(b, d)，代表近似於 b，模糊度為 d。模糊超平面是由底下的歸屬函數所描述：

$$\mu_Y(y)=\begin{cases}1-\left(\left|y-\left(\langle \mathbf{w}\cdot\mathbf{x}\rangle+b\right)\right|\Big/\left(\langle \mathbf{c}\cdot|\mathbf{x}|\rangle+d\right)\right) & \mathbf{x}\neq 0\\ 1 & \mathbf{x}=0, y=0\\ 0 & \mathbf{x}=0, y\neq 0\end{cases} \tag{1}$$

其中Θ代表「模糊零」，它亦是一個三角形模糊數，其中心值為 0，寬度為 O_w，如果下面這個不等式成立時，這組訓練資料被稱為「模糊線性可分割（Fuzzy Linear Separable）」

$$y_i\left(\langle \mathbf{W}\cdot\mathbf{x}_i\rangle+\mathbf{B}\right)\underset{f}{\geq}\mathbf{I}_F \tag{2}$$

其中 $\mathbf{I}_F$ 代表「模糊壹」，它也是一個三角形模糊數，以 1 為中心，I_w 為寬度，$\underset{f}{\geq}$ 表示「模糊大於」，對於兩個三角形模糊數字 $A=(m_A,c_A)$ 與 $B=(m_B,c_B)$，其中 m 為中心，c 為寬度，則：

$$A\underset{f}{\geq}B \text{ 若且維若 } m_A+c_A\geq m_B+c_B \text{ 與 } m_A-c_A\geq m_B-c_B \tag{3}$$

根據公式(1)-(3)，要找出具有最大邊界的模糊超平面來最佳地模糊分割正負

兩個類別，等同於求解下列二次最佳化問題（Quadratic Programming Problem; QPP）：

$$\underset{\mathbf{w},\mathbf{c},b,d,\xi_{1i},\xi_{2i}}{\text{minimize}}\ J=\frac{1}{2}\|\mathbf{w}\|^2+C\left(v\left(\frac{1}{2}\|\mathbf{c}\|^2+d\right)+\frac{1}{N}\sum_{i=1}^{N}\mu_i(\xi_{1i}+\xi_{2i})\right) \tag{4}$$

$$\text{subject to}\ \ y_i\left(\langle\mathbf{w}\cdot\mathbf{x}_i\rangle+b\right)+\left(\langle\mathbf{c}\cdot|\mathbf{x}_i|\rangle+d\right)\geq 1+\mathrm{I}_w-\xi_{1i}$$

$$y_i\left(\langle\mathbf{w}\cdot\mathbf{x}_i\rangle+b\right)-\left(\langle\mathbf{c}\cdot|\mathbf{x}_i|\rangle+d\right)\geq 1-\mathrm{I}_w-\xi_{2i},\ \ d\geq 0,\ \xi_{1i},\xi_{2i}\geq 0\ \ \text{for } i=1,\ldots,N.$$

其中$\|\mathbf{w}\|^2$代表模型的複雜度，最小化$\|\mathbf{w}\|^2$實現統計學習（Statistical Learning）的基礎原理－要獲得絕佳的推理能力（Generalization Ability），必須要同時降低訓練誤差與模型的複雜度，此外，最小化$\|\mathbf{w}\|^2$亦等同於最大化模糊超平面的邊界（Marign），而$\frac{1}{2}\|\mathbf{c}\|^2+d$則代表模型的模糊度，分類模型越模糊，越能夠捕捉到資料的不確定性，而參數 M 控制二者之間的調控。差額變數$\{\xi_i\}_{i=1,..,N}$測量限制條件(2)被違反的程度，參數 C 則是由使用者指定的懲罰參數，C 值越大，對於訓練誤差的懲罰越大。模糊歸屬程度μ_i代表樣本點 $\mathbf{x}_i$ 屬於所對應類別的信心強度，根據拉格朗日（Lagrangian）理論，可得到底下的對偶問題（Dual Problem）：

$$\begin{aligned}\underset{\alpha_{1i},\alpha_{2i}}{\text{maximize}}\ \ &\frac{-1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}y_iy_j(\alpha_{1i}+\alpha_{2i})(\alpha_{1j}+\alpha_{2j})\langle\mathbf{x}_i\cdot\mathbf{x}_j\rangle\\&-\frac{1}{2Cv}\sum_{i=1}^{N}\sum_{j=1}^{N}(\alpha_{1i}-\alpha_{2i})(\alpha_{1j}-\alpha_{2j})\langle|\mathbf{x}_i|\cdot|\mathbf{x}_j|\rangle+\sum_{i=1}^{N}(\alpha_{1i}-\alpha_{2i})\mathrm{I}_w+\sum_{i=1}^{N}(\alpha_{1i}+\alpha_{2i})\end{aligned} \tag{5}$$

$$\text{subject to}\ \ \sum_{i=1}^{N}y_i(\alpha_{1i}+\alpha_{2i})=0,\ \ \sum_{i=1}^{N}(\alpha_{1i}-\alpha_{2i})\leq Cv,\ \ \alpha_{1i},\alpha_{2i}\in\left[0,\frac{C\mu_i}{N}\right]\ \ i=1,...,N$$

求解出上式後，我們得到拉格朗日乘數（Lagrange Multipliers）α_i，權重向量 $\mathbf{w}$ 與 $\mathbf{c}$ 是樣本點$\mathbf{x}_i$與$|\mathbf{x}_i|$的線性組合：

$$\mathbf{w}=\sum_{i=1}^{N}y_i(\alpha_{1i}+\alpha_{2i})\mathbf{x}_i\quad\text{與}\quad\mathbf{c}=\frac{1}{Cv}\sum_{i=1}^{N}(\alpha_{1i}-\alpha_{2i})|\mathbf{x}_i| \tag{6}$$

偏移量參數 b 與 d 的值則可以根據 Karush-Kuhn-Tucker（KKT）最佳化條件來決定，其計算公式如下：

$$b=\frac{-1}{y_i+y_j}\left(y_i\left\langle \mathbf{w}\cdot\mathbf{x}_i\right\rangle+y_j\left\langle \mathbf{w}\cdot\mathbf{x}_j\right\rangle+\left\langle \mathbf{c}\cdot\left|\mathbf{x}_i\right|\right\rangle-\left\langle \mathbf{c}\cdot\left|\mathbf{x}_j\right|\right\rangle-2\right) \tag{7}$$

$$d=\frac{-1}{2}\left(y_i\left\langle \mathbf{w}\cdot\mathbf{x}_i\right\rangle-y_j\left\langle \mathbf{w}\cdot\mathbf{x}_j\right\rangle+\left\langle \mathbf{c}\cdot\left|\mathbf{x}_i\right|\right\rangle+\left\langle \mathbf{c}\cdot\left|\mathbf{x}_j\right|\right\rangle-2I_w\right) \tag{8}$$

對任何 i, j 滿足 $\alpha_{1i}\in(0,C\mu_i/N),\alpha_{2j}\in(0,C\mu_j/N)$ 與 $y_i\cdot y_j=1$。求解出模糊超平面的參數（$\mathbf{w}, \mathbf{c}, b, d$）後，模糊超平面的歸屬函數定義為

$$\mu_{Y_i^*}(y)=1-\frac{\left|y-\left(\sum_{k=1}^{N}y_k(\alpha_{1k}+\alpha_{2k})\left\langle \mathbf{x}_i\cdot\mathbf{x}_k\right\rangle+b\right)\right|}{\left(\frac{1}{Cv}\sum_{k=1}^{N}(\alpha_{1k}-\alpha_{2k})\left\langle|\mathbf{x}_i|\cdot|\mathbf{x}_k|\right\rangle\right)+d} \tag{9}$$

求解出模糊超平面 $Y=\left\langle \mathbf{W}\cdot\mathbf{x}\right\rangle+\mathbf{B}$ 後，對於任何輸入 $\mathbf{x}_i$，$Y_i=\left\langle \mathbf{W}\cdot\mathbf{x}_i\right\rangle+\mathbf{B}$ 是一個對稱三角形模糊數，其中心為 $\left\langle \mathbf{w}\cdot\mathbf{x}\right\rangle+b$ 且寬度為 $\left\langle \mathbf{c}\cdot|\mathbf{x}|\right\rangle+d$。而模糊零 Θ 亦是一個對稱三角形模糊數，其中心為 0 且寬度為 O_w。要判斷新進測試樣本點 $\mathbf{x}$ 屬於那個類別，必須評估它位於模糊超平面哪一邊，亦即，我們必須定義判斷二個三角形模糊數字大小程度的方式，對於任意二個對稱三角形模糊數字 $A=(m_A,c_A)$ 與 $B=(m_B,c_B)$，模糊數 A 大於 B 的模糊程度（亦即 A 位在 B 右邊的模糊程度），是由下列模糊歸屬函數定義

$$R_{\geq B}(A)=R(A,B)=\begin{cases}1 & \text{if } \alpha>0 \text{ and } \beta>0\\ 0 & \text{if } \alpha<0 \text{ and } \beta<0\\ 0.5\left(1+\dfrac{\alpha+\beta}{\max(|\alpha|,|\beta|)}\right) & \text{o.w.}\end{cases} \tag{10}$$

其中 $\alpha=(m_A+c_A)-(m_B+c_B)$ 與 $\beta=(m_A-c_A)-(m_B-c_B)$。注意，當 $m_A=m_B$ 時，$R_{\geq B}(A)=0.5$，當 $m_A<m_B$ 時，$R_{\geq B}(A)<0.5$，反之，當 $m_A>m_B$ 時，$R_{\geq B}(A)>0.5$。因此，本研究中所使用的模糊支持向量機，其決策函數為

$$f(\mathbf{x})=R_{\geq\Theta}\left(\left\langle \mathbf{W}\cdot\mathbf{x}\right\rangle+\mathbf{B}\right)=R\left(\left\langle \mathbf{W}\cdot\mathbf{x}\right\rangle+\mathbf{B},\Theta\right) \tag{11}$$

此決策函數傳回樣本點 $\mathbf{x}$ 屬於正類別的模糊歸屬程度，其值在 0 與 1 之間，如今，正類別與負類別之間存在一條模糊的邊界，這樣更貼近現實界中曖昧與不

精確的現象。要將此模糊超平面延伸到非線性決策曲面，可以透過核心函數（Kernel Function）學習的策略，僅需將訓練樣本透過一個非線性轉換Φ映射至高維度的特徵空間，在高維度特徵空間求得的模糊線性超平面，在原始空間即為模糊非線性決策曲面，支持向量機有個很重要的性質，就是所有的計算都是以資料向量內積（Inner Product）的方式呈現，因此，可以透過定義核心函數 $k(\mathbf{x},\mathbf{y})=\Phi(\mathbf{x})\cdot\Phi(\mathbf{y})$，則無需定義Φ詳盡的函數形式。將公式(5)與(9)中的$\langle \mathbf{x}_i \cdot \mathbf{x}_j \rangle$與$\langle |\mathbf{x}_i| \cdot |\mathbf{x}_j| \rangle$分別用$k(\mathbf{x}_i,\mathbf{x}_j)$與$k(|\mathbf{x}_i|,|\mathbf{x}_j|)$取代，即可得到非線性的模糊決策曲面。

肆、實驗結果

在此章節本研究會著重於「注音、情緒」SVM 分類器所分類的結果，所有的 SVM 實驗皆採用 RBF（Radial Basis Function）核心函數，亦即 $k(\mathbf{x},\mathbf{y})=\exp(-q\|\mathbf{x}-\mathbf{y}\|^2)$，核心函數的參數 q 與誤差逞罰參數 C 的值是透過方格式搜尋決定。在其「表情符號、特殊用語」因採完全比對對照表方式進行正規劃，如有正規化失敗之情形，代表資料庫中無該筆資料，可透過逐步擴充資料庫來提升正確率。經過本研究搜集整理，表情符號共有 372 筆，礙於篇幅限制，表 2 顯示部分的表情符號列表。此外，本研究共蒐集 68 筆特殊用語，表 3 顯示部分的特殊用語列表。針對 SVM 注音分類器的訓練資料，經過本研究搜集整理，共有 5000 筆訓練資料，其中包含 36 種注音符號（ㄥ不在其內），共 118 個類別，表 12 顯示部分的類別與範例。針對 SVM 情緒分類器的訓練資料，經過本研究搜集整理，共有 8000 筆訓練資料，其中 2295 筆為「正面」情緒句子以及 5705 筆為「負面」情緒句子。本研究在建構最大邊界的模糊 SVM 分類器時，訓練樣本的模糊歸屬程度，則是按照 Jiang（2006）學者提出的計算方式去決定，其步驟如下，假設所有樣本皆經由非線性轉換Φ映射至高維度特徵空間，令Φ_+與Φ_-分別表示正類別 C_+與負類別 C_-在特徵空間的中心點，其計算公式為

$$\Phi_+ = \frac{1}{n_+}\sum_{x_i \in C_+}\Phi(x_i) \quad 與 \quad \Phi_- = \frac{1}{n_-}\sum_{x_i \in C_-}\Phi(x_i) \tag{12}$$

其中 n_+ 與 n_- 分別表示正負類別的樣本數，而正與負類別的半徑定義為

$$r_+ = \max_{x_i \in C_+}\|\Phi_+ - \Phi(x_i)\| \quad 與 \quad r_- = \max_{x_i \in C_-}\|\Phi_- - \Phi(x_i)\| \tag{13}$$

則第 i 筆訓練樣本的模糊歸屬度定義如下：

$$\eta_i = \begin{cases} 1-\sqrt{\left\|\Phi_+ - \Phi(x_i)\right\|^2 / (r_+^2+\delta)} & if\ \ y_i = +1 \\ 1-\sqrt{\left\|\Phi_- - \Phi(x_i)\right\|^2 / (r_-^2+\delta)} & if\ \ y_i = -1 \end{cases} \tag{14}$$

其中參數$\delta > 0$用來避免$\eta_i = 0$的情況，明顯地，上式是根據各類別在特徵空間中的中心點與半徑來決定的函數。

表 12：注音符號分類類別（僅顯示部分內容）

注音符號分類類別（本研究整理）			
ㄅ			
類別	範例	類別	範例
不	ㄅ要	巴	ㄅ長
八	醜ㄅ怪	抱	ㄅ歉
叭	喇ㄅ	掰	ㄅㄅ
吧	好ㄅ	暴	ㄅ炸
啵	ㄅ一個	爸	老ㄅ
寶	ㄅ貝	白	ㄅ色
ㄆ			
類別	範例	類別	範例
胚	色ㄆ	笑	ㄆㄆ（笑聲）
砲	大ㄆ		
ㄇ			
類別	範例	類別	範例
嗎	好ㄇ	麼	什ㄇ
媽	老ㄇ	慢	ㄇ吞吞
們	我ㄇ		
⋮			
儿			
類別	範例	類別	範例
兒	儿子	而	儿且

在注音文正規化的分類任務中，會出現類別不平衡（Class Imbalance）的問題，亦即正樣本（目標類別）與負樣本（其餘的類別）的數量相差很大，這種類別不平衡的情況，對於所有的分類器而言（包含類神經網路、貝式分類器等），

都是很棘手的問題。SVM 由於有良好的理論基礎與絕佳的推理能力，所以相較於其它分類器而言，SVM 受到類別不平衡問題的影響較小。SVM 解決類別不平衡的策略有以下幾種（詳見 Batuwita 2012）：

- 重新採樣法（Resampling Methods）：首先在原始類別不平衡資料集上訓練 SVM 分類模型，並且找出對於建構分類模型最具有影響力的樣本點，這些樣本點是位於分類邊界區域的資料點。然後，對於這些有影響力的樣本點進行平衡的重新採樣（並不是盲目地對整個樣本集重新採樣），使得正負樣本的數目相當，這種方法可以大幅減少支持向量機的訓練時間（Batuwita 2010）。
- 集成式學習法（Ensemble Learning Methods）：在此方法中，大類別的資料被分割成數個子集合，使得這些子集合內的資料量與小類別的資料量相當。然後，學習數個支持向量機分類器，每一個分類器都是使用大類別中的一組子集合與小類別整體來進行訓練。最後，將數個分類器的預測結果透過投票的策略來決定新進樣本的類別（Lin 2009）。
- 不同誤差懲罰參數法（Different Error Costs; DEC）：原始的支持向量機對於正 / 負樣本使用相同的誤差懲罰參數（C），這會導致分割超平面往小類別處傾斜，以減少總體的錯誤分類懲罰代價，DEC 方法對於正負樣本給予不同的誤差懲罰參數，假設 C^+ 是正類別的錯誤分類成本，而 C^- 是負類別的錯誤分類成本，同時假設正類別是小類別，負類別是大類別。透過對小類別設定較大的誤差懲罰參數（即 $C^+ > C^-$），即可以減少類別不平衡的影響。也就是說，支持向量機不會再傾向於將分割超平面往小類別處傾斜，因為小類別的樣本現在指派較高的錯誤分類懲罰代價（Veropoulos 1999）。

在本研究中，我們是採取不同誤差懲罰參數法來解決類別不平衡的問題。此外，針對目標類別樣本數目稀少的問題，SVM 有另外一個優點，就是它很適合實作增加式學習（Incremental Learning），增加式學習能夠在給予新進的訓練樣本時，不需要再使用過去全部的訓練樣本來重新學習 SVM 分類模型，而是只需要少數幾個步驟的微調，就可以得到對於新進樣本與過去訓練樣本而言最佳的分類模型（Laskov 2006）。增加式學習很適合線上即時系統，以本系統為例，如果發現有注音分類錯誤的情況時，我們可以將錯誤分類的樣本納入訓練資料集，以增加式學習演算法對 SVM 分類模型很快地進行微調與修正，而無須曠日廢時的重新訓練整個分類模型，透過不斷將錯誤分類樣本納入訓練資料集並且對分類器進行微調，可以使得 SVM 分類器的效能更趨理想，並且解決目標樣本數目稀少的問題。表 13 顯示本研究針對注音文正規化的分類正確率，本研究透過網路爬蟲抓取五位實況主的直播內容，並針對直播內容中的注音文進行正規化，由表 13

顯示，正規化正確率皆在 97%以上，尤其後面三個直播的正確率可達到 100%。

表 13：注音 SVM 分類器－實驗結果（即時資料）

直播平台	實況主	日期	開始時間	結束時間	樣本數	正確率
Twitch	a54****	2017/07/03	17:00:00	18:00:00	154	97.36%
	God**	2017/07/03	23:00:00	23:30:00	34	97.05%
	ky0*******	2017/07/04	15:00:00	15:30:00	7	100%
	Din*******	2017/07/04	21:30:00	22:00:00	70	100%
	Aph*********	2017/07/06	21:00:00	21:30:00	5	100%

接著針對本研究提出的 SVM 注音文分類器進行錯誤分析，列出發生錯誤的情況：

1. 連續注音：在直播一中出現數筆留言「ㄎㄅㄒㄐ」，在連續多筆注音字頭文解讀方面，本研究建立之分類器分類效果極差，其原因為人工難以解讀該文字之意，導致該種類型的訓練資料蒐集相當困難。
2. 留言者打錯字：在直播一中出現 2 筆原始資料有誤（留言者打錯字），將注音符號「ㄧ」輸入成中文字「一」，因此在上一階段特殊用語轉換失敗（注音符號諧音）。
3. 語意模稜兩可：在直播二中出現「4 ㄍ」被分類為「哥」的錯誤結果，其正確分類結果應為「個」，然而「4 ㄍ」同時有可能是「四個」或是「四哥」，必須要考慮上下文的內容才可以判別，無法由單一個使用者留言判別。

表 14 顯示本研究針對用戶留言的情緒分類正確率（accuracy＝$\frac{TP+TN}{TP+TN+FP+FN}$）、精確率（precision＝$\frac{TP}{TP+FP}$）與召回率（recall＝$\frac{TP}{TP+FN}$），其中 *TP* 為真陽率，*TN* 為真陰率，*FP* 為偽陽率，*FN* 為偽陰率。本研究透過網路爬蟲抓取五位實況主的直播內容，並針對直播內容中的留言進行情緒分類，在進行分類正確與否判斷時，則是透過人工方式依照專家主觀意見進行判斷。表 14 的結果是由三位專家來人工裁定判斷結果，當專家意見分歧時，則是採取多數決，專家判定結果的 Fleiss' Kappa 一致性量測分數為 0.8646，一致性屬於理想的。如表 14 所示，本研究使用的模糊支持向量機在情緒分類的正確率略大於傳統支持向量機，而本研究的模糊支持向量機最主要的貢獻，是可以精確地表達留言屬於正（或負）面情緒的模糊歸屬程度，而傳統支持向量機僅能用絕對的二分法，分成正面或負面情緒，而無法處理現實世界中情緒的曖昧模糊的特性。

表 14：情緒 SVM 分類器－實驗結果（即時資料）

直播平台	實況主	日期	開始時間	結束時間	樣本數	支持向量機			模糊支持向量機		
						正確率	精確率	召回率	正確率	精確率	召回率
Twitch	a54****	2017/07/03	17:00:00	18:00:00	4283	89.66%	90.20%	88.33%	89.89%	90.45%	88.58%
	God**	2017/07/03	23:00:00	23:30:00	729	88.20%	79.37%	81.57%	88.75%	80.27%	82.49%
	ky0*******	2017/07/04	15:00:00	15:30:00	145	89.66%	96.43%	65.85%	90.34%	96.55%	68.29%
	Din*******	2017/07/04	21:30:00	22:00:00	732	91.80%	74.50%	83.46%	92.08%	75.17%	84.21%
	Aph*********	2017/07/06	21:00:00	21:30:00	138	76.81%	96.30%	63.41%	77.54%	96.36%	64.63%

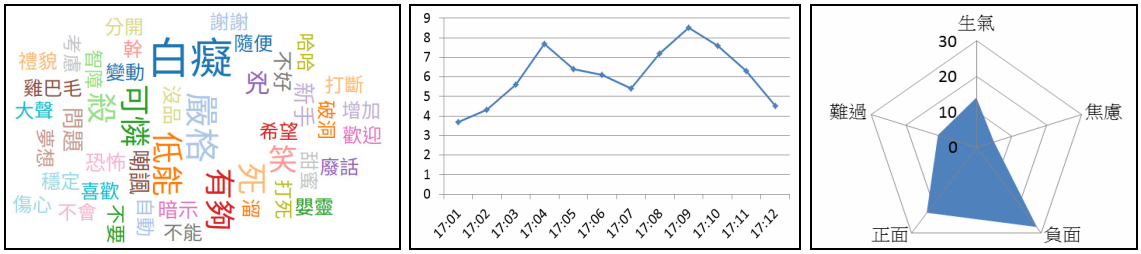
接著針對本研究提出的 SVM 情緒分類器進行錯誤分析，列出錯誤發生的情況。

1. 嘲諷用語：「小件跟電競教父學一學禮貌」，意旨嘲諷負面之意，但「禮貌」為正面詞彙，在未考慮使用正面詞彙進行嘲諷的情境下，造成分類錯誤。
2. 問候用語：在聊天過程中，用於關心與問候的語句，應被分為正面語句，但因句子中含有負面詞彙所造成分類錯誤，如「辛苦了」。
3. 使用負面詞進行強調：利用文字技巧來強調與形容某件事情，例如「文字不足以表達我澎湃的內心」，但用於強調的「不足以」是屬於負面詞彙。
4. 特定情境：在特定情境與主題所出現的聊天內容，本屬無情緒波動之語句，但因句子中出現含情緒詞彙所造成的錯誤分類結果，如本研究第 5 次實驗－實況主「aphrolinll07」，所進行直播內容為 Live 歌唱，在該情境中觀眾透過聊天室輸入歌名進行點歌，此時歌名中若含有情緒詞彙如「迷途羔羊」，就會造成系統將該內容視為有情緒波動的語句。

伍、系統輸出展示與分析

為了讓實況主在不影響直播的進行之下，能夠以較為簡單的方式得知觀眾的情緒反應，讓實況主以較輕鬆的方式得知當下內容是否符合觀眾需求，並可做為內容調整之參酌，本研究以底下幾種統計圖表呈現直播即時情緒分析的結果，圖 5 顯示「文字雲」、「情緒波動圖」與「情緒雷達圖」的呈現結果。文字雲是一種將關鍵詞視覺化的呈現方式，以關鍵字出現頻率定義文字大小，常被用於表示在集合中各個項目的數據大小。在本研究中，文字雲能讓實況主一目了然熱門的情緒關鍵字及其出現次數的多寡。另外，本研究也使用情緒波動圖讓實況主瞭解直播現場的情緒波動情況，我們以一分鐘為間隔，將此時間間隔中直播現場所有的留言內容進行情緒分析，獲得每一句留言的情緒分數，對此時間間隔內所有留言

的情緒分數計算平均值後，以折線圖呈現情緒的波動狀況。注意，如果使用傳統的支持向量機進行情緒分類，每一個留言的情緒分數只有正面（1）與負面（-1）這兩種可能，反之，透過本研究的模糊支持向量機進行情緒分類，可以得到每一個留言屬於正面與負面的歸屬程度，因此可以更精細的呈現直播現場的情緒波動情況。另外，為了更進一步呈現出關鍵字所代表之情緒含意，本研究透過「C-LIW」語文探索辭典將「焦慮」、「生氣」、「難過」三種情緒從負面詞彙獨立出來（黃金蘭 2012），並統計關鍵字在各情緒所占比例後，使用雷達圖呈現結果。



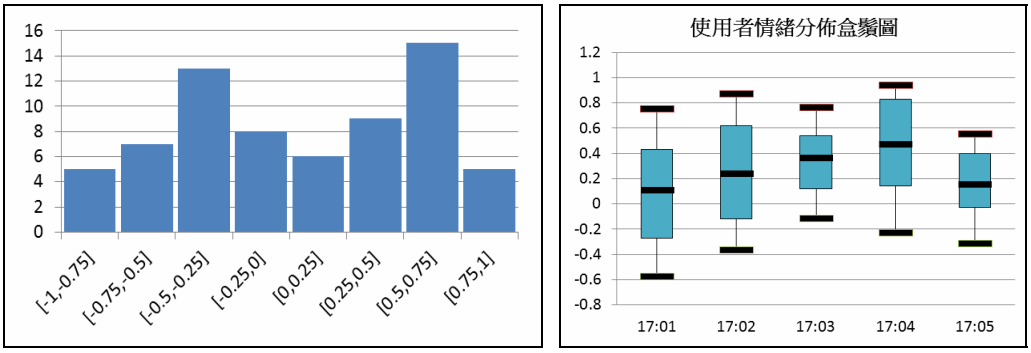
(a) (b) (c)

圖 5：針對全部留言的情緒統計圖表有(a)文字雲、(b)情緒波動圖、與(c)情緒雷達圖

這些統計圖表能即時的讓實況主了解觀眾的情緒轉折，並且立即做出反應，以 a54****的場景為例子，這是一個線上電競遊戲的直播，由於實況主是非常有名氣的電競高手，觀看線上直播的粉絲非常多，在 1 個小時的直播內容中，共有 4283 條留言，平均每秒有 1.19 條留言，因此在直播過程中，實況主根本無法一邊參與電競，同時逐一審閱每條留言並且做出反應，在 17:04 分鐘時，由於實況主採取保守的迴避策略，導致隊友陷入危機，因此直播的觀眾群情激憤，而實況主從情緒波動圖中，看到觀眾情緒劇烈轉為負面，趕快採取積極的策略攻擊對手並成功驅趕敵人，因此才讓情緒波動往正面方向轉折（如圖 5(b)所示）。

注意，上述三種統計圖是針對所有的留言進行即時的情緒分析，並未考慮到個別使用者，然而，有的時候可能少數使用者非常踴躍的發言（洗版），使得實況主誤認為那群少數使用者的情緒就是代表整體使用者的情緒，而產生偏差的認知。為了避免以上情況發生，本研究對每一個使用者的留言進行情緒分析，將時間間隔內使用者留言的情緒分數計算平均值後，即可得到該使用者的情緒分數。在計算完所有使用者的情緒分數後，本研究以「情緒直方圖」與「情緒盒鬚圖」呈現使用者情緒分布的情況，如圖 6 所示。直方圖（Histogram）是一種描述數據分布情況的圖形表示法，它是一種二維統計圖表，沿著橫軸以各組組界為分界，組距為底邊，以各組內樣本的數目為高度，依序在固定的間距上畫出矩形高度所繪製而成之圖形。直方圖可以讓實況主一目了然所有觀眾的情緒是否一致（單

峰）、兩極化（雙峰）、或是雜亂無章（多峰）。盒鬚圖又稱為箱形圖（Box plot），它是一種用作顯示一組數據分散情況的資料統計圖。它能顯示出一組數據的最大值、最小值、中位數、及上下四分位數。我們以一分鐘為間隔，將此時間間隔中直播現場所有的使用者的留言內容進行情緒分析，獲得每一個使用者的情緒分數，「情緒盒鬚圖」可以幫助實況主瞭解情緒最正面與負面的使用者的情緒分數（極端值），也可以知道所有使用者的情緒分數的平均值，同時可以瞭解所有使用者的情緒分數的變異量，即上四分位數（Q3）與下四分位數（Q1）之間的差距（即四分位間距 Interquartile Range），該差距越大，代表直播現場的使用者的情緒變異量越大。同樣地，如果使用傳統支持向量機的二元分類，情緒分數僅有正面（1）與負面（-1）這兩種情形，會造成直方圖與盒鬚圖呈現結果不盡理想。而透過模糊支持向量機能得到屬於正面與負面的歸屬程度，情緒分數落在[-1,1]之間，因此能得到更好呈現效果。



(a) (b)

圖 6：針對全部使用者的情緒統計圖表有(a)情緒直方圖與(b)情緒盒鬚圖

最後，本研究採用問卷調查法來衡量受訪者對於各種情緒統計圖表能夠幫助直播進行的認同程度。問卷設計採取封閉式問卷，並採用李克特的五點尺度量表，量表上標明「非常同意」、「同意」、「普通」、「不同意」、「非常不同意」等五個選項，在資料輸入上給予 5 至 1 分，分數越高表示受訪者對該選項的同意程度越高。調查對象為擔任過直播實況主至少一次的使用者，問卷發放方式包含電郵、面訪與網路問卷，扣除掉無效問卷後，共回收 144 份有效問卷，敘述性統計分析結果顯示於表 15。由表 15 得知，「情緒波動圖」的平均數為最高，「文字雲」次高，而「情緒直方圖」的平均數為最低，「情緒雷達圖」的標準差為最高，「情緒盒鬚圖」次高，而「情緒波動圖」的標準差為最低。各項目認同度都在 4 分（同意）左右，代表實況主認同本研究提供的情緒統計圖表能夠幫助直播

更加順利。其中「情緒波動圖」與「文字雲」由於簡單明瞭，因此最受到實況主的認同，反之，「情緒直方圖」與「情緒盒鬚圖」雖然能夠提供的資訊量較多，但是必須要具備基礎的統計觀念才能解讀其中的內容，因此認同度較低。接著，本研究使用 SPSS 20.0 軟體對問卷調查結果進行信度分析，信度（Reliability）是在反應測量結果的一致性（Consistency）或穩定性（Stability），最常使用的檢測指標為 Cronbach's Alpha 值，Cronbach's Alpha 值小於 0.35 代表低信度，介於 0.35 與 0.7 之間代表中信度；Cronbach's Alpha 值大於 0.7 代表高信度。本研究的 Cronbach's Alpha 值為 0.817，代表問卷信度是可接受的。受限於現今實況主多數是以手機進行直播，因此問卷訪問對象並未侷限於是本系統的使用者，如何將情緒統計資訊以更精簡的畫面顯示在手機螢幕上，以供實況主參考，是本研究未來的研究方向。

表 15：問卷調查結果

受訪者資料分布百分比								
性別	男性	44.4%	年齡	20 歲以下	2.1%	教育程度	國中（含以下）	0%
	女性	55.6%		20-30 歲	63.2%		高中 / 職	9.7%
婚姻狀況	未婚	76.4%		30-40 歲	34.0%		大專（學）	76.4%
	已婚	23.6%		40 歲以上	0.7%		研究所以上	13.9%
職業	商業	15.3%	平均月收入	20000 元以下	47.2%	擔任直播實況主次數	1-5 次	63.9%
	工業（製造業）	5.6%		20001-30000	26.4%		6-10 次	15.3%
	服務業	17.4%		30001-40000	17.4%		11-15 次	11.8%
	軍公教	6.3%		40001-50000	5.6%		16-20 次	5.6%
	家管	9.7%		50001-60000	2.8%		20 次以上	3.5%
	學生	35.4%		60001-70000	0.7%			
	其他	10.4%		70000 以上	0%			

題號	題目	平均數	標準差
1	請問您認為「文字雲」能幫助實況主掌握觀眾情緒並讓直播更順利嗎？	4.23076	0.65518
2	請問您認為「情緒波動圖」能幫助實況主掌握觀眾情緒並讓直播更順利嗎？	4.44055	0.65443
3	請問您認為「情緒雷達圖」能幫助實況主掌握觀眾情緒並讓直播更順利嗎？	4.00699	0.73376
4	請問您認為「情緒直方圖」能幫助實況主掌握觀眾情緒並讓直播更順利嗎？	3.74125	0.67893

5	請問您認為「情緒盒鬚圖」能幫助實況主掌握觀眾情緒並讓直播更順利嗎？	3.76223	0.72626
信度（Cronbach's Alpha）		0.817	

陸、結論

本研究建構一個即時的直播聊天室情緒探勘系統，由於注音文是時下年輕人流行的網路用語，本研究為了正規化網路用語所建置之「注音文 SVM 分類器」，在進行偵測並正規化注音字頭文的準確率平均值為 98.88%，除了原始資料就存在著錯誤（錯字）所導致分類錯誤的情形外，所發生的錯誤，不外乎是訓練資料所涵蓋類別或特徵不夠完善，相信若能增加訓練資料的完整程度以及資料量，就能再提高準確率。在「情緒 SVM 分類器」進行情緒分類準確率平均值為 87.72%，發生分類錯誤的原因皆是只靠句子中出現的少數情緒詞與詞嵌入群集特徵進行情緒分類，並無帶入情境與前後文內容，但本研究所探勘內容皆為極短之句子（聊天內容），且每筆資料皆可能由不同的人進行留言，前後句很難存在著關聯，因此建議後續研究者能將這些問題進行整合，做為資料特徵值選取之參考依據，相信在短文句探勘領域會有很大之研究價值。在這已經資訊爆炸的時代，不論是政府、企業家與研究學者，均想透過各式社群平台分析民意、消費者偏好與各式未來趨勢等等，因此本研究將以最大限度提供所蒐集與整理的資料以及方法，希望能夠為中文短句子探勘領域帶來些許貢獻。

誌謝

本文接受行政院科技部專題研究計畫（MOST 106-2221-E-151-049-）之補助研究經費，順利完成此篇著作之研究工作，謹此致謝。

參考文獻

林彩雯（2015），『以 Google App 評論為字詞權重調整之情緒分析系統』，未出版碩士論文，靜宜大學資訊管理學系，台中市。

邱鴻達（2010），『意見探勘在中文電影評論之應用』，未出版碩士論文，國立交通大學資訊科學與工程研究所，新竹市。

黃金蘭（2012），『中文版語文探索與字詞計算字典之建立』，*中華心理學刊*，第 54 卷，2 期，頁 185-201。

張冬雯、楊鵬飛、許雲峰（2016），『基於 word2vec 和 SVMperf 的中文評論情感分類研究』，*計算機科學（Computer Science）*，第 43 卷，第 6A 期，頁

418-421。

楊亨利、黃泓彰、林青峰（2015），『基於決策樹與二元語言模型的網路用語轉譯系統』，*電子商務學報*，第 17 卷，第一期，頁 25-48。

廖偉帆（2015），『熱門華語流行音樂歌詞情緒分析與趨勢發展』，未出版碩士論文，實踐大學資訊科技與管理學系碩士班，台北市。

甯格致、賴昆棋（2010），『基於網路社群之旅遊經驗及對應情境之情感意見分析研究』，*第廿二屆自然語言與語音處理研討會（ROCLING 2010）*，南投縣，台灣，9 月 1-2 日。

Alshari, E., Azman, A., Alkeshr, M., Doraisamy, S.C. and Mustapha, N. (2017). 'Improvement of sentiment analysis based on clustering of Word2Vec features', *2017 28th International Workshop on Database and Expert Systems Applications (DEXA 2017)*, pp. 123-126.

Batuwita, R. and Palade, V. (2012), 'Class imbalance learning methods for support vector machines', in He H. and Ma Y. (Eds.), *Imbalanced Learning: Foundations, Algorithms, and Applications*, John Wiley & Sons, Inc, Chapter 6.

Batuwita, R. and Palade, V. (2010), 'Efficient resampling methods for training support vector machines with imbalanced datasets', in *Proceedings of the International Joint Conference on Neural Networks*, pp. 1-8.

Bengio, Y., Courville, A. and Vincent, P. (2013), 'Representation learning: A review and new perspectives' , *IEEE TPAMI*, Vol. 35, No. 8, pp. 1798-1828.

Bollen, J., Mao, H. and Zeng X. (2011), 'Twitter mood predicts the stock market', *Journal of Computational Science*, Vol. 2, No.1, pp. 1-8.

Collobert, R., Weston, J., Bottou, L., Karlen, M., Kavukcuoglu, K. and Kuksa, P. (2011), 'Natural language processing (almost) from scratch', *Journal of Machine Learning Research*, Vol. 12, pp. 2493-2537.

Cortes C. and Vapnik V. (1995), 'Support-vector networks', *Machine Learning*, Vol. 20, No.3, pp. 273-297.

Feldman, R. (2013), 'Techniques and applications for sentiment analysis', *Communications of the ACM*, Vol. 56, No.4, pp. 82-89.

Ghose, A. and Ipeirotis, P.G. (2011), 'Estimating the helpfulness and economic impact of product reviews: Mining text and reviewer characteristics', *IEEE Transactions on Knowledge and Data Engineering (TKDE)*, Vol. 23, No.10, pp. 1498-1512.

Hao, P.-Y. (2016), 'Support vector classification with fuzzy hyperplane', *Journal of Intelligent & Fuzzy Systems*, Vol. 30, No. 3, pp. 1431-1443.

Hinton, E. and Salakhutdinov, R. (2006), 'Reducing the dimensionality of data with

neural networks', *Science*, Vol. 313, No. 5786, pp. 504-507.

Jahanbakhsh, K. and Moon, Y. (2014), 'The predictive power of social media: On the predictability of U.S. presidential elections using Twitter', *arXiv:1407.0622*.

Jiang, X., Yi, Z. and Lv, J.C. (2006), 'Fuzzy SVM with a new fuzzy membership function', *Neural Computing & Applications*, Vol. 15, No. 3-4, pp. 268-276.

Kaiser, K., Miksch, S. (2005), 'Information extraction-a survey', *Technical Report for Asgaard-TR-2005-6*, Vienna University of Technology, Institute of Software Technology and Interactive Systems, Vienna.

Kang, D. and Park Y. (2013), 'Review-based measurement of customer satisfaction in mobile service: Sentiment analysis and VIKOR approach', *Expert Systems with Applications*, Vol. 41, No. 4, Part 1, pp. 1041-1050.

Ku L.-W. and Chen, H.-H. (2007), 'Mining opinions from the web: Beyond relevance retrieval', *Journal of the American Society for Information Science and Technology*, Vol. 58, No. 12, pp. 1838-1850.

Laskov, P., Gehl, C., Kruger, S. and Muller, K.-R. (2006), 'Incremental support vector learning: Analysis, implementation and applications', *Journal of Machine Learning Research*, Vol. 7, pp. 1909-1936.

Li, Y.-M. and Li, T.-Y. (2013), 'Deriving market intelligence from microblogs', *Decision Support Systems*., Vol. 55, No. 1, pp. 206-217.

Li, X., Xie, H. and Chen L. (2014), 'News impact on stock price return via sentiment analysis', *Knowledge-Based Systems*, Vol.69, pp. 14-23.

Lin, Z., Hao, Z., Yang, X. and Liu, X. (2009), 'Several svm ensemble methods integrated with under-sampling for imbalanced data learning' , in *Proceedings of the 5th International Conference on Advanced Data Mining and Applications*, pp. 536-544, Springer-Verlag.

Liu, J.-S. and Cheng, Y.-W. (2006), 'Textual data error detection based on string features', in *ROCLING 2006*.

Liu, B. and Zhang, L. (2012).' A survey of opinion mining and sentiment analysis', in *Mining Text Data* (pp. 415-463), Springer.

Mikolov, T., Sutskever, I,. Chen, K., Corrado, G. and Dean, J. (2013), 'Distributed representations of words and phrases and their compositionality', in *Proceedings of NIPS 2013*.

Nikfarjam, A., Sarker, A., O'Connor, K., Ginn, R. and Gonzalez, G. (2015), 'Pharmacovigilance from social media: Mining adverse drug reaction mentions using sequence labeling with word embedding cluster features', *Journal of the*

American Medical Informatics Association, Vol. 22, No. 3, pp. 671-681.

Ravi, K. and Ravi, V. (2015), 'A survey on opinion mining and sentiment analysis: Tasks, approaches and applications', *Knowledge-Based Systems*, Vol. 89, pp. 14-46.

Rui, H., Liu, Y. and Whinston A. (2013), 'Whose and what chatter matters? The effect of tweets on movie sales', *Decision Support Systems*, Vol. 55, No. 4, pp. 863-870.

Singh, S. (2016), 'Trump win: Poll experts failed but AI by an Indian got it right in October', *India Today*, available at http://indiatoday.intoday.in/technology/story/trump-win-poll-experts-failed-but-ai-by-an-indian-got-it-right-in-october/1/807123.html (accessed 12 January 2018).

Vapnik, V.N. (1995), *The Nature of Statistical Learning Theory*, Springer-Verlag, New York.

Veropoulos, K., Campbell, C. and Cristianini, N. (1999), 'Controlling the sensitivity of support vector machines', *Proceedings of the International Joint Conference on Artificial Intelligence*, pp. 55-60.

Yu, L.-C., Wu, J.-L., Chang, P.-C. and Chu, H.-S. (2013), 'Using a contextual entropy model to expand emotion words and their intensity for the sentiment classification of stock market news', *Knowledge-Based Systems*, Vol. 41, pp. 89-97.

Zadeh, L.A. (1965), 'Fuzzy sets', *Information and Control*, Vol. 8, pp. 338-353.